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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 December 2011 (08.12.2011)

PCT

(10) International Publication Number
WO 2011/150855 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *A61K 31/7105* (2006.01)
C12N 15/11 (2006.01) *A61P 35/00* (2006.01)

(21) International Application Number:
PCT/CN2011/075132

(22) International Filing Date:
2 June 2011 (02.06.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PCT/CN2010/073557 4 June 2010 (04.06.2010) CN

(71) Applicant *(for all designated States except US)*: **FUDAN UNIVERSITY** [CN/CN]; 220 Handan Rd. Yangpu District, Shanghai 200433 (CN).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **WU, Ying** [CN/CN]; 21/F Kerry Everbright Building, 218 Tian Mu Xi Road, Shanghai 200070 (CN). **ZHU, Hongguang** [CN/CN]; 21/F Kerry Everbright Building, 218 Tian Mu Xi Road, Shanghai 200070 (CN). **WANG, Shuyang** [CN/CN]; 21/F Kerry Everbright Building, 218 Tian Mu Xi Road, Shanghai 200070 (CN). **NA, Yima** [CN/CN]; 21/F Kerry Everbright Building, 218 Tian Mu Xi Road, Shanghai 200070 (CN). **LI, Jian** [CN/CN]; 21/F Kerry Everbright Building, 218 Tian Mu Xi Road, Shanghai 200070 (CN).

(74) Agent: **NTD UNIVATION INTELLECTUAL PROPERTY AGENCY LTD.**; Room 1802, 18th Floor, Block A, Investment Plaza, 27 Jinrongdajie, Xicheng District, Beijing 100033 (CN).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, 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, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

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

Published:

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

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

WO 2011/150855 A1

(54) **Title**: MICRO-RNA BIOMARKERS AND METHODS FOR DIAGNOSIS OF EARLY COLORECTAL CARCINOMA AND HIGH-GRADE ADENOMA

(57) **Abstract**: The invention provides diagnostic kits for identifying one or more mammalian target cells exhibiting early colorectal carcinoma or high-grade adenoma. The kits comprise a plurality of nucleic acid molecules encoding microRNA sequences, wherein the nucleic acid molecules are differentially expressed in target cells and control cells, and the nucleic acid molecules together represent a nucleic acid expression biomarker that is indicative for the presence of early colorectal carcinoma or high-grade adenoma. The invention further provides methods for identifying one or more mammalian target cells exhibiting early colorectal carcinoma or high-grade adenoma by using the nucleic acid molecules. Methods and pharmaceutical compositions for preventing or treating early colorectal carcinoma or high-grade adenoma are also disclosed.

MICRO-RNA BIOMARKERS AND METHODS FOR DIAGNOSIS OF EARLY COLORECTAL CARCINOMA AND HIGH-GRADE ADENOMA

FIELD OF THE INVENTION

The present invention relates to validated miRNA biomarkers and corresponding methods for reliably diagnosis of early colorectal carcinoma or high-grade adenoma in the colorectal surgical and biopsy tissues. Particularly of discriminating early colorectal carcinoma from high-grade adenoma. The early colorectal carcinoma is Dukes' A carcinoma. The high-grade adenoma is high-grade intraepithelial neoplasm.

BACKGROUND OF THE INVENTION

Most cancers are epithelial in origin and arise through a stepwise progression from normal cells, through dysplasia, into malignant cells that invade surrounding tissues and have metastatic potential. Colorectal cancer (CRC; also referred to as colon cancer or large bowel cancer) is one prominent type of cancer undergoing such tumor progression.

CRC includes cancerous growth in the colon, rectum and appendix. Colorectal cancer (CRC) is the most significant human cancer with an incidence of about 1.067.000 new cases worldwide in 2009. It is the third most common cancer and the fourth leading cause of cancer deaths in the world (reviewed, e.g., in Gryfe, R. et al. (1997) *Curr. Probl. Cancer* **21**, 233-300; Petersen, G.M. et al. (1999) *Cancer* **86**, 2540-2550). CRC is curable if diagnosed at an early stage of development. At this early stage, most patients have no phenotypic symptoms of the disease. Early detection can markedly improve chances of long-term survival.

Initially, CRC is characterized by the occurrence of a hyper-proliferative (dysplastic) epithelium in the colon, which first turns into inflammatory adenomatous polyps, then into adenomas, which are intraepithelial neoplasms (i.e. benign tumors) in the inner lining of the colon or rectum. Usually, only a small subset of the adenomas formed (occurring with an incidence of 60-70% by age 60) progress into malignant adenocarcinomas. More than 95% of the cases of CRC are manifested as adenocarcinomas (Muto, T. et al. (1975) *Cancer* **36**, 2251-2270; Fearon, E.R. and Vogelstein, B. (1990) *Cell* **61**, 759-767).

World Health Organization Classification of Tumors declared that the morphology of high-grade adenoma is similar to that of carcinoma, while the only difference between these two lesions is invasion into the stroma or not (Stanley R et al. (2003) *IARC Press* 103-142). Currently, colonoscopy biopsy remains the golden standard for diagnosis of colorectal cancer. Often, pathologists find high-grade adenoma in the colonoscopy biopsy specimens. It is very hard to judge whether the lesion is invasive because mostly a biopsy cannot be beyond the muscularis mucosae. Pathologists cannot diagnose carcinoma without seeing the invasion. This is very difficult for a surgeon to make a resection decision before the operation. Therefore, there is a significantly unmet need for clinically validated molecular biomarkers in discriminating carcinoma from high-grade adenoma in colonoscopy biopsy tissues, Particularly of discriminating early colorectal carcinoma from high-grade adenoma.

Molecular studies have shown that the etiology of colon carcinogenesis results from an accumulation of multiple epigenetic and genetic alterations including *inter alia* activating mutations of the *K-ras* proto-oncogene, inactivating mutations of *APC* and *p53* tumor suppressor genes and DNA repair genes (cf., e.g., Forrester, K. et al. (1987) *Nature* **327**, 298-303; Baker, S.J. et al. (1989) *Science* **244**, 217-221).

Genomic instability is another crucial step in progression from adenomas to adenocarcinomas and occurs in two ways in CRC (Lengauer, C. et al. (1997) *Nature* **386**, 623-627). DNA mismatch repair deficiency leading to microsatellite instability, explains only about 15% of the cases of adenoma to carcinoma progression (Umar, A. et al. (2004) *J. Natl. Cancer Inst.* **96**, 261-268; di Pietro, M. et al. (2005) *Gastroenterology* **129**, 1047-1059). In the other 85%, genomic instability occurs at the chromosomal level (CIN), giving rise to aneuploidy. Chromosomal aberrations frequently reported in CRC are 7pq, 8q, 13q, and 20q gains and 4pq, 5q, 8p, 15q, 17p, and 18q losses (Douglas, E.J. et al. (2004) *Cancer Res.* **64**, 4817-4825).

However, no specific molecular markers have been identified so far that allow for a reliable diagnosis of CRC, preferably CRC manifested as an adenocarcinoma, and/or the progression of a benign adenoma into such a malignant tumor, even though cDNA microarray analyses revealed a set of differentially expressed genes apparently involved in the development of CRC (Kitahara, O. et al. (2001) *Cancer Res.* **61**, 3544-3549).

The identification of such molecular markers and development of the matched clinical tests would be of utmost clinical importance, particularly if these markers enable a diagnosis at an early stage of tumor progression in order to allow early stage treatment of carcinomas while avoiding unnecessary surgical intervention. Ideally, such markers should enable the identification of a carcinoma at a stage where the presence of malignant cells is not yet detectable by microscopic analysis of biopsy or resection materials.

Many diagnostic assays are also hampered by the fact that they are typically based on the analysis of only a single molecular marker, which might affect detection reliability and/or accuracy. In addition, a single marker normally does not enable detailed predictions concerning latency stages, tumor progression, and the like. Thus, there is still a continuing need for the identification of alternative molecular markers and assay formats overcoming these limitations.

One approach to address this issue might be based on small regulatory RNA molecules, in particular on microRNAs (miRNAs) which, constitute an evolutionary conserved class of endogenously expressed small non-coding RNAs of 20-25 nucleotides (nt) in size that can mediate the expression of target mRNAs and thus – since their discovery about ten years ago – have been implicated with critical functions in cellular development, differentiation, proliferation, and apoptosis.

miRNAs are produced from primary transcripts that are processed to stem-loop structured precursors (pre-miRNAs) by the RNase III Drosha. After transport to the cytoplasm, another RNase III termed Dicer cleaves of the loop of the pre-miRNA hairpin to form a short double-stranded (ds) RNA, one strand of which is incorporated as mature miRNA into a miRNA-protein (miRNP). The miRNA guides the miRNPs to their target mRNAs where they exert their function (reviewed, e.g. in Bartel, D.P. (2004) *Cell* **23**, 281-292; He, L. and Hannon, G.J. (2004) *Nat. Rev. Genet.* **5**, 522-531).

Depending on the degree of complementarity between the miRNA and its target, miRNAs can guide different regulatory processes. Target mRNAs that are highly complementary to miRNAs are specifically cleaved by mechanisms identical to RNA interference (RNAi). Thus, in such scenario, the miRNAs function as short interfering RNAs (siRNAs). Target mRNAs with less complementarity to miRNAs are either directed to cellular degradation pathways or are translationally repressed without affecting the mRNA level. However, the mechanism of how miRNAs repress

translation of their target mRNAs is still a matter of controversy.

Emerging data available indicate that dysregulation of miRNA expression may *inter alia* be associated with the development and/or progression of certain types of cancer. For example, two miRNAs, *miR-15* and *miR-16-1*, were shown to map to a genetic locus that is deleted in chronic lymphatic leukemia (CLL) and it was found that in about 70% of the CLL patients, both miRNA genes are deleted or down-regulated. Furthermore, down-regulation of *miR-143* and *miR-145* was observed in colorectal neoplasia, whereas expression of the miRNA *let-7* is frequently reduced in lung cancers (Michael, M.Z. et al. (2003) *Mol. Cancer Res.* **1**, 882-891; Mayr, C. et al. (2007) *Science* **315**, 1576-1579).

In fact, it has been speculated based on cancer-associated alterations in miRNA expression and the observation that miRNAs are frequently located at genomic regions involved in cancers that miRNAs may act both as tumor suppressors and as oncogenes (reviewed, e.g., in Esquela-Kerscher, A. and Slack, F.J (2006) *Nat. Rev. Cancer* **6**, 259-269; Calin, G.A. and Croce, C.M. (2007) *J. Clin. Invest.* **117**, 2059-2066; Blenkiron, C. and Miska, E.A. (2007) *Hum. Mol. Genet.* **16**, R106-R113).

Thus, there still remains a need for (a set of) diagnostic markers, particularly in form of a "expression biomarker " or a "molecular footprint", that enable the rapid, reliable and cost-saving identification of cells exhibiting colorectal carcinoma or adenoma. In addition, there is also a continuing need for corresponding methods for the identification of target cells displaying such a tumor phenotype.

OBJECT AND SUMMARY OF THE INVENTION

It is an objective of the present invention to provide validated miRNA biomarkers and corresponding methods for reliably diagnosis of early colorectal carcinoma or high-grade adenoma in colorectal surgical and biopsy tissues. Particularly of discriminating early colorectal carcinoma from high-grade adenoma. Specifically, A early colorectal carcinoma or high-grade adenoma by determining a plurality of nucleic acid molecules, each nucleic acid molecule encoding a microRNA (miRNA) sequence, wherein one or more of the plurality of nucleic acid molecules are differentially expressed in the target cells analyzed as compared to control cells, and wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic

acid expression biomarker that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

More specifically, it is an object of the invention to provide validated miRNA biomarkers for reliably diagnosis of early colorectal carcinoma or high-grade adenoma. Particularly of discriminating early Dukes' A carcinoma from high-grade adenoma. Furthermore, it is an object of the invention to provide corresponding methods for reliably diagnosis of early colorectal carcinoma or high-grade adenoma. Particularly of discriminating early colorectal carcinoma from high-grade adenoma.

The early colorectal carcinoma is manifested as a Dukes' A adenocarcinoma. The high-grade adenoma manifested is manifested as a high-grade intraepithelial neoplasm.

These objectives as well as others, which will become apparent from the ensuing description, are attained by the subject matter of the independent claims. Some of the preferred embodiments of the present invention are defined by the subject matter of the dependent claims.

In a first aspect, the present invention relates to a diagnostic kit of miRNA biomarkers for identifying one or more mammalian target cells exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprising a plurality of nucleic acid molecules, each nucleic acid molecule encoding a microRNA sequence, wherein one or more of the plurality of nucleic acid molecules are differentially expressed in the target cells and in one or more control cells, and wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic acid expression signature that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

The early colorectal carcinoma is manifested as a Dukes' A adenocarcinoma. The high-grade adenoma is manifested as a high-grade intraepithelial neoplasm.

In preferred embodiments of the invention, the nucleic acid expression biomarkers comprises at least one nucleic acid molecule encoding a microRNA sequence whose expression is up-regulated in the one or more target cells compared to the one or more control cells and at least one nucleic acid molecule encoding a microRNA sequence whose expression is down-regulated in the one or more target cells compared to the one or more control cells.

The nucleic acid expression biomarker, as defined herein, may comprise at least three nucleic acid molecules, preferably at least six nucleic acid molecules.

In preferred embodiments, the nucleic acid expression biomarker comprises any one or more of the nucleic acid molecules encoding hsa-miR-375, hsa-miR-92a and hsa-miR-99a.

Particularly preferably, the expression of any one or more of the nucleic acid molecules encoding hsa-miR-92a and hsa-miR-99a are up-regulated and the expression of hsa-miR-375 is down-regulated in the one or more target cells compared to the one or more control cells.

In further preferred embodiments, the nucleic acid expression biomarker comprises any one or more nucleic acid combinations encoding hsa-miR-92a/hsa-miR-375 and hsa-miR-99a/hsa-miR-375.

Particularly preferably, the expression of any one or more of the nucleic acid molecule combinations encoding hsa-miR-92a/hsa-miR-375 and hsa-miR-99a/hsa-miR-375 are up-regulated compared to the one or more control cells.

In more preferred embodiments, the nucleic acid expression biomarkers comprises nucleic acid molecules encoding hsa-miR-125b, hsa-miR-375, hsa-miR-424, hsa-miR-92a, hsa-miR-99a and hsa-miR-7.

Particularly preferably, the expression of any one or more of the nucleic acid molecules encoding hsa-miR-125b, hsa-miR-424, hsa-miR-92a, hsa-miR-99a and hsa-miR-7 are up-regulated and the expression of hsa-miR-375 is down-regulated in the one or more target cells compared to the one or more control cells.

In further preferred embodiments, the nucleic acid expression biomarker comprises any one or more nucleic acid combinations encoding hsa-miR-125b/hsa-miR-375, hsa-miR-424/hsa-miR-375, hsa-miR-92a/hsa-miR-375, hsa-miR-99a/hsa-miR-375 and hsa-miR-7/hsa-miR-375.

Particularly preferably, the expression of any one or more of the nucleic acid molecule combinations encoding hsa-miR-125b/hsa-miR-375, hsa-miR-424/hsa-miR-375, hsa-miR-92a/hsa-miR-375, hsa-miR-99a/hsa-miR-375 and hsa-miR-7/hsa-miR-375 are up-regulated compared to the one or more control cells.

In a second aspect, the present invention relates to a method for identifying one or more mammalian target cells exhibiting early colorectal carcinoma or high-grade adenoma, the method comprising: (a) collecting a biopsy or surgical tissue

from a patient; (b) preparing tissue section on a slide; (c) hybridizing at least one nucleic acid molecule biomarker encoding a microRNA sequence to the section on the slide, (d) quantifying the miRNA expression under microscope or by digital pathology solution; (e) determining in the one or more target cells the expression levels of a plurality of nucleic acid molecules, each nucleic acid molecule encoding a microRNA sequence; (f) determining the expression levels of the plurality of nucleic acid molecules in one or more control cells; and (g) identifying from the plurality of nucleic acid molecules one or more nucleic acid molecules that are differentially expressed in the target and control cells by comparing the respective expression levels obtained in steps (e) and (f), wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic acid expression biomarker, as defined herein, that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

The early colorectal carcinoma is Dukes' A carcinoma. The high-grade adenoma manifested is high-grade intraepithelial neoplasm.

Particularly preferably, the method is manifested as *in situ* hybridization.

For quantitative determination, 6 validated miRNA biomarkers are used: hsa-miR-125b, hsa-miR-375, hsa-miR-424, hsa-miR-92a, hsa-miR-99a and hsa-miR-7.

For normalizing the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-423-5p may be preferably used, which is stably expressed in colorectal tissues. For negative control of the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-122 may be preferably used, which does not expressed in colorectal tissues.

In a third aspect, the present invention relates to a method for preventing or treating colorectal cancer, the method comprising: (a) identifying a nucleic acid expression biomarker by using a method, as defined herein; and (b) modifying the expression of one or more nucleic acid molecules encoding a microRNA sequence that is/are comprised in the nucleic acid expression biomarker in such way that the expression of a nucleic acid molecule whose expression is up-regulated is down-regulated and the expression of a nucleic acid molecule whose expression is down-regulated is up-regulated.

In a fourth aspect, the present invention relates to a pharmaceutical composition for the prevention and/or treatment of colorectal cancer, the composition comprising one or more nucleic acid molecules, each nucleic acid molecule encoding a sequence that is at least partially complementary to a microRNA sequence encoded by a nucleic acid molecule whose expression is up-regulated from colorectal cancer patients, as defined herein, and/or that corresponds to a microRNA sequence encoded by a nucleic acid molecule whose expression is down-regulated from colorectal cancer patients, as defined herein.

In a final aspect, the present invention relates to the use of said pharmaceutical composition for the manufacture of a medicament for the prevention and/or treatment of colorectal cancer cancer.

Other embodiments of the present invention will become apparent from the detailed description hereinafter.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts a flow chart schematically illustrating the essential method steps in the fifth aspect for determining an expression biomarker according to the present invention for identifying one or more target cells exhibiting colorectal carcinoma or adenoma. Particularly of discriminating early colorectal carcinoma from high-grade adenoma using *in situ* hybridization method.

Figure 2 illustrates the human miRNAs comprised in particularly preferred expression biomarkers in the first aspect according to the present invention for identifying one or more target cells exhibiting Dukes' A carcinoma or high-grade adenoma in frozen surgical tissues. Also indicates the expression levels and accuracy of these miRNAs in the patients with Dukes' A carcinoma as compared to high-grade adenoma cells (i.e. an up-regulation or a down-regulation). The data indicate that early colorectal carcinoma can be reliably discriminated from high-grade adenoma in frozen surgical tissues.

Figure 3A illustrates the human miRNAs comprised in particularly preferred expression biomarkers in the first aspect according to the present invention for identifying one or more target cells exhibiting Dukes' A carcinoma or high-grade adenoma in FFPE surgical tissues. Also indicates the expression levels and accuracy of these miRNAs in the patients with Dukes' A carcinoma as compared to high-grade

adenoma cells (i.e. an up-regulation or a down-regulation). The data indicate that early colorectal carcinoma can be reliably discriminated from high-grade adenoma in FFPE surgical tissues.

Figure 3B depicts the ROC curve analysis of two combinations (hsa-miR-92a/hsa-miR-375 and hsa-miR-99a/hsa-miR-375) in the first aspect according to the present invention for identifying one or more target cells exhibiting Dukes' A carcinoma or high-grade adenoma in FFPE surgical tissues.. The data indicate that early colorectal carcinoma can be reliably discriminated from high-grade adenoma in FFPE surgical tissues.

Figure 4 illustrates the human miRNAs comprised in particularly preferred expression biomarkers in the first aspect according to the present invention for identifying one or more target cells exhibiting early colorectal carcinoma or high-grade adenoma in FFPE biopsy tissues from colonscope. Also indicates the expression levels and accuracy of these miRNAs in the patients with colorectal carcinoma as compared to high-grade adenoma cells (i.e. an up-regulation or a down-regulation). The data indicate tha colorectal carcinoma can be reliably discriminated from high-grade adenoma in FFPE biopsy tissues from colonscope.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based on the unexpected finding that cells exhibiting early colorectal carcinoma or high-grade adenoma, preferably an adenocarcinoma can be reliably identified based on a particular expression biomarker both with high accuracy and sensitivity, wherein the expression biomarker as defined herein typically comprises both up- and down-regulated human miRNAs. More specifically, said miRNA expression biomarker – by analyzing the respective individual miRNA expression level(s) and/or the miRNA expression pattern of the miRNA combinations – allows the diagnosis of early colorectal carcinoma or high-grade adenoma, particularly of discriminating early colorectal carcinoma from high-grade adenoma.

The early colorectal carcinoma is Dukes' A carcinoma. The high-grade adenoma manifested is high-grade intraepithelial neoplasm.

The present invention illustratively described in the following may suitably be practiced in the absence of any element or elements, limitation or limitations,

not specifically disclosed herein. The present invention will be described with respect to particular embodiments and with reference to certain drawings but the invention is not limited thereto but only by the claims. The drawings described are only schematic and are to be considered non-limiting.

Where the term "comprising" is used in the present description and claims, it does not exclude other elements or steps. For the purposes of the present invention, the term "consisting of" is considered to be a preferred embodiment of the term "comprising of". If hereinafter a group is defined to comprise at least a certain number of embodiments, this is also to be understood to disclose a group, which preferably consists only of these embodiments.

Where an indefinite or definite article is used when referring to a singular noun e.g. "a" or "an", "the", this includes a plural of that noun unless specifically stated otherwise.

The term "about" in the context of the present invention denotes an interval of accuracy that the person skilled in the art will understand to still ensure the technical effect of the feature in question. The term typically indicates deviation from the indicated numerical value of ± 10%, and preferably ± 5%.

Furthermore, the terms first, second, third, (a), (b), (c), and the like in the description and in the claims, are used for distinguishing between similar elements and not necessarily for describing a sequential or chronological order. It is to be understood that the terms so used are interchangeable under appropriate circumstances and that the embodiments of the invention described herein are capable of operation in other sequences than described or illustrated herein.

Further definitions of term will be given in the following in the context of which the terms are used.

The following terms or definitions are provided solely to aid in the understanding of the invention. These definitions should not be construed to have a scope less than understood by a person of ordinary skill in the art.

It is an objective of the present invention to provide validated miRNA biomarkers and corresponding methods for reliably diagnosis of early colorectal carcinoma or high-grade adenoma in colorectal surgical and biopsy tissues. Particularly of discriminating early colorectal carcinoma from high-grade adenoma. Specifically, A early colorectal carcinoma or high-grade adenoma by determining a plurality of nucleic

acid molecules, each nucleic acid molecule encoding a microRNA (miRNA) sequence, wherein one or more of the plurality of nucleic acid molecules are differentially expressed in the target cells analyzed as compared to control cells, and wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic acid expression biomarker that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

The term "colorectal", as used herein, relates to the colon, the rectum and/or the appendix, i.e. the complete large intestine.

The term "cancer" (also referred to as "carcinoma"), as used herein, generally denotes any type of malignant neoplasm, that is, any morphological and/or physiological alterations (based on genetic re-programming) of target cells exhibiting or having a predisposition to develop characteristics of a carcinoma as compared to unaffected (healthy) wild-type control cells. Carcinoma is defined as invasion through the muscularis mucosae. Examples of such alterations may relate *inter alia* to cell size and shape (enlargement or reduction), cell proliferation (increase in cell number), cell differentiation (change in physiological state), apoptosis (programmed cell death) or cell survival. Hence, the term "colorectal cancer" refers to cancerous growths in the colon, rectum, and appendix.

Preferably, the colorectal cancer is manifested as an adenocarcinoma.

The most common colorectal cancer (CRC) cell type is adenocarcinoma that accounts for about 95% of cases. Other types of CRC include *inter alia* lymphoma and squamous cell carcinoma.

The term "adenocarcinoma", as used herein, relates to a malignant neoplasm of epithelial cells of the colorectal mucosa. Typically, adenocarcinoma is a type of cancer that originates in glandular tissue. This tissue is part of a more general type of tissue known as epithelial tissue. Epithelial tissue includes skin, glands and a variety of other tissues lining/surrounding the cavities and organs of the body.

Embryologically, the epithelium is derived from ectoderm, endoderm and mesoderm. In order to be classified as adenocarcinoma, the cells do not necessarily need to be part of a gland, as long as they have secretory properties. Hence, adenocarcinomas are also often referred to as "glandular cancer" or "glandular carcinoma". Highly differentiated adenocarcinomas tend to resemble the glandular tissue that they are derived from, while poorly differentiated may not.

The occurrence of a hyper-proliferative epithelium in the colon is the first step in cancer progression. This dysplastic epithelium turns into adenomatous polyps, subsequently into adenomas, which are abnormal but benign neoplasms (i.e. tumors) in the inner lining of the colon or rectum. Thus, the term "adenoma or intraepithelial neoplasm ", as used herein, thus relates to a benign epithelial neoplasm. Adenomas are usually well circumscribed and can be flat or polypoid. The neoplastic cells of benign adenomas do not infiltrate or invade adjacent tissue and rarely metastasize. The term "low-grade adenoma or low-grade intraepithelial neoplasm " is understood as equivalent to "non-progressed adenoma". The term "high-grade adenoma or high-grade intraepithelial neoplasm " is understood as equivalent to "high-risk adenoma" for the transformation of colorectal cancer. Malignant adenocarcinomas, however, invade other tissues and often metastasize given enough time to do so. Malignant cells are often characterized by progressive and uncontrolled growth. They can spread locally or through the blood stream and lymphatic system to other parts of the body. Particularly, hepatic metastases (i.e. metastases in the liver) are commonly found to be associated with adenocarcinomas. The occurrence of such metastases may be considered a late stage (or even a post-cancerous stage) of colorectal cancer.

The terms "progressed adenoma", as used herein, refer to an adenoma that harbors a focus of a cancer. This is also called a "malignant polyp". Colorectal adenomas are common in the elderly population, but only a small proportion of these pre-malignant tumors (estimated approximately 5%) progresses to malignant tumors. Such malignant tumors are herein referred to as (colorectal) "adenocarcinomas".

Adenocarcinomas may be classified according to the Dukes system (Dukes, C.E. (1932) *J. Pathol. Bacteriol.* **35**, 323-325), which identifies the following stages: Dukes A – a tumour confined to the intestinal wall; Dukes B – a tumor invading through the intestinal wall; Dukes C – a tumor also involving the lymph node(s); and Dukes D – a tumor with distant metastasis.

The mammalian target cells employed in the present invention may be of human or non-human origin. However, the invention is typically performed with human cells. The term "one or more cells", as used herein, is to be understood not only to include individual cells but also tissues, organs, and organisms. The term "target cell", as used herein, refers to a cell being at least supposed to exhibit or to have a predisposition to develop colorectal cancer, whereas the term "control cell" typically

denotes a (healthy) wild-type cell not having characteristics of such a cancerous phenotype. However, in some applications, for example, when comparing cells exhibiting different cancerous or pre-cancerous states, the cells having the less severe disease characteristics are typically considered the "control cells".

Typically, the target and control cells used are derived from biological samples collected from the subjects to be diagnosed for the presence of colorectal cancer or adenoma. Furthermore, in order to corroborate the data obtained "comparative samples" may also be collected from subjects having a given known disease state. The biological samples may include body tissues and fluids, such as blood, sputum, and urine. Furthermore, the biological sample may contain a cell extract derived from or a cell population including an epithelial cell, preferably a cancerous epithelial cell or an epithelial cell derived from tissue suspected to be cancerous. Even more preferably the biological sample comprises a cell population derived from a glandular tissue. Furthermore, the cell may be purified from the obtained body tissues and fluids if necessary, and then used as the biological sample. According to the present invention, the expression level of the nucleic acid markers of the present invention is determined in the subject-derived biological sample(s).

The sample used for detection in the *in vitro* methods of the present invention should generally be collected in a clinically acceptable manner, preferably in a way that nucleic acids (in particular RNA) are preserved. The samples to be analyzed are typically colorectal biopsies or resections. A biopsy or resection may contain a majority of normal cells and only a minority of adenocarcinoma cells or adenoma cells. To increase the specificity and sensitivity of the detection, a probe encoding a microRNA sequence may be hybridized to the tissue section on the slide and quantified under microscope or by digital pathology system. Even if the total number of carcinoma cells or adenoma cells in the biopsy or resection is limited, the carcinoma cells or adenoma cells may be reliably detected by the combination of morphology and miRNA quantification.

The term "microRNA" (or "miRNA"), as used herein, is given its ordinary meaning in the art (reviewed, e.g. in Bartel, D.P. (2004) *Cell* **23**, 281-292; He, L. and Hannon, G.J. (2004) *Nat. Rev. Genet.* **5**, 522-531). Accordingly, a "microRNA" denotes a RNA molecule derived from a genomic locus that is processed from

transcripts that can form local RNA precursor miRNA structures. The mature miRNA is usually 20, 21, 22, 23, 24, or 25 nucleotides in length, although other numbers of nucleotides may be present as well, for example 18, 19, 26 or 27 nucleotides.

The miRNA encoding sequence has the potential to pair with flanking genomic sequences, placing the mature miRNA within an imperfect RNA duplex (herein also referred to as stem-loop or hairpin structure or as pre-miRNA), which serves as an intermediate for miRNA processing from a longer precursor transcript. This processing typically occurs through the consecutive action of two specific endonucleases termed Drosha and Dicer, respectively. Drosha generates from the primary transcript (herein also denoted "pri-miRNA") a miRNA precursor (herein also denoted "pre-miRNA") that typically folds into a hairpin or stem-loop structure. From this miRNA precursor a miRNA duplex is excised by means of Dicer that comprises the mature miRNA at one arm of the hairpin or stem-loop structure and a similar-sized segment (commonly referred to miRNA*) at the other arm. The miRNA is then guided to its target mRNA to exert its function, whereas the miRNA* is degraded. In addition, miRNAs are typically derived from a segment of the genome that is distinct from predicted protein-coding regions.

The term "miRNA precursor" (or "precursor miRNA" or "pre-miRNA"), as used herein, refers to the portion of a miRNA primary transcript from which the mature miRNA is processed. Typically, the pre-miRNA folds into a stable hairpin (i.e. a duplex) or a stem-loop structure. The hairpin structures typically range from 50 to 80 nucleotides in length, preferably from 60 to 70 nucleotides (counting the miRNA residues, those pairing to the miRNA, and any intervening segment(s) but excluding more distal sequences).

The term "nucleic acid molecule encoding a microRNA sequence", as used herein, denotes any nucleic acid molecule coding for a microRNA (miRNA). Thus, the term does not only refer to mature miRNAs but also to the respective precursor miRNAs and primary miRNA transcripts as defined above. Furthermore, the present invention is not restricted to RNA molecules but also includes corresponding DNA molecules encoding a microRNA, e.g. DNA molecules generated by reverse transcribing a miRNA sequence. A nucleic acid molecule encoding a microRNA sequence according to the invention typically encodes a single miRNA sequence (i.e. an individual miRNA). However, it is also possible that such nucleic acid molecule

encodes two or more miRNA sequences (i.e. two or more miRNAs), for example a transcriptional unit comprising two or more miRNA sequences under the control of common regulatory sequences such as a promoter or a transcriptional terminator.

The term "nucleic acid molecule encoding a microRNA sequence", as used herein, is also to be understood to include "sense nucleic acid molecules" (i.e. molecules whose nucleic acid sequence (5'→ 3') matches or corresponds to the encoded miRNA (5'→ 3') sequence) and "anti-sense nucleic acid molecules" (i.e. molecules whose nucleic acid sequence is complementary to the encoded miRNA (5'→ 3') sequence or, in other words, matches the reverse complement (3'→ 5') of the encoded miRNA sequence). The term "complementary", as used herein, refers to the capability of an "anti-sense" nucleic acid molecule sequence of forming base pairs, preferably Watson-Crick base pairs, with the corresponding "sense" nucleic acid molecule sequence (having a sequence complementary to the anti-sense sequence).

Within the scope of the present invention, two nucleic acid molecules (i.e. the "sense" and the "anti-sense" molecule) may be perfectly complementary, that is, they do not contain any base mismatches and/or additional or missing nucleotides. Alternatively, the two molecules comprise one or more base mismatches or differ in their total numbers of nucleotides (due to additions or deletions). Preferably, the "complementary" nucleic acid molecule comprises at least ten contiguous nucleotides showing perfect complementarity with a sequence comprised in corresponding "sense" nucleic acid molecule.

Accordingly, the plurality of nucleic acid molecules encoding a miRNA sequence that are comprised in a diagnostic kit of the present invention may include one or more "sense nucleic acid molecules" and/or one or more "anti-sense nucleic acid molecules". In case, the diagnostic kit includes one or more "sense nucleic acid molecules" (i.e. the miRNA sequences as such), said molecules are to be considered to constitute the totality or at least a subset of differentially expressed miRNAs (i.e. molecular markers) being indicative for the presence of or the disposition to develop a particular condition, here colorectal cancer, preferably colorectal cancer manifested as an adenocarcinoma. On the other hand, in case a diagnostic kit includes one or more "anti-sense nucleic acid molecules" (i.e. sequences complementary to the miRNA sequences), said molecules may comprise *inter alia* probe molecules (for performing hybridization assays) and/or oligonucleotide primers (e.g., for reverse transcription or

PCR applications) that are suitable for detecting and/or quantifying one or more particular (complementary) miRNA sequences in a given sample.

A plurality of nucleic acid molecules as defined within the present invention may comprise at least two, at least ten, at least 50, at least 100, at least 200, at least 500, at least 1.000, at least 10.000 or at least 100.000 nucleic acid molecules, each molecule encoding a miRNA sequence.

The term "differentially expressed", as used herein, denotes an altered expression level of a particular miRNA in the target cells as compared to the control cells, which may be an up-regulation (i.e. an increased miRNA concentration in the target cells) or a down-regulation (i.e. a reduced or abolished miRNA concentration in the target cells). In other words, the nucleic acid molecule is activated to a higher or lower level in the target cells than in the control cells.

Within the scope of the present invention, a nucleic acid molecule is to considered differentially expressed if the respective expression levels of this nucleic acid molecule in target cells and control cells typically differ by at least 5% or at least 10%, preferably by at least 20% or at least 25%, and most preferably by at least 30% or at least 50%. Thus, the latter values correspond to an at least 1.3-fold or at least 1.5-fold up-regulation of the expression level of a given nucleic acid molecule in the target cells compared to the wild-type control cells or *vice versa* an at least 0.7-fold or at least 0.5-fold down-regulation of the expression level in the target cells, respectively.

The term "expression level", as used herein, refers to extent to which a particular miRNA sequence is transcribed from its genomic locus, that is, the concentration of a miRNA in the one or more cells to be analyzed.

As outlined above, the term "control cell" typically denotes a (healthy) wild-type cell not having characteristics of a CRC phenotype. However, in some applications, for example, when comparing cells exhibiting different cancerous or pre-cancerous states, the cells having the less severe disease characteristics are typically considered the "control cells".

The determining of expression levels typically follows established standard procedures well known in the art (cf., for example, Sambrook, J. et al. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; Ausubel, F.M. et al. (2001) *Current Protocols in Molecular Biology*. Wiley & Sons, Hoboken, NJ). Determination may occur at the RNA

level, for example by Northern blot analysis using miRNA-specific probes, or at the DNA level following reverse transcription (and cloning) of the RNA population, for example by quantitative PCR or real-time PCR techniques. The term "determining", as used herein, includes the analysis of any nucleic acid molecules encoding a microRNA sequence as described above. However, due to the short half-life of pri-miRNAs and pre-mRNAs typically the concentration of only the mature miRNA is measured.

In specific embodiments, the standard value of the expression levels obtained in several independent measurements of a given sample (for example, two, three, five or ten measurements) and/or several measurements within a population of target cells or control cells is used for analysis. The standard value may be obtained by any method known in the art. For example, a range of mean ± 2 SD (standard deviation) or mean ± 3 SD may be used as standard value.

The difference between the expression levels obtained for one or more target cells and one or more control cells may be normalized to the expression level of further control nucleic acids, e.g. endogenous controls whose expression levels are known not to differ depending on the disease states of the cell. Exemplary endogenous controls include small nuclear RNAs and housekeeping genes.

In preferred embodiments, the control nucleic acid for normalizing the expression levels obtained is another miRNA known to be stably expressed during the various non-cancerous and (pre-)cancerous states of the cell.

However, instead of determining in any experiment the expression levels for one or more control cells it may also be possible to define based on experimental evidence and/or prior art data on or more cut-off values for a particular cell phenotype (i.e. a disease state). In such scenario, the respective expression levels for the one or more target cells can be determined by using a stably expressed control miRNA for normalization. If the "normalized" expression levels calculated are higher than the respective cutoff value defined, then this finding would be indicative for an up-regulation of gene expression. *Vice versa*, if the "normalized" expression levels calculated are lower than the respective cutoff value defined, then this finding would be indicative for a down-regulation of gene expression.

In the context of the present invention, the term "identifying one or more mammalian target cells exhibiting colorectal carcinoma or adenoma " is intended to also encompass predictions and likelihood analysis (in the sense of "diagnosing"). The

compositions and methods disclosed herein are intended to be used clinically in making decisions concerning treatment modalities, including therapeutic intervention, diagnostic criteria such as disease stages, and disease monitoring and surveillance for the disease. According to the present invention, an intermediate result for examining the condition of a subject may be provided. Such intermediate result may be combined with additional information to assist a doctor, nurse, or other practitioner to diagnose that a subject suffers from the disease. Alternatively, the present invention may be used to detect cancerous cells in a subject-derived tissue, and provide a doctor with useful information to diagnose that the subject suffers from the disease.

Within the present invention, one or more differentially expressed nucleic acid molecules identified together represent a nucleic acid expression signature that is indicative for the presence of or the predisposition to develop colorectal cancer in the target cells. The term "expression signature", as used herein, denotes a set of nucleic acid molecules (e.g., miRNAs), wherein the expression level of the individual nucleic acid molecules differs between the (cancerous) target cells and the (non-cancerous) control cells. Herein, a nucleic acid expression signature is also referred to as a set of markers and represents a minimum number of (different) nucleic acid molecules, each encoding a miRNA sequence that is capable for identifying a phenotypic state of a target cell.

In a first aspect, the present invention relates to a diagnostic kit of miRNA biomarkers for identifying one or more mammalian target cells exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprising a plurality of nucleic acid molecules, each nucleic acid molecule encoding a microRNA sequence, wherein one or more of the plurality of nucleic acid molecules are differentially expressed in the target cells and in one or more control cells, and wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic acid expression signature that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

The early colorectal carcinoma is Dukes' A carcinoma. The high-grade adenoma is high-grade intraepithelial neoplasm.

In preferred embodiments of the invention, the nucleic acid expression biomarkers comprises at least one nucleic acid molecule encoding a microRNA sequence whose expression is up-regulated in the one or more target cells compared to the one or more control cells and at least one nucleic acid molecule encoding a

microRNA sequence whose expression is down-regulated in the one or more target cells compared to the one or more control cells.

Typically, the nucleic acid molecules comprised in the nucleic acid expression signature are human sequences (hereinafter designated "hsa" (Homo sapiens)).

The nucleic acid expression biomarker, as defined herein, may comprise at least three nucleic acid molecules, preferably at least six nucleic acid molecules.

In preferred embodiments, the nucleic acid expression biomarker comprises any one or more of the nucleic acid molecules encoding hsa-miR-375 (SEQ ID NO:2), hsa-miR-92a (SEQ ID NO:1) and hsa-miR-99a (SEQ ID NO:5).

For normalizing the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-423-5p (SEQ ID NO:7) may be preferably used, which is stably expressed in colorectal tissues. For negative control of the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-122 (SEQ ID NO:8) may be preferably used, which does not expressed in colorectal tissues.

The nucleic acid sequences of the above-referenced miRNAs are listed in Table 1

TABLE 1

miRNA	**Sequence (5' → 3')**
Biomarker	
hsa-miR-92a	uauugcacuugucccggccugu
hsa-miR-375	uuuguucguucggcucgcguga
hsa-miR-99a	aacccguagauccgaucuugug
Control	
hsa-miR-423-5p	ugaggggcagagagcgagacuuu
hsa-miR-122	uggagugugacaaugguguuug

All miRNA sequences disclosed herein have been deposited in the miRBase database (http://microrna.sanger.ac.uk/; see also Griffiths-Jones S. et al. (2008) *Nucl. Acids Res.* **36**, D154-D158).

Particularly preferably, the expression of any one or more of the nucleic acid molecules encoding hsa-miR-92a and hsa-miR-99a are up-regulated and the expression of hsa-miR-375 is down-regulated in the one or more target cells compared to the one or more control cells.

The terms "any one or more of the plurality of nucleic acid molecules" or "any one or more of the plurality of nucleic acid molecules" as used herein, may relate to any subgroup of the plurality of nucleic acid molecules, e.g., any one, any two, any three, any four, any five, any six, any seven, any eight, any nine, any ten, and so forth nucleic acid molecules, each encoding a microRNA sequence that are comprised in the nucleic acid expression biomarker, as defined herein.

In further preferred embodiments, the nucleic acid expression biomarker comprises any one or more nucleic acid combinations encoding hsa-miR-92a (SEQ ID NO:1)/hsa-miR-375 (SEQ ID NO:2) and hsa-miR-99a (SEQ ID NO:5)/hsa-miR-375 (SEQ ID NO:2).

The term “nucleic acid combinations”, as used herein, refers to the usage of at least two nucleic acid expression levels as a whole. Preferably may use the relative changes or calculate results through a formulation as a whole.

In more preferred embodiments, the nucleic acid expression biomarkers comprises nucleic acid molecules encoding hsa-miR-125b (SEQ ID NO:3), hsa-miR-375 (SEQ ID NO:2), hsa-miR-424 (SEQ ID NO:4), hsa-miR-92a (SEQ ID NO:1), hsa-miR-99a (SEQ ID NO:5) and hsa-miR-7 (SEQ ID NO:6).

For normalizing the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-423-5p (SEQ ID NO:7) may be preferably used, which is stably expressed in colorectal tissues. For negative control of the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-122 (SEQ ID NO:8) may be preferably used, which does not expressed in colorectal tissues.

The nucleic acid sequences of the above-referenced miRNAs are listed in Table 2

TABLE 2

miRNA	**Sequence (5' → 3')**
Biomarker	

hsa-miR-92a	uauugcacuugucccggccugu
hsa-miR-375	uuuguucguucggcucgcguga
hsa-miR-125b	ucccugagacccuaacuuguga
hsa-miR-424	cagcagcaauucauguuuugaa
hsa-miR-99a	aacccguagauccgaucuugug
hsa-miR-7	uggaagacuagugauuuuguugu
Control	
hsa-miR-423-5p	ugaggggcagagagcgagacuuu
hsa-miR-122	uggagugugacaaugguguuug

All miRNA sequences disclosed herein have been deposited in the miRBase database (http://microrna.sanger.ac.uk/; see also Griffiths-Jones S. et al. (2008) *Nucl. Acids Res.* **36**, D154-D158).

Particularly preferably, the expression of any one or more of the nucleic acid molecules encoding hsa-miR-125b, hsa-miR-424, hsa-miR-92a, hsa-miR-99a and hsa-miR-7 is up-regulated and the expression of hsa-miR-375 is down-regulated in the one or more target cells compared to the one or more control cells.

In further preferred embodiments, the nucleic acid expression biomarker comprises any one or more nucleic acid combinations encoding hsa-miR-125b (SEQ ID NO:3) / hsa-miR-375 (SEQ ID NO:2), hsa-miR-424 (SEQ ID NO:4) / hsa-miR-375 (SEQ ID NO:2), hsa-miR-92a (SEQ ID NO:1) / hsa-miR-375 (SEQ ID NO:2), hsa-miR-99a (SEQ ID NO:5) / hsa-miR-375 (SEQ ID NO:2) and hsa-miR-7 (SEQ ID NO:6) / hsa-miR-375 (SEQ ID NO:2).

In a second aspect, the present invention relates to a method for identifying one or more mammalian target cells exhibiting early colorectal carcinoma or high-grade adenoma, the method comprising:

(a) collecting a biopsy or surgical tissue from a patient;

(b) preparing tissue section on a slide;

(c) hybridizing at least one nucleic acid molecule encoding a microRNA sequence to the section on the slide;

(d) quantifying the miRNA expression under microscope or by digital pathology solution;

(e) determining in the one or more target cells the expression levels of a plurality of nucleic acid molecules, each nucleic acid molecule encoding a microRNA sequence;

(f) determining the expression levels of the plurality of nucleic acid molecules in one or more control cells; and

(g) identifying from the plurality of nucleic acid molecules one or more nucleic acid molecules that are differentially expressed in the target and control cells by comparing the respective expression levels obtained in steps (e) and (f), wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic acid expression biomarker, as defined herein, that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

The method of the present invention comprises the validated biomarkers and *in situ* hybridization. The method determins and comparies the expression levels of a plurality of nucleic acid molecules encoding a microRNA sequence both in one or more target cells supposed to exhibit early colorectal carcinoma or high-grade adenoma and in one or more control cells, i.e. typically wild-type cells not showing the characteristics of such a cancerous phenotype.

For quantitative determination, six validated miRNA biomarkers listed in Table 2 are used in the method: hsa-miR-125b (SEQ ID NO:3), hsa-miR-375 (SEQ ID NO:2), hsa-miR-424 (SEQ ID NO:4), hsa-miR-92a (SEQ ID NO:1), hsa-miR-99a (SEQ ID NO:5) and hsa-miR-7 (SEQ ID NO:6).

For normalizing the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-423-5p (SEQ ID NO:7) may be preferably used, which is stably expressed in colorectal tissues. For negative control of the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-122 (SEQ ID NO:8) may be preferably used, which does not expressed in colorectal tissues.

In situ hybridization techniques allow specific nucleic acid sequences to be detected in morphologically preserved chromosomes, cells or tissue sections. In combination with immunocytochemistry, *in situ* hybridization can relate microscopic topological information to gene activity at the DNA, mRNA, and protein level. There are two types of nonradioactive hybridization methods: direct and indirect. Direct

methods using fluorescein or other fluorochromes directly coupled to the nucleotide (Baumann, J. G. J. et al. ((1980) *Exp. Cell Res.* **138,** 485–490). Indirect methods using digoxigenin (detected by specific antibodies) and biotin (detected by streptavidin) (Leary, J. L et al (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4045–4049).

In a third aspect, the invention relates to a method for preventing or treating colorectal cancer, preferably manifested as an adenocarcinoma, in one or more mammalian target cells, the method comprising:

(a) identifying in one or more target cells a nucleic acid expression signature by using a method, as defined herein; and

(b) modifying in the one or more cells the expression of one or more nucleic acid molecules encoding a microRNA sequence that is/are comprised in the nucleic acid expression signature in such way that the expression of a nucleic acid molecule whose expression is up-regulated in the one or more target cells is down-regulated and the expression of a nucleic acid molecule whose expression is down-regulated in the one or more target cells is up-regulated.

The term "modifying the expression of a nucleic acid molecule encoding a miRNA sequence", as used herein, denotes any manipulation of a particular nucleic acid molecule resulting in an altered expression level of said molecule, that is, the production of a different amount of corresponding miRNA as compared to the expression of the "wild-type" (i.e. the unmodified control). The term "different amount", as used herein, includes both a higher amount and a lower amount than determined in the unmodified control. In other words, a manipulation, as defined herein, may either up-regulate (i.e. activate) or down-regulate (i.e. inhibit) the expression (i.e. particularly transcription) of a nucleic acid molecule.

Within the present invention, expression of one or more nucleic acid molecules encoding a microRNA sequence comprised in the nucleic acid expression signature is modified in such way that the expression of a nucleic acid molecule whose expression is up-regulated in the one or more target cells is down-regulated and the expression of a nucleic acid molecule whose expression is down-regulated in the one or more target cells is up-regulated. In other words, the modification of expression of a particular nucleic acid molecule encoding a miRNA sequence occurs in an anti-cyclical pattern to the regulation of said molecule in the one or more cancerous target cells in order to interfere with the "excess activity" of an up-regulated molecule and/or to

restore the "deficient activity" of a down-regulated molecule in the one or more target cells.

In a preferred embodiment of the inventive method, down-regulating the expression of a nucleic acid molecule comprises introducing into the one or more target cells a nucleic acid molecule encoding a sequence that is complementary to the microRNA sequence encoded by nucleic acid molecule to be down-regulated.

The term "introducing into a cell", as used herein, refers to any manipulation allowing the transfer of one or more nucleic acid molecules into a cell. Examples of such techniques include *inter alia* transfection or transduction techniques all of them well established in the art (cf., for example, Sambrook, J. et al. (1989) *Molecular, Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; Ausubel, F.M. et al. (2001) *Current Protocols in Molecular Biology*, Wiley & Sons, Hoboken, NJ).

The term "complementary sequence", as used herein, is to be understood that the "complementary" nucleic acid molecule (herein also referred to as an "anti-sense nucleic acid molecule") introduced into the one or more cells is capable of forming base pairs, preferably Watson-Crick base pairs, with the up-regulated endogenous "sense" nucleic acid molecule.

Two nucleic acid molecules (i.e. the "sense" and the "anti-sense" molecule) may be perfectly complementary, that is, they do not contain any base mismatches and/or additional or missing nucleotides. In other embodiments, the two molecules comprise one or more base mismatches or differ in their total numbers of nucleotides (due to additions or deletions). In further embodiments, the "complementary" nucleic acid molecule comprises at least ten contiguous nucleotides showing perfect complementarity with a sequence comprised in the up-regulated "sense" nucleic acid molecule.

The "complementary" nucleic acid molecule (i.e. the nucleic acid molecule encoding a nucleic acid sequence that is complementary to the microRNA sequence encoded by nucleic acid molecule to be down-regulated) may be a naturally occurring DNA- or RNA molecule or a synthetic nucleic acid molecule comprising in its sequence one or more modified nucleotides which may be of the same type or of one or more different types.

For example, it may be possible that such a nucleic acid molecule

comprises at least one ribonucleotide backbone unit and at least one deoxyribonucleotide backbone unit. Furthermore, the nucleic acid molecule may contain one or more modifications of the RNA backbone into 2'-*O*-methyl group or 2'-*O*-methoxyethyl group (also referred to as "2'-*O*-methylation"), which prevented nuclease degradation in the culture media and, importantly, also prevented endonucleolytic cleavage by the RNA-induced silencing complex nuclease, leading to irreversible inhibition of the miRNA. Another possible modification – which is functionally equivalent to 2'-*O*-methylation – involves locked nucleic acids (LNAs) representing nucleic acid analogs containing one or more LNA nucleotide monomers with a bicyclic furanose unit locked in an RNA-mimicking sugar conformation (cf., e.g., Orom, U.A. et al. (2006) *Gene* **372**, 137-141).

Another class of silencers of miRNA expression was recently developed. These chemically engineered oligonucleotides, named "antagomirs", represent single-stranded 23-nucleotide RNA molecules conjugated to cholesterol (Krutzfeldt, J. et al. (2005) *Nature* **438**, 685–689). As an alternative to such chemically modified oligonucleotides, microRNA inhibitors that can be expressed in cells, as RNAs produced from transgenes, were generated as well. Termed "microRNA sponges", these competitive inhibitors are transcripts expressed from strong promoters, containing multiple, tandem binding sites to a microRNA of interest (Ebert, M.S. et al. (2007) *Nat. Methods* **4**, 721-726).

In particularly preferred embodiments of the inventive method, the one nucleic acid molecule whose expression is to be down-regulated encode microRNA sequence hsa-miR-375.

In a further preferred embodiment of the inventive method, up-regulating the expression of a nucleic acid molecule comprises introducing into the one or more target cells a nucleic acid molecule encoding the microRNA sequence encoded by nucleic acid molecule to be up-regulated. In other words, the up-regulation of the expression of a nucleic acid molecule encoding a miRNA sequence is accomplished by introducing into the one or more cells another copy of said miRNA sequence (i.e. an additional "sense" nucleic acid molecule). The "sense" nucleic acid molecule to be introduced into the one or more target cells may comprise the same modification as the "anti-sense" nucleic acid molecules described above.

In a particularly preferred embodiment, the one or more nucleic acid molecules whose expression is to be up-regulated encode microRNA sequences selected from the group consisting of hsa-miR-125b, hsa-miR-424, hsa-miR-92a, hsa-miR-99a and hsa-miR-7.

The "sense" and/or the "anti-sense" nucleic acid molecules to be introduced into the one ore more target cells in order to modify the expression of one or more nucleic acid molecules encoding a microRNA sequence that is/are comprised in the nucleic acid expression signature may be operably linked to a regulatory sequence in order to allow expression of the nucleotide sequence.

In order to unravel any potential implication of the miRNAs identified in the cancerous or pre-cancerous samples preliminary functional analyses may be performed with respect to the identification of mRNA target sequences to which the miRNAs may bind. Based on the finding that miRNAs may be involved in both tumor suppression and tumorigenesis (reviewed, e.g., in Esquela-Kerscher, A. and Slack, F.J (2006) *supra*; Calin, G.A. and Croce, C.M. (2007) *supra*; Blenkiron, C. and Miska, E.A. (2007) *supra*) it is likely to speculate that mRNA target sites for such miRNAs include tumor suppressor genes as well as oncogenes.

A nucleic acid molecule is referred to as "capable of expressing a nucleic acid molecule" or capable "to allow expression of a nucleotide sequence" if it comprises sequence elements which contain information regarding to transcriptional and/or translational regulation, and such sequences are "operably linked" to the nucleotide sequence encoding the polypeptide. An operable linkage is a linkage in which the regulatory sequence elements and the sequence to be expressed (and/or the sequences to be expressed among each other) are connected in a way that enables gene expression.

The precise nature of the regulatory regions necessary for gene expression may vary among species, but in general these regions comprise a promoter which, in prokaryotes, contains both the promoter *per se*, i.e. DNA elements directing the initiation of transcription, as well as DNA elements which, when transcribed into RNA, will signal the initiation of translation. Such promoter regions normally include 5' non-coding sequences involved in initiation of transcription and translation, such as the -35/-10 boxes and the Shine-Dalgarno element in prokaryotes or the TATA box, CAAT sequences, and 5'-capping elements in eukaryotes. These regions can also include enhancer or repressor elements as well as translated signal and leader sequences for

targeting the native polypeptide to a specific compartment of a host cell.

In addition, the 3' non-coding sequences may contain regulatory elements involved in transcriptional termination, polyadenylation or the like. If, however, these termination sequences are not satisfactory functional in a particular host cell, then they may be substituted with signals functional in that cell.

Furthermore, the expression of the nucleic molecules, as defined herein, may also be influenced by the presence, e.g., of modified nucleotides (cf. the discussion above). For example, locked nucleic acid (LNA) monomers are thought to increase the functional half-life of miRNAs *in vivo* by enhancing the resistance to degradation and by stabilizing the miRNA-target duplex structure that is crucial for silencing activity (cf., e.g., Naguibneva, I. et al. (2006) *Biomed. Pharmacother.* **60**, 633–638).

Therefore, a nucleic acid molecule of the invention to be introduced into the one or more cells provided may include a regulatory sequence, preferably a promoter sequence, and optionally also a transcriptional termination sequence.

The promoters may allow for either a constitutive or an inducible gene expression. Suitable promoters include *inter alia* the *E. coli lac*UV5 and *tet* (tetracycline-responsive) promoters, the T7 promoter as well as the SV40 promoter or the CMV promoter.

The nucleic acid molecules of the invention may also be comprised in a vector or other cloning vehicles, such as plasmids, phagemids, phages, cosmids or artificial chromosomes. In a preferred embodiment, the nucleic acid molecule is comprised in a vector, particularly in an expression vector. Such an expression vector can include, aside from the regulatory sequences described above and a nucleic acid sequence encoding a genetic construct as defined in the invention, replication and control sequences derived from a species compatible with the host that is used for expression as well as selection markers conferring a selectable phenotype on transfected cells. Large numbers of suitable vectors such as pSUPER and pSUPERIOR are known in the art, and are commercially available.

In a third aspect, the invention relates to a pharmaceutical composition for the prevention and/or treatment of colorectal cancer, preferably manifested as an adenocarcinoma, in one or more mammalian target cells, the composition comprising one or more nucleic acid molecules, each nucleic acid molecule encoding a sequence that is at least partially complementary to a microRNA sequence encoded by a nucleic

acid molecule whose expression is up-regulated in the one or more target cells, as defined herein, and/or that corresponds to a microRNA sequence encoded by a nucleic acid molecule whose expression is down-regulated in the one or more target cells, as defined herein.

In a final aspect, the invention is directed to the use of such a pharmaceutical composition for the manufacture of a medicament for the prevention and/or treatment of colorectal cancer, preferably manifested as an adenocarcinoma.

In the context of the present invention, suitable pharmaceutical compositions include those suitable for oral, rectal, nasal, topical (including buccal and sub-lingual), peritoneal and parenteral (including intramuscular, subcutaneous and intravenous) administration, or for administration by inhalation or insufflation. Administration may be local or systemic. Preferably, administration is accomplished via the oral, rectal or intravenous routes. The formulations may be packaged in discrete dosage units.

Pharmaceutical compositions according to the present invention include any pharmaceutical dosage forms established in the art, such as *inter alia* capsules, microcapsules, cachets, pills, tablets, powders, pellets, multi-particulate formulations (e.g., beads, granules or crystals), aerosols, sprays, foams, solutions, dispersions, tinctures, syrups, elixirs, suspensions, water-in-oil emulsions such as ointments, and oil-in water emulsions such as creams, lotions, and balms.

The ("sense" and "anti-sense") nucleic acid molecules described above can be formulated into pharmaceutical compositions using pharmacologically acceptable ingredients as well as established methods of preparation (Gennaro, A.L. and Gennaro, A.R. (2000) *Remington: The Science and Practice of Pharmacy*, 20th Ed., Lippincott Williams & Wilkins, Philadelphia, PA; Crowder, T.M. et al. (2003*) A Guide to Pharmaceutical Particulate Science.* Interpharm/CRC, Boca Raton, FL; Niazi, S.K. (2004) *Handbook of Pharmaceutical Manufacturing Formulations*, CRC Press, Boca Raton, FL).

In order to prepare the pharmaceutical compositions, pharmaceutically inert inorganic or organic excipients (i.e. carriers) can be used. To prepare e.g. pills, tablets, capsules or granules, for example, lactose, talc, stearic acid and its salts, fats, waxes, solid or liquid polyols, natural and hardened oils may be used. Suitable excipients for the production of solutions, suspensions, emulsions, aerosol mixtures or

powders for reconstitution into solutions or aerosol mixtures prior to use include water, alcohols, glycerol, polyols, and suitable mixtures thereof as well as vegetable oils. The pharmaceutical composition may also contain additives, such as, for example, fillers, binders, wetting agents, glidants, stabilizers, preservatives, emulsifiers, and furthermore solvents or solubilizers or agents for achieving a depot effect. The latter is to be understood that the nucleic acid molecules may be incorporated into slow or sustained release or targeted delivery systems, such as liposomes, nanoparticles, and microcapsules.

To target most tissues within the body, clinically feasible noninvasive strategies are required for directing such pharmaceutical compositions, as defined herein, into cells. In the past years, several approaches have achieved impressive therapeutic benefit following intravenous injection into mice and primates using reasonable doses of siRNAs without apparent limiting toxicities.

One approach involves covalently coupling the passenger strand (miRNA* strand) of the miRNA to cholesterol or derivatives/conjugates thereof to facilitate uptake through ubiquitously expressed cell-surface LDL receptors (Soutschek, J. et al. (2004) *Nature* **432**, 173-178). Alternatively, unconjugated, PBS-formulated locked-nucleic-acid-modified oligonucleotides (LNA-antimiR) may be used for systemic delivery (Elmen, J. et al. (2008) *Nature* **452**, 896-899). Another strategy for delivering miRNAs involves encapsulating the miRNAs into specialized liposomes formed using polyethylene glycol to reduce uptake by scavenger cells and enhance time spent in the circulation. These specialized nucleic acid particles (stable nucleic acid-lipid particles or SNALPs) delivered miRNAs effectively to the liver (and not to other organs (cf., e.g., Zimmermann, T.S. et al. (2006) *Nature* **441**, 111-114). Recently, a new class of lipid-like delivery molecules, termed lipidoids (synthesis scheme based upon the conjugate addition of alkylacrylates or alkyl-acrylamides to primary or secondary amines) has been described as delivery agents for RNAi therapeutics (Akinc, A. et al. (2008) *Nat. Biotechnol.* **26**, 561-569).

A further cell-specific targeting strategy involves the mixing of miRNAs with a fusion protein composed of a targeting antibody fragment linked to protamine, the basic protein that nucleates DNA in sperm and binds miRNAs by charge (Song, E. et al. (2005) *Nat. Biotechnol.* **23**, 709-717). Multiple modifications or variations of the above basic delivery approaches have recently been developed. These techniques are

known in the art and reviewed, e.g., in de Fougerolles, A. et al. (2007) *Nat. Rev. Drug Discov.* **6**, 443-453; Kim, D.H. and Rossi, J.J. (2007) *Nat. Genet.* **8**, 173-184).

The invention is further described by the figures and the following examples, which are solely for the purpose of illustrating specific embodiments of this invention, and are not to be construed as limiting the scope of the invention in any way.

EXAMPLES

Example 1: Patient materials

In the discovery study, 47 frozen surgical tissues from CRC patients were obtained from the tissue bank at Shanghai Medical College in Fudan University between 2007 and 2009. All patients who participated in the study had given informed consent. The collection of the tissue specimens in accordance with the protocol was approved by the Institutional Review Board of Shanghai Medical College. The specimens included 7 high-grade adenoma, 20 Dukes' A carcinoma and 20 matched normal colorectal tissues (at least 10 cm from tumor loci).

In the validation study, 128 of formalin-fixed, paraffin-embedded (FFPE) tissues from CRC patients were obtained from Shanghai Huashan Hospital and Shanghai Zhongshan Hospital between 2006 and 2009. Of the 128 FFPE tissues, 53 were colonoscopy biopsy specimens from the patients who had both colonoscopy biopsy and surgical specimens. In the biopsy tissues, all of them were pathologically diagnosed as high-grade adenoma. In the surgical resection tissues, 24 were confirmed as high-grade adenoma, while 29 were re-diagnosed as carcinomas. Baseline characteristics of the tumor specimens for both discovery and validation studies are are shown in Table 3.

Patient data (age, sex, imaging data, therapy, other medical conditions, family history, and the like) were derived from the hospital databases for matching the various samples collected. Pathologic follow-up (for example, histological analysis via hematoxylin and eosin (H&E) staining) was used for evidently determining the disease state (i.e. healthy control, adenoma, adenocarcinoma or intermediate state) of a given sample as well as to ensure a consistent classification of the specimens.

Table 3

Characteristics of tumor specimens in the discovery and validation studies

	Discovery (Frozen)	Validation (FFPE)	
Patient characteristics	Surgical tissue	Surgical tissue	Biopsy tissue

Dukes' A carcinoma	20	30	
High-grade adenoma	7	24	53
Adjacent normal tissue	20	21	0
Total samples	47	75	53

Example 2: Sample preparations

In the discovery study, laser-capture micro-dissection was performed for each frozen tissue specimen in order to specifically isolate tumor cell populations (about 200.000 cells). In brief, a transparent transfer film is applied to the surface of a tissue section or specimen. Under a microscope, the thin tissue section is viewed through the glass slide on which it is mounted and clusters of cells are identified for isolation. When the cells of choice are in the center of the field of view, a near IR laser diode integral with the microscope optics is activated. The pulsed laser beam activates a spot on the transfer film, fusing the film with the underlying cells of choice. The transfer film with the bonded cells is then lifted off the thin tissue section (reviewed, e.g., in Emmert-Buck, M.R. et al. (1996). Science 274, 998-1001; Espina, V. et al. (2007) Expert Rev. Mol. Diagn. 7, 647-657). The preparation of the cryostat sections and the capturing step using a laser capture microspope (Arcturus VeritasTM Laser Capture Microdissection Instrument (Molecular Devices, Inc., Sunnyvale, CA, USA) were performed essentially according to the instructions of the manufacturer.

To aid the transition from exploratory research to clinical implementation, FFPE surgical and biopsy tissues were used in the validation study. Once FFPE tissues were selected for the analysis, H&E-stained sections were prepared in order to check the proportion of tumour material in each sample. If a tumour has more than 75% neoplastic cells, it shall be deemed suitable for analysis without further purification of tumour cells. If, however, histology shows the tumour to have <75% neoplastic cells, it will be selected and marked tumours for macrodissection. In addition to tumour lesions, control tissue shall be derived from at least 10 cm from tumour loci.

Total RNA was extracted from the tissue sections by using mirVana miRNA isolation kit according to the instructions from the manufacturer (Ambion, Austin, TX). The concentration was quantified by NanoDrop 1000 Spectrophotometer (NanoDrop Technologies, Waltham, MA). The quality control of RNA was performed by a 2100 Bioanalyzer using the RNA 6000 Pico LabChip kit (Agilent Technologies, Santa Clara, CA).

Example 3: The microarray data

In the discovery study, a qualitative analysis of the miRNAs differentially expressed in a particular sample may optionally be performed using the Agilent miRNA microarray platform (Agilent Technologies, Santa Clara, CA, USA). The microarray contains probes for 723 human miRNAs from the Sanger database v.10.1. Total RNA (100 ng) derived from each of 47 LCM-selected colorectal samples were used as inputs for labeling via Cy3 incorporation. Microarray slides were scanned by XDR Scan (PMT100, PMT5). The labeling and hybridization were performed according to the protocols in the Agilent miRNA microarray system. The raw data obtained for single-color (CY3) hybridization were normalized by applying a Quantile method and using GeneSpring GX10 software (Agilent Technologies, Santa Clara, CA, USA) known in the art.

Unpaired t-test after Fisher test (F-test) was used to identify top miRNA signatures between adenoma vs. carcinomas. For the specificity and sensitivity of the individual miRNA as diagnostic biomarkers, MedCalc software was used to perform receiver operating characteristic (ROC) curve analysis of the individual miRNA in adenoma vs. carcinomas. 95% confidence interval was used to determine the significance.

The experimental data on the array analysis of 6 key candidate miRNAs in the first aspect in the frozen surgical tissues for discriminating early colorectal carcinoma from high-grade adenoma are shown in Table 4.

Table 4
Candidate miRNA biomarkers in the frozen surgical tissues for discriminating Dukes'A carcinoma from high-grade adenoma

	Dukes'A carcinoma/high-grade adenoma				
	p-value	Fold change	Sensitivity	Specificity	AUC
Individual miRNA					
hsa-miR-125b	7.9E-03	4.4	65%	67%	0.672
hsa-miR-375	1.3E-02	0.3	80%	67%	0.722
hsa-miR-424	1.3E-03	4.5	85%	63%	0.735
hsa-miR-92a	1.6E-03	4.3	85%	63%	0.733
hsa-miR-99a	1.9E-03	5.9	70%	71%	0.752
hsa-miR-7	6.9E-03	4.6	80%	67%	0.712
miRNA combination					

hsa-miR-125b/ hsa-miR-375	1.6E-06	12.1	80%	88%	0.882
hsa-miR-424/ hsa-miR-375	2.6E-03	9.0	85%	83%	0.825
hsa-miR-92a/ hsa-miR-375	1.2E-07	12.8	95%	88%	0.942
hsa-miR-99a/ hsa-miR-375	4.2E-05	10.3	75%	88%	0.848
hsa-miR-7/ hsa-miR-375	5.3E-06	13.3	90%	83%	0.902

Example 3: Validation of the microarray data in FFPE surgical tissues

For validation of the miRNA expression data acquired on microarrays, an established quantitative RT-PCR employing a TaqMan MicroRNA assay (Applied Biosystems, Foster City, CA, USA) was used according to the manufacturer's instructions. The assays were performed for hsa-miR-125b (SEQ ID NO: 3), hsa-miR-375 (SEQ ID NO: 2), hsa-miR-424 (SEQ ID NO: 4), hsa-miR-92a (SEQ ID NO: 1), hsa-miR-99a (SEQ ID NO: 5) and hsa-miR-7 (SEQ ID NO: 6) using 75 FFPE surgical tissue specimens. The expression level of the small nuclear RNA U47 was used as the normalization control. All assays were carried out in triplicate.

Briefly, reverse transcription (RT) was performed with Taqman microRNA RT Kits according to the instruction from Applied Biosystem. 100ng total RNA was reverse-transcripted in 15ul RT solution mix that contains 1X Reverse Transcription Buffer, 1X RT primer, 1nM dNTP, 4U RNase Inhibitor and 50U MultiScribe Reverse Transcriptase. Then the RT solutions were performed by using the thermal program of 16°C, 30min; 42°C, 30min; 85°C, 5min on the PCR machine (Thermal cycler alpha engine, Bio-rad). Quantitative PCR was performed with TaqMan Universal PCR Master Mix kit and and Taqman microRNA assays kits according to the instruction from Applied Biosystem. 2ul RT products were PCR amplified in 1X TaqMan Universal PCR Master Mix, No AmpErase UNG, 1X TaqMan MicroRNA Assay mix. The real-time PCR was performed in Roch Light Cycling 480 machine with the program of 96°C, 5min initial heating; then 45 or 50 cycles of 95°C, 15s; 60°C, 60s. Cp value was calculated with 2nd derivative method in LC480 software. Then miRNAs were absolutely quantified with the standard samples Cp values.

Unpaired t-test after Fisher test (F-test) was used to identify top miRNA biomarkers between early colorectal carcinoma vs. high-grade adenoma. For the specificity and sensitivity of the individual miRNA as diagnostic biomarkers, MedCalc software was used to perform receiver operating characteristic (ROC) curve analysis of the individual miRNA. 95% confidence interval was used to determine the significance.

The experimental data on the 6 validated miRNA biomarkers in FFPE surgical tissues in the first aspect for discriminating Dukes'A carcinoma from high-grade adenoma show in Table 5. Particularly preferred hsa-miR-92a (SEQ ID NO: 1), hsa-miR-375 (SEQ ID NO: 2) and hsa-miR-99a (SEQ ID NO: 5) are shown in bold.

Table 5
Validated miRNA biomarkers in FFPE surgical tissues for discriminating Dukes'A carcinoma from high-grade adenoma

	Dukes'A carcinoma/high-grade adenoma				
	p-value	Fold change	Sensitivity	Specificity	AUC
Individual miRNA					
hsa-miR-125b	7.9E-03	4.4	65%	67%	0.672
hsa-miR-375	1.3E-02	0.3	80%	67%	0.722
hsa-miR-424	1.3E-03	4.5	85%	63%	0.735
hsa-miR-92a	1.6E-03	4.3	85%	63%	0.733
hsa-miR-99a	1.9E-03	5.9	70%	71%	0.752
hsa-miR-7	6.9E-03	4.6	80%	67%	0.712
miRNA combination					
hsa-miR-125b/ hsa- miR-375	1.6E-06	12.1	80%	88%	0.882
hsa-miR-424 / hsa- miR-375	2.6E-03	9.0	85%	83%	0.825
hsa-miR-92a / hsa- miR-375	1.2E-07	12.8	95%	88%	0.942
hsa-miR-99a / hsa- miR-375	5.3E-06	13.3	90%	83%	0.902
hsa-miR-7 / hsa- miR-375	4.2E-05	10.3	75%	88%	0.848

Example 4: Validation of the microarray data in FFPE biopsy tissues

53 biopsy specimens from the patients who had both colonoscopy biopsy and surgical specimens (see Table 1). In the biopsy tissues, all of them were pathologically diagnosed as high-grade adenoma. In the surgical resection tissues, 24 were confirmed as high-grade adenoma, while 29 were re-diagnosed as carcinomas.

Quantitative RT-PCR was performed on the 6 validated miRNA biomarkers in the FFPE biopsy tissues from colonoscopy. The expression level of the small nuclear RNA U47 was used as the normalization control. All assays were carried out in triplicate. Unpaired t-test after Fisher test (F-test) was used to differentially expressed miRNAs between high-grade adenoma and carcinoma. For the specificity and sensitivity of the individual miRNA as diagnostic biomarkers, MedCalc software was used to perform receiver operating characteristic (ROC) curve analysis of the individual miRNA in adenoma vs. carcinomas. 95% confidence interval was used to determine the significance.

The experimental data on the validated miRNA biomarkers in the first aspect in FFPE biopsy tissues from colonoscopy for discriminating high-grade adenoma from carcinoma are shown in Table 6. Particularly preferred hsa-miR-7 (SEQ ID NO: 6), hsa-miR-375 (SEQ ID NO: 2) and hsa-miR-99a (SEQ ID NO: 5) are shown in bold.

Table 6
Validated miRNA biomarkers in FFPE biopsy tissues for discriminating carcinoma from high-grade adenoma

	Carcinoma/ high-grade adenoma				
	p-value	Fold change	Sensitivity	Specificity	RUC
Individual miRNA					
hsa-miR-125b	2.1E-03	3.9	79%	71%	0.748
hsa-miR-375	2.0E-02	0.4	45%	91%	0.666
hsa-miR-424	1.5E-04	4.9	86%	79%	0.786
hsa-miR-92a	5.5E-04	4.1	90%	58%	0.793
hsa-miR-99a	2.9E-03	2.7	83%	63%	0.733
hsa-miR-7	2.4E-06	7.8	72%	88%	0.861
miRNA combination					
hsa-miR-125b/hsa-miR-375	2.8E-05	8.9	83%	77%	0.855
hsa-miR-424/hsa-miR-375	3.8E-06	12.4	71%	86%	0.843
hsa-miR-92a/hsa-miR-375	1.9E-05	10.0	69%	86%	0.825
hsa-miR-99a/hsa-miR-375	1.8E-05	6.4	83%	86%	0.849
hsa-miR-7/hsa-miR-375	1.1E-06	19.4	79%	86%	0.875

Example 5: Method for quantifying miRNA biomarkers

A quantitative analysis of the miRNA biomarkers (differentially) expressed in a particular sample may optionally be performed by quantitative RT-PCR employing a TaqMan MicroRNA assay (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions.

Particularly preferably, the quantification of the miRNA biomarkers is performed by using *in situ* hybridization (Fig.1). The method comprising: (a) collecting a biopsy or surgical tissue from a patient; (b) preparing tissue section on a slide; (c) hybridizing at least one nucleic acid molecule encoding a microRNA sequence to the section on the slide; (d) quantifying the miRNA expression under microscope or by digital pathology solution; (e) determining in the one or more target cells the expression levels of a plurality of nucleic acid molecules, each nucleic acid molecule encoding a microRNA sequence; (f) determining the expression levels of the plurality of nucleic

acid molecules in one or more control cells; and (g) identifying from the plurality of nucleic acid molecules one or more nucleic acid molecules that are differentially expressed in the target and control cells by comparing the respective expression levels obtained in steps (e) and (f), wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic acid expression biomarker, as defined herein, that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

For normalizing the expression levels obtained for the nucleic acid molecules encoding microRNA sequences that are comprised in the nucleic acid expression signature hsa-miR-423-5p (SEQ ID NO:7) may be preferably used, which is stably expressed in colorectal tissues. For negative control of the expression levels obtained for the nucleic acid molecules encoding microRNA sequences that are comprised in the nucleic acid expression signature hsa-miR-122 (SEQ ID NO:8) may be preferably used, which does not expressed in colorectal tissues.

Probe for hybridization is a synthesized unmodified sequence that hybridizes to the target miRNA including a 30 base residue (GGGGGTCCTATATGGCTCCACTTCTCCCCC). The residue sequence is shown in bold. The probe is labelled with a fluorophore at the 5' end. Singe or multiple probes with separate fluorescent dyes can be hybridized in parallel. The probe sequences for *in situ* hybridization in the invention are given in Table 7.

The residue forms 5' hairpin (GGGGG-CCCCC pair) to stabilize the hybridization between the probe and target miRNA and increase the hybridization specificity. The same principle can be used to design probes for other miRNA biomarkers using *in situ* hybridization.

TABLE 7
The probes for *in situ* hybridization

miRNA	Target sequence (5' to 3')	Probe sequence (5' to 3')
Biomarker		
hsa-miR-92a	uauugcacuugucccggccugu	**GGGGGTCCTATATGGCTCCACTTCTCCCCC** ACAGGCCGGGACAAGTGCAATA
hsa-miR-375	uuuguucguucggcucgcguga	**GGGGGTCCTATATGGCTCCACTTCTCCCCC** TCACGCGAGCCGAACGAACAAA
hsa-miR-125b	ucccugagacccuaacuuguga	**GGGGGTCCTATATGGCTCCACTTCTCCCCC**

		TCACAAGTTAGGGTCTCAGGGA
hsa-miR-424	cagcagcaauucauguuuugaa	**GGGGGTCCTATATGGCTCCACTTCTCCCCC** TTCAAAACATGAATTGCTGCTG
hsa-miR-99a	aacccguagauccgaucuugug	**GGGGGTCCTATATGGCTCCACTTCTCCCCC** CACAAGATCGGATCTACGGGTT
hsa-miR-7	uggaagacuagugauuuuguugu	**GGGGGTCCTATATGGCTCCACTTCTCCCCC** ACAACAAAATCACTAGTCTTCCA
Controls		
hsa-miR-122	uggagugugacaaugguguuug	**GGGGGTCCTATATGGCTCCACTTCTCCCCC** CAAACACCATTGTCACACTCCA
hsa-miR-423-5p	ugaggggcagagagcgagacuuu	**GGGGGTCCTATATGGCTCCACTTCTCCCCC** AAAGTCTCGCTCTCTGCCCCTCA

The present invention illustratively described herein may suitably be practiced in the absence of any element or elements, limitation or limitations, not specifically disclosed herein. Thus, for example, the terms "comprising", "including", "containing", etc. shall be read expansively and without limitation. Additionally, the terms and expressions employed herein have been used as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by embodiments and optional features, modifications and variations of the inventions embodied therein may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention.

The invention has been described broadly and generically herein. Each of the narrower species and sub-generic groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein.

Other embodiments are within the following claims. In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group.

CLAIMS:

1. Diagnostic kit of molecular markers for identifying one or more mammalian target cells in colorectal surgical and biopsy tissues exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprising a plurality of nucleic acid molecules, each nucleic acid molecule encoding a microRNA sequence,

wherein one or more of the plurality of nucleic acid molecules are differentially expressed in the target cells and in one or more control cells, and

wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic acid expression biomarker that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

2. The kit of claim 1, wherein the early colorectal carcinoma is Dukes' A carcinoma, and the high-grade adenoma is high-grade intraepithelial neoplasm.

3. The kit of any of claims 1 to 2, wherein the nucleic acid expression biomarker at least one nucleic acid molecule encoding a microRNA sequence whose expression is up-regulated in the one or more target cells compared to the one or more control cells and at least one nucleic acid molecule encoding a microRNA sequence whose expression is down-regulated in the one or more target cells compared to the one or more control cells.

4. The kit of any of claims 1 to 3, for the use of identifying a early colorectal carcinoma or high-grade adenoma, wherein the nucleic acid expression biomarker at least three nucleic acid molecules, preferably at least six nucleic acid molecules.

5. The kit of any of claim 1 to 4, wherein the nucleic acid expression biomarker comprises any one or more nucleic acid molecules encoding microRNA sequences selected from the group consisting of hsa-miR-375, hsa-miR-92a and hsa-miR-99a, and preferably comprises any one or more nucleic acid molecules encoding

microRNA sequences selected from the group consisting of hsa-miR-125b, hsa-miR-375, hsa-miR-424, hsa-miR-92a, hsa-miR-99a and hsa-miR-7.

6. The kit of any of claims 1 to 5, wherein the expression of any one or more of the nucleic acid molecules encoding hsa-miR-125b, hsa-miR-424, hsa-miR-92a, hsa-miR-99a and hsa-miR-7 are up-regulated and the expression of any one or more of the nucleic acid molecule hsa-miR-375 is down-regulated in the in the one or more target cells compared to the one or more control cells.

7. The kit of any of claim 1 to 4, wherein the nucleic acid expression biomarker comprises any one or more nucleic acid molecule combinations encoding hsa-miR-92a/hsa-miR-375 and hsa-miR-99a/hsa-miR-375, and preferably comprises any one or more nucleic acid molecule combinations hsa-miR-125b/hsa-miR-375, hsa-miR-424/hsa-miR-375, hsa-miR-92a/hsa-miR-375, hsa-miR-99a/hsa-miR-375 and hsa-miR-7/hsa-miR-375.

8. The kit of any of claims 1 to 4 and 7, wherein the expression of any one or more of the nucleic acid molecule combinations encoding hsa-miR-125b/hsa-miR-375, hsa-miR-424/hsa-miR-375, hsa-miR-92a/hsa-miR-375, hsa-miR-99a/hsa-miR-375 and hsa-miR-7/hsa-miR-375 are up-regulated compared to the one or more control cells.

9. Method for identifying one or more mammalian target cells exhibiting early colorectal carcinoma or high-grade adenoma, the method comprising:

(a) collecting a biopsy or surgical tissue from a patient;

(b) preparing tissue section on a slide;

(c) hybridizing at least one nucleic acid molecule encoding a microRNA sequence to the section on the slide;

(d) quantifying the miRNA expression under microscope or by digital pathology solution;

(e) determining in the one or more target cells the expression levels of a plurality of nucleic acid molecules, each nucleic acid molecule encoding a microRNA sequence;

(f) determining the expression levels of the plurality of nucleic acid molecules in one or more control cells; and

(g) identifying from the plurality of nucleic acid molecules one or more nucleic acid molecules that are differentially expressed in the target and control cells by comparing the respective expression levels obtained in steps (e) and (f), wherein the one or more differentially expressed nucleic acid molecules together represent a nucleic acid expression biomarker, as defined in any of claims 1 to 8, that is indicative for the presence of early colorectal carcinoma or high-grade adenoma.

10. The kit of any of claims 1 to 8, wherein, the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences are normalized by the nucleic acid expression molecule encoding hsa-miR-423-5p, which is stably expressed in colorectal tissues. For negative control of the expression levels obtained for the nucleic acid molecule biomarkers encoding microRNA sequences, the nucleic acid expression molecule encoding hsa-miR-122 is used, which does not expressed in colorectal tissues.

11. Method for preventing or treating colorectal cancer, preferably manifested as an adenocarcinoma, in one or more mammalian target cells, the method comprising:

(a) identifying in one or more target cells a nucleic acid expression biomarker by using a method as defined in claim 1 to 8; and

(b) modifying in the one or more cells the expression of one or more nucleic acid molecules encoding a microRNA sequence that is/are comprised in the nucleic acid expression signature in such way that the expression of a nucleic acid molecule whose expression is up-regulated in the one or more target cells is down-regulated and the expression of a nucleic acid molecule whose expression is down-regulated in the one or more target cells is up-regulated.

12. Pharmaceutical composition for the prevention and/or treatment of colorectal cancer, preferably manifested as an adenocarcinoma, in one or more mammalian target cells, the composition comprising one or more nucleic acid molecules, each nucleic acid molecule encoding a sequence that is at least partially complementary to a microRNA sequence encoded by a nucleic acid molecule whose expression is up-regulated in the one or more target cells and/or that corresponds to a microRNA sequence encoded by a nucleic acid molecule whose expression is down-regulated in the one or more target cells, as defined in any of claims 1 to 8.

13. Use of the pharmaceutical composition of claim 12 for the manufacture of a medicament for the prevention and/or treatment of colorectal cancer, preferably manifested as an adenocarcinoma.

FIGURE 1

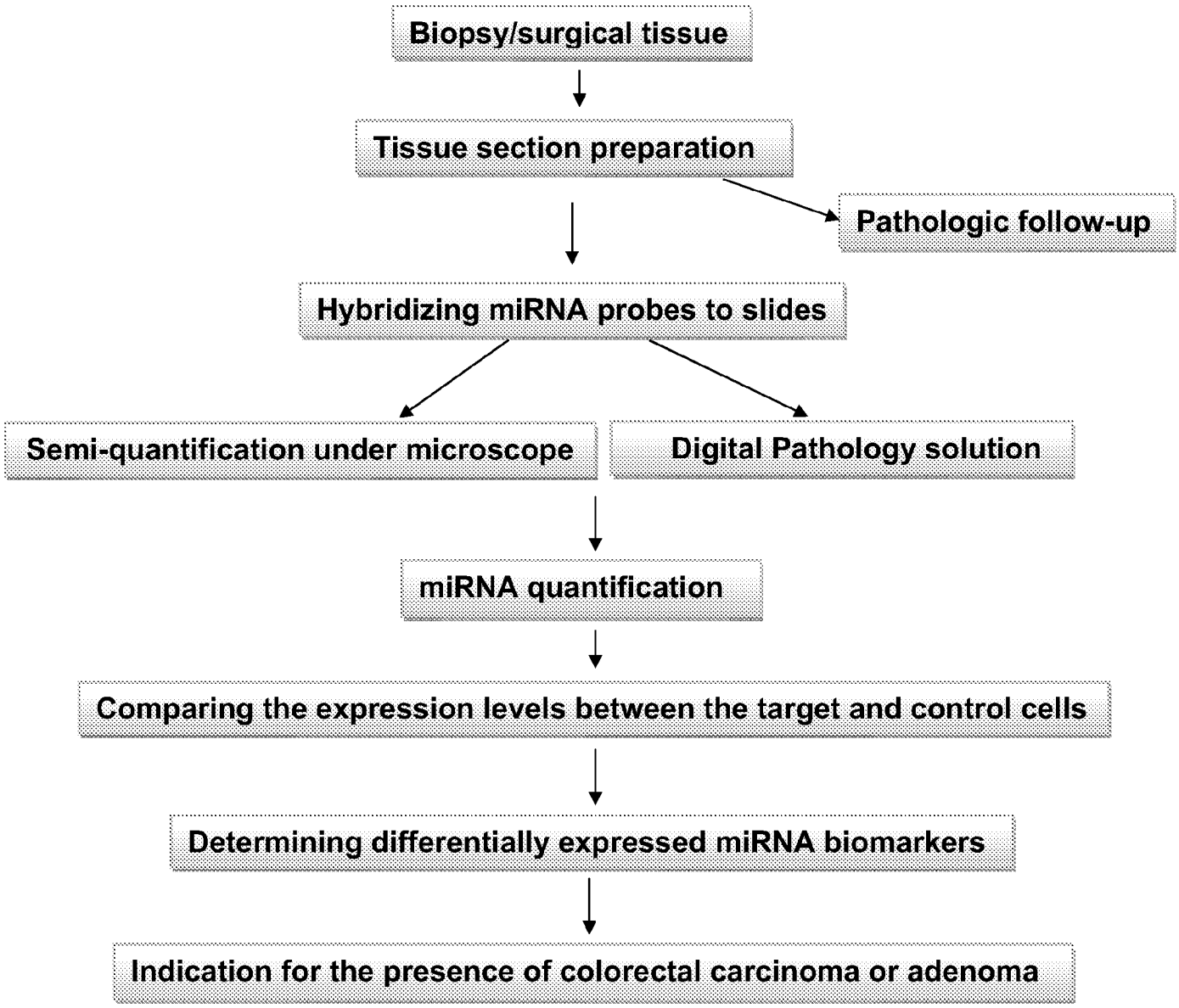
Biopsy/surgical tissue
Tissue section preparation
Pathologic follow-up
Hybridizing miRNA probes to slides
Semi-quantification under microscope
Digital Pathology solution
miRNA quantification
Comparing the expression levels between the target and control cells
Determining differentially expressed miRNA biomarkers
Indication for the presence of colorectal carcinoma or adenoma

FIGURE 2

	Dukes'A carcinoma/high-grade adenoma		
	Expression	**Accuracy**	**AUC**
Individual miRNA			
hsa-miR-125b	UP	66%	0.672
hsa-miR-375	DOWN	74%	0.722
hsa-miR-424	UP	74%	0.735
hsa-miR-92a	UP	74%	0.733
hsa-miR-99a	UP	71%	0.752
hsa-miR-7	UP	74%	0.712
miRNA combination			
hsa-miR-125b/ hsa-miR-375	UP	84%	0.882
hsa-miR-424/ hsa-miR-375	UP	84%	0.825
hsa-miR-92a/ hsa-miR-375	UP	92%	0.942
hsa-miR-99a/ hsa-miR-375	UP	82%	0.848
hsa-miR-7/ hsa-miR-375	UP	87%	0.902

FIGURE 3A

	Dukes' A carcinoma/high-grade adenoma		
	Expression	**Accuracy**	**AUC**
Individual miRNA			
hsa-miR-99a	UP	71%	0.752
hsa-miR-424	UP	74%	0.735
hsa-miR-92a	UP	74%	0.733
hsa-miR-375	DOWN	74%	0.722
hsa-miR-7	UP	74%	0.712
hsa-miR-125b	UP	66%	0.672
miRNA combination			
hsa-miR-92a/hsa-miR-375	UP	92%	0.942
hsa-miR-99a/hsa-miR-375	UP	87%	0.902
hsa-miR-125b/hsa-miR-375	UP	84%	0.882
hsa-miR-7/hsa-miR-375	UP	82%	0.848
hsa-miR-424/hsa-miR-375	UP	84%	0.825

FIGURE 3B

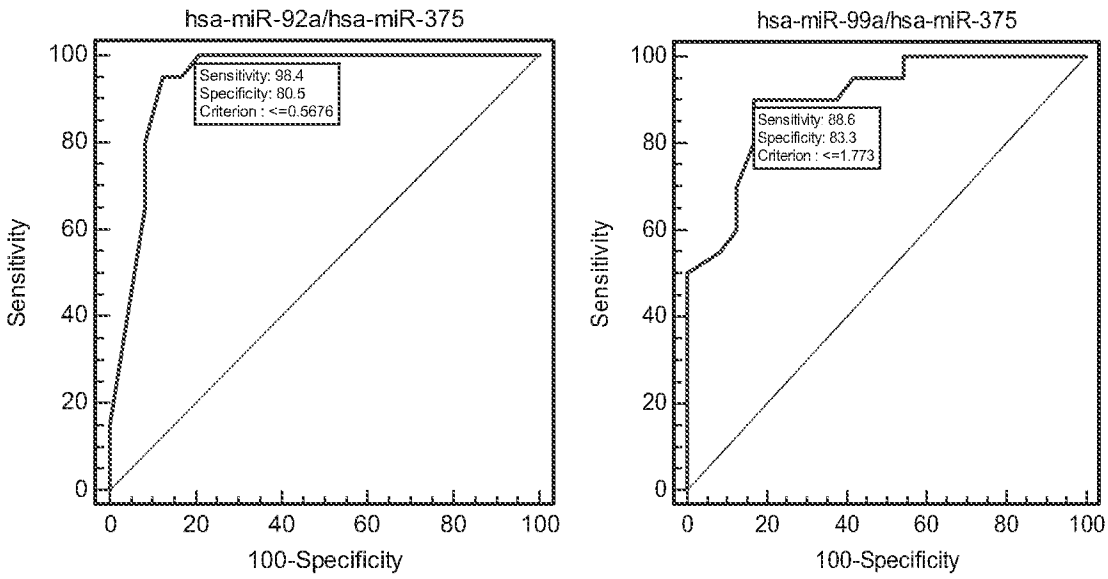

FIGURE 4

	Carcinoma/high-grade adenoma		
	Expression	**Accuracy**	**RUC**
Individual miRNA			
hsa-miR-7	UP	80%	0.861
hsa-miR-92a	UP	74%	0.793
hsa-miR-424	UP	82%	0.786
hsa-miR-125b	UP	75%	0.748
hsa-miR-99a	UP	73%	0.733
hsa-miR-375	DOWN	68%	0.666
miRNA combination			
hsa-miR-7/hsa-miR-375	UP	83%	0.875
hsa-miR-125b/hsa-miR-375	UP	80%	0.855
hsa-miR-99a/hsa-miR-375	UP	85%	0.849
hsa-miR-424/hsa-miR-375	UP	79%	0.843
hsa-miR-92a/hsa-miR-375	UP	78%	0.825

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2011/075132

A. CLASSIFICATION OF SUBJECT MATTER

SEE EXTRA SHEET

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)

IPC: C12Q , C12N , A61K , A61P

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)

CNABS,CPRSABS,TWABS,MOABS,DWPI,SIPOABS,CNKI,Elsevier SD,CA: microRNA, miRNA, colorectal carcinoma, adenoma,

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN101389770A(UNIV OHIO STATE),18 Mar. 2009(18.03.2009), see p.12 table1a-p.36 para.1, p.58 para.6-p.61 para.2 in description, claims	1-5,9,12-13(all partially)
A	CN101400361A (UNIV OHIO STATE), 1 Apr. 2009(01.04.2009),see the whole document	1-10,12-13

☐ Further documents are listed in the continuation of Box C. ☒ 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
30 AUG.2011(30.08.2011)	**22 Sep. 2011 (22.09.2011)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer ZHAO,Yanhao Telephone No. (86-10)62411043

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2011/075132

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

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

1. ☒ Claims Nos.: 11
because they relate to subject matter not required to be searched by this Authority, namely:
The subject matter of claim 11 relates to a method of treatment of the human/animal body(Rule 39.1(iv)).

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

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

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

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

See extra sheet

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

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

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

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

Claims 1-5,9,12-13(all partially)

Remark on protest

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

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

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2011/075132

Continuation of Box A:

CLASSIFICATION OF SUBJECT MATTER:

C12Q 1/68 (2006.01) i
C12N 15/11 (2006.01) i
A61K 31/7105 (2006.01) i
A61P 35/00 (2006.01) i

Continuation of Box No. III:

This International Searching Authority found 55 inventions in this international application, as follows:

1. claims 1-5,9,12-13(all partially), relate to a diagnostic kit of molecular markers for identifying one or more mammalian target cells in colorectal surgical and biopsy tissues exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprises an expression biomarker which comprises a nucleic acid molecules encoding has-miR-92a.

2. claims1-5,9,12-13(all partially), relate to a diagnostic kit of molecular markers for identifying one or more mammalian target cells in colorectal surgical and biopsy tissues exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprises an expression biomarker which comprises a nucleic acid molecules encoding has-miR-92a and has-miR-7.

3. claims 1-5,9,12-13(all partially), relate to a diagnostic kit of molecular markers for identifying one or more mammalian target cells in colorectal surgical and biopsy tissues exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprises an expression biomarker which comprises a nucleic acid molecules encoding has-miR-92a, has-miR-99a and has-miR-7.

. . .

36. claims1-6,9,12-13(all partially), relate to a diagnostic kit of molecular markers for identifying one or more mammalian target cells in colorectal surgical and biopsy tissues exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprises an expression biomarker which comprises a nucleic acid molecules encoding has-miR-125b, has-miR-424 ,has-miR-92a, has-miR-99a ,has-miR-7, has-miR-375.

37. claims 1-7,9,12-13 (all partially), relate to a diagnostic kit of molecular markers for identifying one or more mammalian target cells in colorectal surgical and biopsy tissues exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprises an expression biomarker which comprises nucleic acid molecule combinations encoding has-miR-92a/has-miR-375.

. . .

54. claims 1-9,12-13 (all partially), relate to a diagnostic kit of molecular markers for identifying one or more mammalian target cells in colorectal surgical and biopsy tissues exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprises an expression biomarker which comprises nucleic acid molecule combinations encoding has-miR-92a/has-miR-375, has-miR125b/has-miR-375, has-miR-424/has-miR-375, has-miR-99a/has-miR-375, has-miR-7/has-miR-375.

55. claims 1-10 (all partially), relate to a diagnostic kit of molecular markers for identifying one or more mammalian target cells in colorectal surgical and biopsy tissues exhibiting early colorectal carcinoma or high-grade adenoma, the kit comprises an expression biomarker which comprises two nucleic acid molecules encoding has-miR-423-5p and has-miR-122.

The foregoing 55 inventions don't have the same or corresponding special technical features. The application hence does not meet the requirements of unity of invention as defined in Rules 13.1 and 13.2 PCT.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2011/075132

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101389770 A	18.03.2009	WO 2007081740 A3	10.07.2008
		AU 2007205163 A1	19.07.2007
		US 2008306006 A1	11.12.2008
		CA 2633754 A1	19.07.2007
		EP 1969147 A2	17.09.2008
		JP 2009531019 T	03.09.2009
		WO 2007081740 A2	19.07.2007
CN 101400361 A	01.04.2009	WO 2007081720 A2	19.07.2007
		AU 2007205234 A1	19.07.2007
		EP 1978986 A2	15.10.2008
		CA 2633674 A1	19.07.2007
		JP 2009531018 T	03.09.2009
		US 2008306017 A1	11.12.2008
		WO 2007081720 A3	14.02.2008