Screening for differentially-expressed microRNA biomarkers in Saudi colorectal cancer patients by small RNA deep sequencing

YAZEED A. AL-SHEIKH, HAZEM K. GHNEIM, KHALID K. ALHARBI and MOURAD A.M. ABOUL-SOUD

Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, Riyadh 11433, Kingdom of Saudi Arabia

Received April 18, 2019; Accepted August 5, 2019

DOI: 10.3892/ijmm.2019.4362

Abstract. Colorectal cancer (CRC) is mostly diagnosed at late stages leading to high mortality rates due to the scarcity of efficient screening approaches exhibiting high diagnostic utility. The current study employed a small-RNA deep-sequencing approach for screening microRNA (miRNA) differentially expressed genes (DEGs), and evaluating their potential as early diagnostic circulating biomarkers for CRC in clinical plasma and tissue samples from a Saudi patient population. The cohort followed a paired-study design composed of 20 CRC patients, providing plasma (P) and tissue (T) samples of CRC, and adjacent normal mucosa (CT). Also, control plasma (CP) samples were obtained from neoplasm-free healthy individuals to compare its miRNA levels with those in P samples. Illumina high-throughput (HiSeq 2000) sequencing was performed for the identification of known and novel miRNA genes that were differentially expressed in the plasma and tissues of CRC patients compared with CT and CP controls. While we identified only one known (*hsa-miR-182-5p*, significantly upregulated) and no novel DEGs at the most stringent significance level (P<0.001) in the P-CP comparison, we found 3 and none at P<0.01, 7 and 9 at P<0.05 level, respectively. In the T-CT comparison, the results revealed 24 known and 196 novel miRNA DEGs (P<0.001), 31 and 204 (P<0.01), 41 and 213 (P<0.05), respectively. Sequencing data were then analyzed by bioinformatics for potential diagnostic miRNAs. Network functional analysis for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway implicated two pathways rooted to signal transduction [Wnt and mitogen-activated protein kinase (MAPK)] that were enriched in CRC patients. Our results suggest that characterizing plasma and tissue profiles of CRC by deep sequencing may be a good strategy for identifying known and novel miRNAs and that the validated miRNAs described here may serve as potential CRC-associated biomarkers. Further research is necessary for determining their screen index values and diagnostic utility for the diagnosis of CRC.

Correspondence to: Professor Mourad A.M. Aboul-Soud, Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, P.O. Box 10219, Riyadh 11433, Kingdom of Saudi Arabia
E-mail: maboulsoud@ksu.edu.sa

Key words: colorectal cancer, microRNA, Biomarkers, MAPK, Wnt, deep sequencing, small RNA

Introduction

Cancer represents a significant worldwide threat to public health as it is one of the primary causes of death. In 2018, while 18.1 million newly diagnosed cases were reported, mortalities and 5-year survival rates were 9.6 million and 43.8 million individuals, respectively. In 2030, it is estimated that these rising trends will persist reaching a striking 26.4 million newly diagnosed cases and 17 million cancer-related deaths. Globally, colorectal cancer (CRC) is the third most prevalent type of cancer (1.8 million cases, 10.2% of the total) accounting for 9.2% of cancer-related mortalities, second only to lung cancer (18.4% of the total) (1). By 2030, CRC incidence in Saudi Arabia has been predicted to rise 4-fold in both sexes; a projected increase in incidence and mortality rates by 350 and 160%, respectively, has been reported (2,3). Currently employed therapeutic strategies have been comprehensively effective in accomplishing notable amelioration in the 5-year survival rate, however the mortality rate remains as high as 60% in some cases, thereby emphasizing the need for an effective early diagnostic approach as a pivotal goal for any healthcare system (4). In this context, higher survival rates have been reported to exhibit positive correlation with early diagnosis and treatment of CRC (5). For instance, a 5-year survival rate of 93.2% for CRC patients diagnosed at stage I has been reported compared with only 8.1% for stage IV (5). Hence, effective therapy and successful clinical outcome of CRC patients are tightly dependent on early diagnosis.

Since the discovery of *lin-4* and *let-7* in 1993 and 2000, respectively, research into microRNAs (miRNAs) has gained a tremendous worldwide interest ascribed by the pivotal regulatory role that they play in various physiological and pathological cellular processes, including development, cell cycle progression, cell differentiation, proliferation and apoptosis (6-8). miRNAs are a ubiquitous class of short, single-stranded and non-coding RNAs that may post-transcriptionally regulate ~60% of all human genes and control hundreds of cognate gene targets through

their oncogenic or anti-oncogenic (i.e., tumor-suppressive) activity (9). This regulatory role and target recognition is considered to occur via direct binding to the 3' untranslated region (3' UTR) of a target mRNA, giving rise to its translational repression and loss of function (2). This sequence-specific miRNA-mRNA target-binding process may be affected by the secondary structure of regions flanking the target sequence and the complementarity rate of seed sequence (9). miRNAs are generally categorized into families, targeting common cognate mRNAs, according to the identity of seed sequences and the similarity of mature sequences (10). miRNAs display aberrant miRNA expression exhibiting unique temporal, spatial and tissue-specific profiles and have been associated with several types of hematological and solid malignant tumors (11,12), including CRC (2,13-18), emphasizing their potential for diagnostic and prognostic applications for human malignancies and their classification (19,20). In this context, various studies, including our previous research (2), have recently highlighted the utility of circulating miRNAs as stable biomarkers for multi-stage carcinogenesis in solid tumors and several other malignancies (21-24).

Despite the fact that reverse transcription-quantitative polymerase chain reaction (RT-qPCR) (2,25) and microarrays (26) are indispensable in deciphering the molecular mechanism underlying individual miRNA gene function and its expression level, comprehensive miRNOME analysis employing a high-throughput next-generation sequencing (NGS) approach on clinical samples is required to elucidate the clinical significance of miRNAs in CRC (27). The present study employed a deep sequencing approach for screening miRNA DEGs, and evaluating their potential as early diagnostic circulating biomarkers for CRC in clinical plasma and tissue samples from a Saudi patient population. A network analysis was performed based on a bioinformatics approach to identify associations between CRC, miRNA and enrichment of specific signaling transduction pathways and biological functions. This study successfully identified a number of interesting novel and known miRNAs that are differentially expressed in CRC context and may possess diagnostic and prognostic importance at the onset, development and progression of CRC.

Subjects and methods

Ethical statement and patients. All patients recruited into this study (CRC patients, n=20; healthy neoplasm-free controls, n=20) were subjected to a rigorous selection and eligibility criteria, as detailed elsewhere by our group (2). Patients were excluded if they had undergone chemotherapy or radiotherapy before blood sampling, or clinically diagnosed with familial adenomatous polyposis or hereditary non-polyposis CRC. The tumor-node-metastasis (TNM) classification system accepted by the Union for International Cancer Control (UICC 2006) (28) was employed for staging of the malignant tumors. Signed informed consent forms, in compliance with the ethical principles of the World Medical Association (Declaration of Helsinki), were collected from all the subjects that participated in this study for the use of their blood and tissue samples. Research protocols conducted in this project were approved by the Institutional Medical Ethics Review Board of the College of Applied Medical Sciences affiliated to King Saud University (Riyadh, Kingdom of Saudi Arabia).

Tissue isolation and RNA extraction. Paired fresh CRC tissue specimens and their adjacent non-cancerous normal mucosa were collected from 20 patients who underwent surgical resection of tumors by surgeons and were subsequently examined by pathologists. All CRC tissues were histologically confirmed to be adenocarcinomas of the colon. Tissue specimens were collected in RNAlater® RNA Stabilization Reagent (Qiagen GmbH) tubes, snap-frozen in liquid nitrogen, and stored at -80˚C until further analysis. Total RNA, including miRNAs, was extracted from 50-100 mg of cryo-preserved tissues by use of TRIzol® reagent (Invitrogen; Thermo Fisher Scientific, Inc.) as described in the manufacturer's protocol. In order to maximize RNA yield, a homogenization step was carried out by use of a TissueLyser LT with 5-mm stainless-steel beads (Qiagen GmbH).

Plasma preparation and circulating RNA extraction. Blood was drawn from the above mentioned 20 CRC patients recruited into this study, as well as 20 age-matched neoplasm-free healthy subjects. The cancer-free status of the blood-donating healthy control subjects was confirmed a priori based on the negative results of health examinations, including blood test, chest X-ray, abdominal ultrasound examination, fecal occult-blood test (FOBT), rectal touch, CT scan and colonoscopy. None of these controls had been previously diagnosed with any type of malignancy. Between 8 to 10 ml of peripheral whole blood were collected from each individual into BD Vacutainer® blood collection tubes (Becton, Dickinson and Company; EDTA spray-coated). Plasma was fractionated from whole blood samples according to the procedure previously described by our group (2). Freshly drawn whole blood was processed for plasma fractionation and the obtained plasma was frozen at -80˚C within 4 h. All plasma samples were spectrophotometrically analyzed to be free from haemoglobin (29). Hemolyzed plasma samples were excluded from further analysis. All centrifugation steps were conducted at 4˚C. Whole blood was centrifuged at low speed (1,700 x g) for 10 min. The obtained cloudy supernatant was transferred to a fresh tube and centrifuged at 2,000 x g for 10 min. Subsequently, the obtained supernatant was centrifuged at 12,000 x g for 10 min to pellet any remaining cellular debris. Circulating RNA extraction was essentially conducted on 1 ml plasma volume using the Plasma/Serum Circulating and Exosomal RNA Purification kit (Slurry Format, Norgen Biotek) following the manufacturer's instructions. The resulting eluate was subjected to an additional concentration step by the use of the RNA Clean-Up and Concentration kit (Norgen Biotek) using 20 μl elution buffer to collect the RNA.

RNA quality. The concentration of the isolated RNA from tissue specimens was assessed by measuring the optical density at 260 nm (OD260) and 280 nm (OD280) using a Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Inc.) and a Nano-Drop 1000 spectrophotometer (PEQLAB Biotechnologie GmbH), whereas the quality of the RNA purified from plasma could only be assessed by PCR amplification curves and

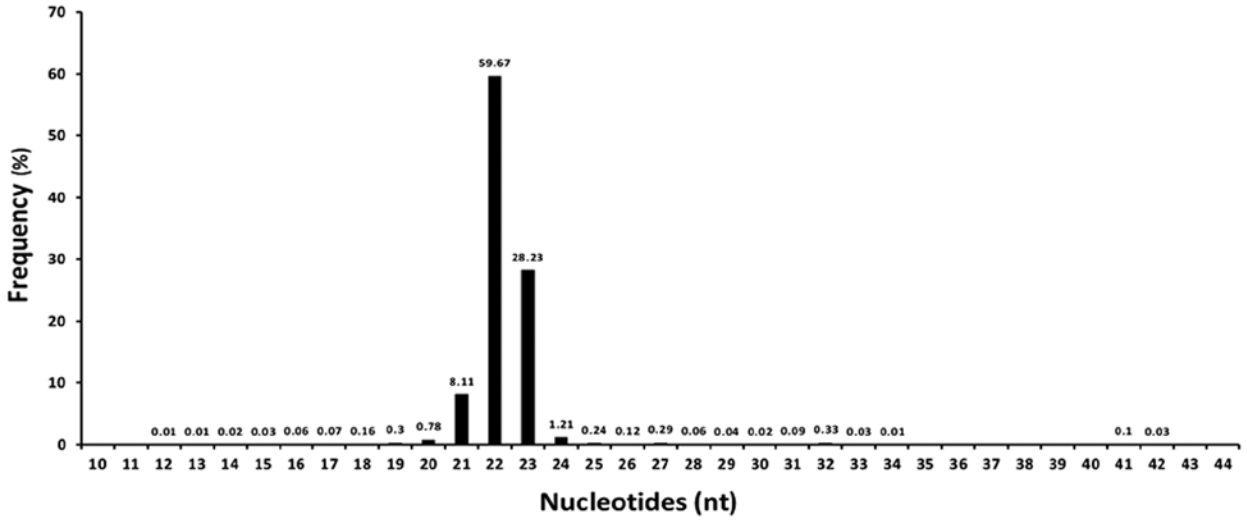

Figure 1. Total nucleotide length of miRNAs detected by deep sequencing. Length distribution of tags. The horizontal coordinate is tag lengths and the vertical coordinate is percent of tags. Sequences with 21-23 nt in length occupied most parts in all detected sequences. miRNA, microRNA.

efficiencies, due to the absence of ribosomal RNA. The purity of RNA was evaluated by the ratio of the absorbance at OD260 to OD280. Analysis of the RNA integrity number (RIN) was conducted by use of the Agilent 2100 Bioanalyzer with RNA 6000 series II Nano LabChip analysis kit (both from Agilent Technologies, Inc.). The 2100 Bioanalyzer generates numerical RIN values. RIN is an incremental scale which spans from 0 to 10, whereas increasing RNA integrity correlates with increasing RIN value. All total RNA samples extracted from CRC and adjacent neoplasm-free mucosal tissues were of high integrity as judged by the obtained RIN values of ≥8.0.

miRNA sequencing and sequence analysis. Small RNA libraries were generated from the purified RNA using the Illumina Small RNA v1.5 Sample Preparation kit (Illumina, Inc.), according to the manufacturer's instructions. Briefly, each RNA sample was size fractionated, and 18-30 nt RNA was isolated and purified (6-8 μl). After 5' and 3' adaptor ligation, RNA was reverse-transcribed and amplified using 14 PCR cycles of 98˚C for 10 sec and 72˚C for 15 sec to generate small RNA libraries. Then, the libraries were validated by loading 1 μl onto the Agilent Technologies 2100 Bioanalyzer to check the size (average molecule length) and purity, and by quantitative PCR using EvaGreen® dye (Jena Bioscience) to verify the concentration. Validated libraries were subsequently sequenced on an Illumina HiSeq 2000 (Illumina, Inc.). Qualified libraries were amplified on cBot to generate the cluster on the flowcell using Illumina TruSeq SE Cluster kit V3-cBot-HS (Illumina, Inc.). Then, the amplified flowcell was single-end sequenced on the HiSeq 2000 system using Illumina's TruSeq V3-SBS-HS kit. Read lengths of 50 bp were the most common sequencing strategy.

Bioinformatic analysis. The obtained miRNA sequences were handled and analyzed by miRWalk (http://mirwalk.uni-hd.de) and miRBase (http://www.mirbase.org). Moreover, pathway analyses and Gene Ontology (GO) were performed by the use of multiple publicly available tools and algorithms that identify pathways and GO according to sequencing datasets with the objective of pinpointing cognate miRNA target genes and their corresponding pathways. Subsequently, GO pathway enrichment analysis of biological processes (hierarchical level 3, 4 and 5) was applied on the set of predicted miRNA target genes, whereby the human genome was employed as a reference set. In this context, the Z-test was utilized for the determination of GO terms or Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment. The tools included Ingenuity Systems Bioinformatics Software (Ingenuity Systems; Qiagen, Inc.) and Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/gds). Furthermore, the Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics tools (http://david.abcc.ncifcrf.gov) was employed to thoroughly examine the functional characterization results of the enriched cognate mRNA target genes via comparative analysis of biological functions, enrichment analysis of signaling pathways, and assignment of disease categories. Significance P-values for DEGs were calculated by Student's t-test.

Results

Clinicopathological characteristics of CRC patients. A total of 20 patients diagnosed with CRC were recruited in the study, and each provided 3 samples (CRC tissues, T; adjacent normal mucosa tissues as control tissues, CT; and CRC plasma, P). Additionally, 20 healthy neoplasm-free individuals were recruited and provided control plasma (CP) samples. All participants were ≥18 and ≤80 years of age. The clinicopathological parameters of the 20 patients (9 men, 11 women) recruited into the study and the general inclusion criteria for the patients have been detailed in our previously published study (2).

The average age of the subjects was 61±10.6 years. None of the participants exhibited any evidence of other disease complications. In addition, 5 participants (25%) had stage I CRC, 3 (15%) had stage II, 7 (35%) had stage III, while the remaining 5 (25%) were diagnosed with stage IV of the disease.

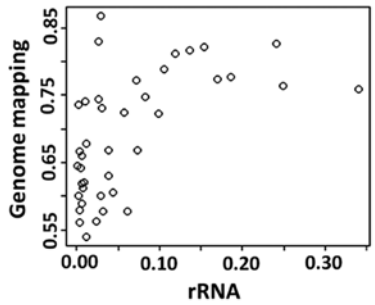

Figure 2. rRNA and genome mapping contribution in small RNA sequencing samples.

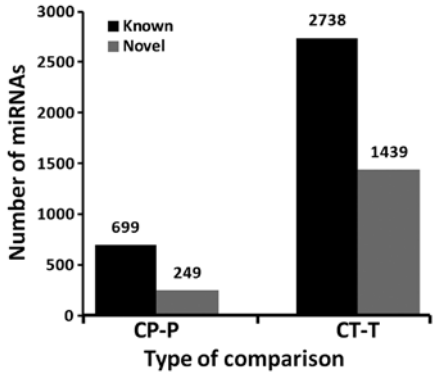

Figure 3. Known and novel miRNA DEGs found in CRC plasma and tissues. miRNA, microRNA; DEGs, differentially expressed genes; CRC, colorectal cancer.

In terms of tumor location, 7 patients had tumors localized to the rectum, 5 to the distal colon and 8 to the proximal colon. Histological examination revealed that all of the diagnosed CRC tumors were of the adenocarcinoma type; 16 adenocarcinoma, 2 mucous adenocarcinoma and 2 signet ring cell.

Deep sequencing of the small RNA transcriptome of 20-paired CRC tissue and plasma samples. In this key miRNA screening step, Illumina-based HiSeq 2000 high-throughput sequencing was conducted to identify differential expression patterns of miRNAs between cancerous (T) and adjacent non-cancerous control (CT) colorectal tissues, as well as colorectal plasma (P) relative to control plasma (CP). Approximately 22,022,602 effective reads were obtained of which 21,944,812 (99.65%) were clean high-quality reads. In these reads, the most abundant length was 22 nt (59.67%) followed by 23 nt (28.23%) and 21 nt (8.11%) length classes (Fig. 1).

The contribution of rRNA in the reads was <0.35%, with indications of negligible degradation. The genome mapping rate was >0.55, which is indicative of very little impurity and high quality (Fig. 2).

Known and novel DEGs in CRC tissues and plasma. DEGs in CRC tissues were greatly significantly higher relative to CRC plasma. In CP-P comparison, a total of 948 DEGs were found, of which 699 were known miRNAs and 249 were novel miRNAs. By contrast, a total of 4,177 DEGs were identified in the CT-T comparison of which 1,439 were known and 2,738 were novel miRNAs (Fig. 3).

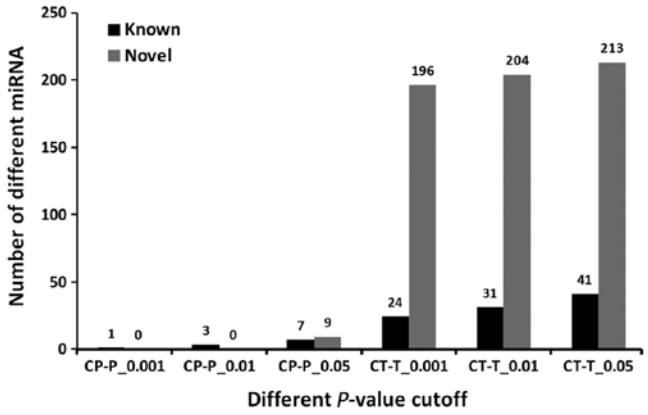

Figure 4. Statistically significant DEGs at different P-values at fold-change cut-off value of >2.0. DEGs, differentially expressed genes.

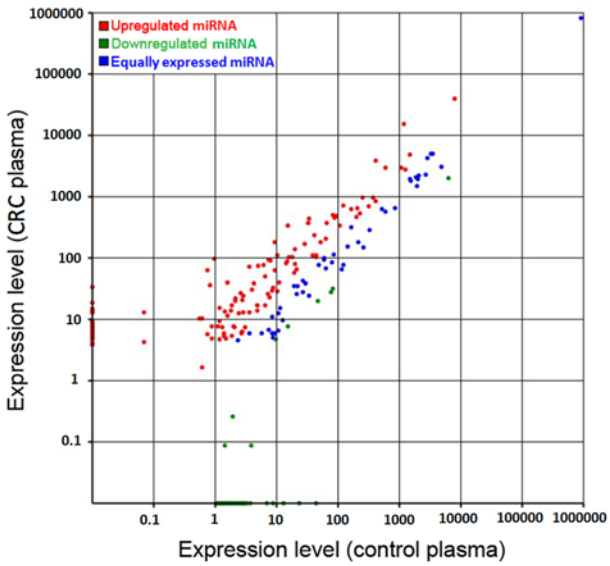

Figure 5. A representative scatter plot for the miRNA DEGs in CRC plasma as compared with control plasma from healthy subjects. Red dots, upregulated miRNAs; green dots, downregulated miRNAs; blue dots, equally expressed miRNAs. miRNA, microRNA; DEGs, differentially expressed genes; CRC, colorectal cancer.

In order to focus on the most significant DEGs, we applied more stringent selection parameters filtering out DEGs with a fold-change cut-off value of <2.0, and subsequently examined the resulting data at three different significance levels: P<0.05, P<0.01, and P<0.001 (Fig. 4).

While we identified only one known (*miR-182-5p*) and no novel DEGs at the most stringent significance level (P<0.001) in the P-CP comparison, 3 and none were detected at P<0.01, while 7 and 9 at the P<0.05 level, respectively. In the T-CT comparison, the results revealed 24 known and 196 novel miRNA DEGs (P<0.001), 31 and 204 (P<0.01), 41 and 213 (P<0.05), respectively (Fig. 4). A representative scatter plot for the miRNA DEGs in CRC plasma as compared with control plasma obtained from healthy subject is depicted in Fig. 5.

Results of novel and known top 10 miRNA DEGs in CRC tissues (T) compared with neighboring neoplasm-free control tissues (CT), and in CRC plasma (P) compared with

Table I. Top 10 miRNA DEGs in CRC tissues (T) compared with neighboring neoplasm-free control tissues (CT).

miRNA name	No. of reads (CRC/control)	Expression ratio (CRC/control)	P-value
hsa-miR-125a-5p	89,382.88/295,517	0.275569	0.043351
hsa-miR-145-5p	1,536.69/9,608	0.160113	0.00282
hsa-miR-145-3p	1,585.5/7,008	0.200085	0.015226
hsa-miR-133a-3p	295.25/2,561	0.088228	0.001124
hsa-miR-204-5p	221.38/983	0.206761	0.032151
hsa-miR-143-5p	127.32/745	0.132749	0.014126
hsa-miR-504-5p	20.69/163	0.105502	0.01235
hsa-miR-490-3p	5/158	0.025243	3.80E-05
hsa-miR-1224-5p	19.56/125	0.118523	0.024104
hsa-miR-548ba	13.56/127	0.087614	0.007696

miRNA, microRNA; DEGs, differentially expressed genes; CRC, colorectal cancer.

Table II. Differentially expressed miRNA DEGs in CRC plasma (P) compared with control plasma (CP).

miRNA name	No. of reads (CRC/control)	Expression ratio (CRC/control)	P-value
hsa-miR-182-5p	66,074.5/19,600.65	0.219551	8.95E-05
hsa-miR-501-5p	16.167/0	0.000954	0.006364
hsa-miR-424-5p	0/51.18	347.6291	0.008168
hsa-let-7b-5p	17,472.67/9,778.35	0.340973	0.012261
hsa-miR-122-5p	8,685.67/1,981.77	0.16204	0.017114
hsa-miR-708-3p	7.33/0	0.001919	0.024503
TAACTGATCTCTACTTTGGCT	32.67/0	0.000131	0.025301
CTGGGACCAGACTTCCTCAGGCAC	11.5/0	0.001612	0.031381
ACGGACCGGGCATTAGGGGCATGC	11/0	0.001685	0.032149
GATGAGGACTTGGAGACACT	20.67/0	0.000207	0.032998

miRNA, microRNA; DEGs, differentially expressed genes; CRC, colorectal cancer.

control plasma (CP) from healthy subjects are presented in Tables I and II, respectively. While, it is evident that hsa-miR-125a-5p is the most significantly downregulated gene in T (89,382.88 reads) compared with CT (295,517 reads) (Table I), *hsa-miR-182-5p* is the most upregulated gene in P (66,074.5 reads) compared with CP (19,600.65 reads) (Table II).

Detailed GO data analysis in comparison of CT-T and CP-P. Based on the top 10 miRNA DEGs in the comparison of CT-T and CP-P, the biological processes involved in CRC development and progression were globally assessed by the use of the web-based GO analysis bioinformatics resources tool of DAVID v6.8 (https://david.ncifcrf.gov/). GO is an international standardized classification system for gene function, which supplies a set of controlled vocabulary to comprehensively describe the property of genes and gene products. There are three ontologies in GO: Molecular function, cellular component and biological process. The basic unit of GO is GO-term representing gene product properties, each of which belongs to one defined type of ontology. After utilizing the blast application to align the targets of the top 10 miRNA DEGs to the database we employed Hypergeometric Distribution Mathematic Model to get the P-value of GO. If the P-value was <0.05, the targets are enriched in the GO, and were chosen among the top 5 of major detects (biological process, BP; cellular component, CC; molecular function, MF) of the GO (Fig. 6).

The diagrams of GO analyses included GO terms rooted to BPs significantly regulated in CRC tissues vs. control non-cancerous tissues (T-CT) and CRC plasma vs. plasma from health donors (P-CP). The major mRNA gene targets were enriched in BP. T-CT and P-CP comparison was conducted and target genes were identified to be involved in cell proliferation, differentiation and development. It is worth noting that several GO terms were observed to be co-modulated in CRC plasma and tissues. For instance, 5 GO terms (i.e., regulation of biosynthesis process, cellular metabolic process, transcription and intracellular signaling cascade, regulation of cell differentiation, regulation of cell development) were significantly modulated in CRC clinical samples; thus, indicating common cellular biological effects in response to the carcinogenesis process in CRC patients (Fig. 6).

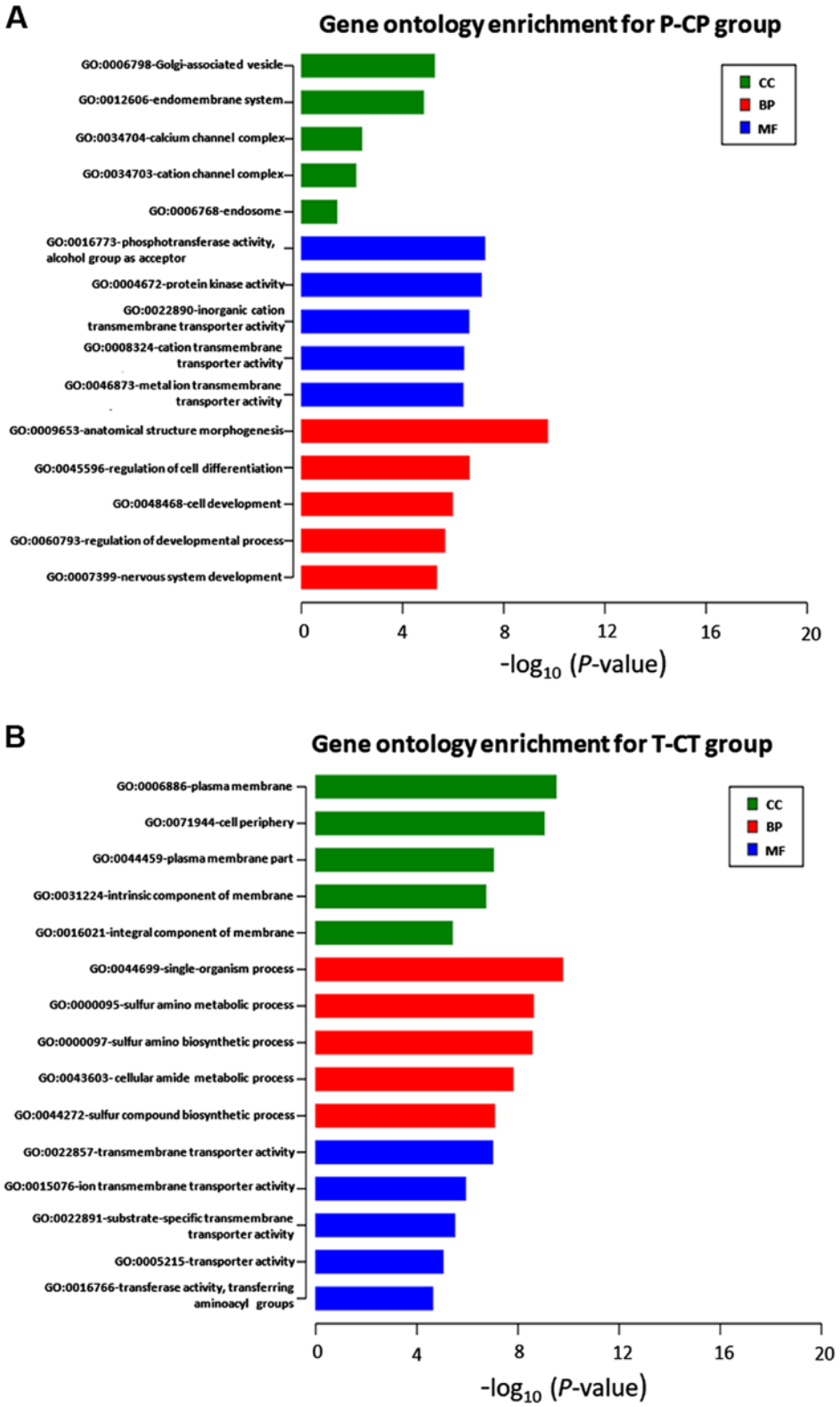

Figure 6. GO analysis. The GO enrichment analysis for the top 10 expressed miRNA targets. The x-axis denotes the -$\log_{10}$(P-value) and the P-value is generated by hypergeometric distribution. The y-axis denotes the GO term. Enriched (P>0.05) BPs were (A) anatomical structure morphogenesis, regulation of cell differentiation, and cell development for CP-P, and (B) single-organism process, sulfur amino acids biosynthetic and metabolic processes for CT-T comparison. GO, Gene Ontology; miRNA, microRNA; CC, cellular component; BP, biological process; MF, molecular function.

Comparison of detailed KEGG data analysis of T-CT and P-CP. Next, we conducted KEGG (https://www.genome.jp/kegg/) pathway analysis using KEGG release 76.0 application, and the same enrichment rule was applied as in GO; i.e., when the P-value and the Q-value was <0.05, the KEGG pathway under study was considered to be enriched and included in the results for both T-CT and P-CP comparisons. Mitogen-activated protein kinase (MAPK) and Wnt signaling

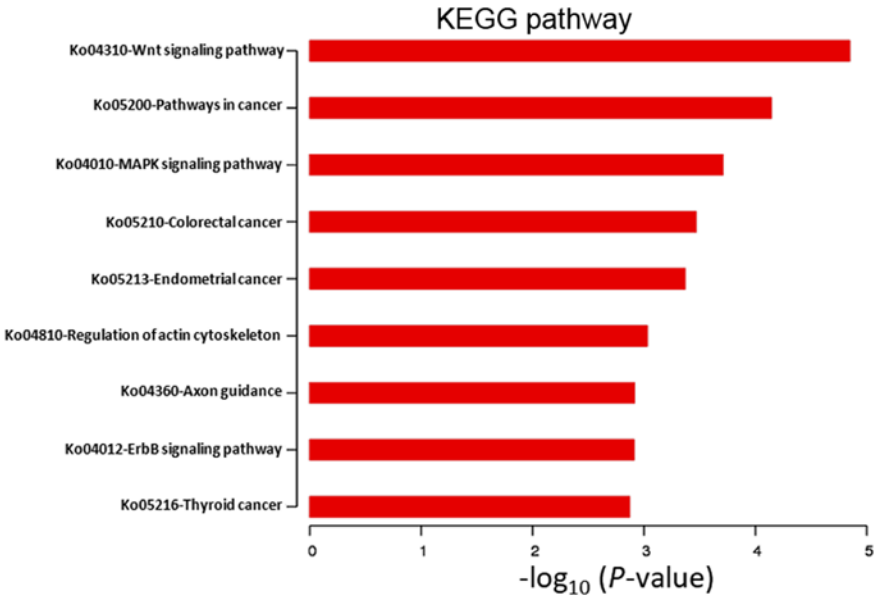

Figure 7. KEGG pathway analysis.

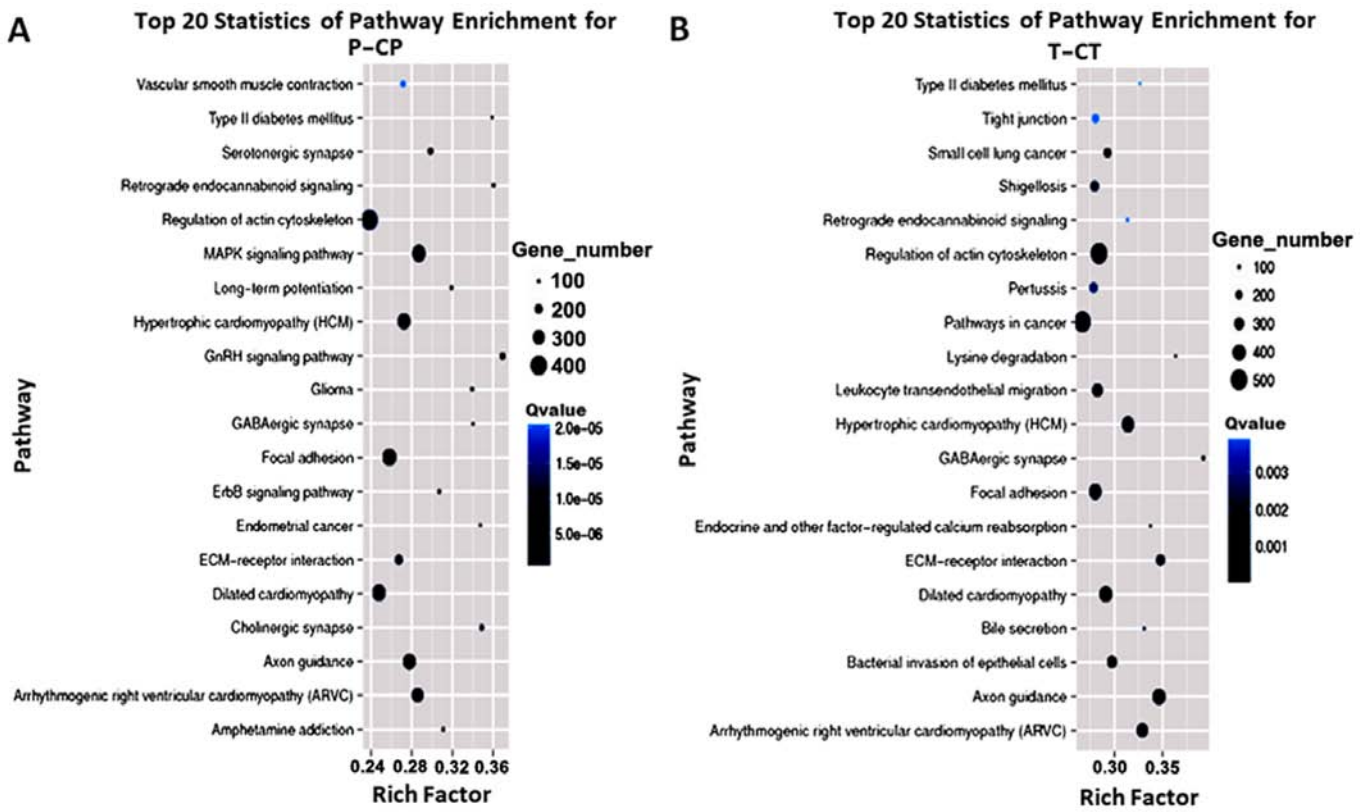

Figure 8. KO annotation and pathway enrichment in CRC samples. (A) CP-P comparison. (B) CT-T comparison. The Rich Factor is the ratio of DEG numbers annotated in this pathway term to all gene numbers annotated in this pathway term. Greater Rich Factor means greater intensiveness. Q-value is corrected P-value ranging from 0.0 to 1.0. When the P-value is small it indicates a greater intensiveness (enrichment). KO, KEGG Orthology; CRC, colorectal cancer; DEG, differentially expressed gene.

pathways were found to be enriched, among others, in CRC plasma and tissue samples compared with control (Fig. 7).

The KEGG Orthology (KO) database is a database of molecular functions represented in terms of functional orthologs. A functional ortholog is manually defined in the context of KEGG molecular networks, including KEGG pathway maps and KEGG modules (https://www.genome.jp/kegg/ko.html). In Fig. 8A and B, KO annotations for CT-T and CP-P are represented by dot graphs displaying the top 20 pathway terms enriched in CRC samples.

It can be noted that MAPK signaling pathway and regulation of actin cytoskeleton are amongst the most significantly

enriched pathways in CRC plasma samples when compared with plasma of healthy individuals (Fig. 8A). Similarly, it can be observed that the regulation of actin cytoskeleton, in addition to cancer signaling pathways, is amongst the most significantly enriched pathways in CRC tissue samples compared with neoplasm-free controls (Fig. 8B).

Discussion

Being one of the most predominant types of cancers worldwide, CRC is estimated to cause 608,700 mortalities per year. The average 5-year survival rate could be significantly extended by 91.3% with early diagnosis of CRC patients at stage I compared with the patients diagnosed late at advanced stage IV (30). Thus, a great opportunity to defeat CRC lies within its early diagnosis. Currently practiced techniques, such as FOBT, X-ray barium enema, and colonoscopy examinations are beneficial for the early CRC detection. However, they are not as yet implemented in the frame of mass nationwide screening programs ascribed to their high cost and invasiveness (31,32). Moreover, currently employed circulating tumor biomarkers have poor sensitivity and specificity (33), justifying the urgent requirement for novel, non-invasive, highly sensitive and specific biomarkers to fill in this evident gap. Hence, the aim of this study was to conduct comprehensive screening of miRNA DEGs in plasma and tissues of Saudi CRC patients as a first step towards the assessment of their diagnostic utility as circulating biomarkers.

In this study, novel and known miRNA DEGs were identified and the top 10 DEGs in each comparison, i.e. CP-P and CT-T, are listed in Tables I and II. It is evident that *miR-182-5p* gene is significantly and differentially overexpressed in CRC plasma compared with control plasma and could potentially serve as a circulating biomarker for CRC. Therefore, further investigation to ascertain its diagnostic utility and index, and define its sensitivity and specificity is necessary. *miR-182-5p* is a member of the *miR-183/96/182* cluster and has been experimentally shown to exhibit expression levels correlating with various cancers (34-41).

However, the exact role that *miR-182-5p* plays during carcinogenesis remains controversial as it has been reported to be dependent on the type of cancer. Our deep sequencing findings conform to other reports in hepatocellular carcinoma (HCC) (34), ovarian cancer (35), breast cancer (36), and melanoma (37), implicating *miR-182-5p* as an oncogene. By contrast, *miR-182-5p* has been advocated as an antioncogene (i.e., tumor suppressor) in the context of renal cell carcinoma (38,39) and glioblastoma (40,41). Recently, *miR-182-5p* has been reported to induce metastasis via the negative regulation of metastasis suppressor 1 *(MTSS1)* gene (34). Of note, increased expression of *miR-182-5p* heightens drug resistance in cisplatin-treated HCC cells by targeting the tumor protein 53-induced nuclear protein 1 *(TP53INP1)* (42).

KEGG pathway analyses provided evidence that MAPK and Wnt signaling pathways are significantly enriched, among others, in CRC plasma and tissue samples compared with control (Fig. 7). Wnt pathway is one of the central signaling transduction cascades regulating development and stemness, and its aberrant signaling has been specifically correlated with carcinogenesis. Aberrant Wnt signaling transduction has been conspicuously demonstrated for CRC, and also observed in the context of several other cancers (43). The MAPK pathway fine-tunes gene expression, cellular growth, and survival (44). Abnormal MAPK signaling may result in accelerated or unrestrained cell proliferation and avoidance of apoptosis in defiance of normal controls, eventually progressing to cancerous biomass (45). Impaired MAPK signaling is involved in diverse cancer types and takes place by means of numerous mode-of-actions, encompassing perturbed expression of genetic mutations and/or pathway receptors that culminate into activation of downstream cognate signaling effectors in the under supply of relevant stimuli (45,46). It has been reported that miRNAs may modify MAPK signaling transduction pathway in CRC through direct association with its vital gene components (47). In this context, a recent report by Slattery *et al* (47) revealed that miRNA:mRNA associations between miRNA and their cognate target MAPK genes take place via seed region matches. Results indicated that in case of direct binding inversely correlated differential expression between the miRNA and mRNA was observed. Direct associations (upstream of *ERK1/ERK2, JNK*, and *p38*) between *RASGRP3* and *PRKCB* with *miR-203a*, and *TGFBR1* with *miR-6071* and *miR-2117* have been reported (47).

In conclusion, the current study may be beneficial in highlighting cancer-specific signatures and potentially useful circulating biomarkers for the early diagnosis of CRC, such as miR-182-5p. Further research is necessary for determining their screen index values and diagnostic utility for CRC. The GO and KEGG network analyses conducted for miRNA:mRNA target gene interactions may assist in pinpointing perturbed miRNA regulatory networks that are implicated in CRC pathogenesis and carcinogenesis.

Acknowledgements

The authors would like to thank Fei Teng (BGI Tech Solutions Co., Ltd.) for the valuable assistance and feedback offered on the bioinformatics analyses. Thanks are also extended to Dr Omar Al-Obeed (Head of the Colorectal Research Center, Department of Surgery, College of Medicine, King Saud University) for the provision of clinical CRC blood and tissue samples.

Funding

The study was financially supported by the National Science, Technology and Innovation strategic Plan (NSTIP) strategic technologies programs (MAARIFAH), KACST (grant no. 11-MED1770-02) of the Kingdom of Saudi Arabia to YAAS.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

Conceptualization, MAMAS and YAAS; methodology, MAMAS and HKG; validation, MAMAS and YAAS; formal

analysis, YAAS and MAMAS; investigation, MAMAS, YAAS, KKA and HKG; resources, YAAS; writing and preparation of the original draft, MAMAS; writing, reviewing and editing, MAMAS and KKA; visualization, HKG and MAMAS; project administration, YAAS; funding acquisition, YAAS.

Ethics approval and consent to participate

Signed informed consent forms, in compliance with the ethical principles of the World Medical Association (Declaration of Helsinki), were collected from all the subjects that participated in this study, for the use of their blood and tissue samples. Research protocols conducted in this project were approved by the Institutional Medical Ethics Review Board of the College of Applied Medical Sciences affiliated to King Saud University (Riyadh, Kingdom of Saudi Arabia).

Patient consent for publication

Not applicable.

Competing interests

The authors declare no issues of interest, sponsorship or affiliation that could be considered as conflicts of interest in the context of this research.

References

1. The International Agency for Research on Cancer (IARC), Global Cancer Observatory, World Health Organization (WHO): Report on the latest estimates on the global burden of cancer. Date of release: September 12, 2018 in Geneva, Switzerland. https://www.iarc.fr/wp-content/uploads/2018/09/pr263_E.pdf. Accessed December 13, 2018.
2. Al-Sheikh YA, Ghneim HK, Softa KI, Al-Jobran AA, Al-Obeed O, Mohamed MA, Abdulla M and Aboul-Soud MA: Expression profiling of selected microRNA signatures in plasma and tissues of Saudi colorectal cancer patients by qPCR. Oncol Lett 11: 1406-1412, 2016.
3. Ibrahim EM, Zeeneldin AA, El-Khodary TR, Al-Gahmi AM and Bin Sadiq BM: Past, present and future of colorectal cancer in the Kingdom of Saudi Arabia. Saudi J Gastroenterol 14: 178-182, 2008.
4. Kumar R, Price TJ, Beeke C, Jain K, Patel G, Padbury R, Young GP, Roder D, Townsend A, Bishnoi S, *et al*: Colorectal cancer survival: An analysis of patients with metastatic disease synchronous and metachronous with the primary tumor. Clin Colorectal Cancer 13: 87-93, 2014.
5. O'Connell JB, Maggard MA and Ko CY: Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. J Natl Cancer Inst 96: 1420-1425, 2004.
6. Lee RC, Feinbaum RL and Ambros V: The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75: 843-854, 1993.
7. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR and Ruvkun G: The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 403: 901-906, 2000.
8. Herranz H and Cohen SM: MicroRNAs and gene regulatory networks: Managing the impact of noise in biological systems. Genes Dev 24: 1339-1344, 2010.
9. Friedman RC, Farh KK, Burge CB and Bartel DP: Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19: 92-105, 2009.
10. Kozomara A and Griffiths-Jones S: miRBase: Integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39: D152-D157, 2011.
11. Hannafon BN, Sebastiani P, de las Morenas A, Lu J and Rosenberg CL: Expression of microRNA and their gene targets are dysregulated in preinvasive breast cancer. Breast Cancer Res 13: R24, 2011.
12. Miko E, Czimmerer Z, Csánky E, Boros G, Buslig J, Dezso B and Scholtz B: Differentially expressed microRNAs in small cell lung cancer. Exp Lung Res 35: 646-664, 2009.
13. Akao Y, Nakagawa Y and Naoe T: MicroRNA-143 and -145 in colon cancer. DNA Cell Biol 26: 311-320, 2007.
14. Bandrés E, Cubedo E, Agirre X, Malumbres R, Zárate R, Ramirez N, Abajo A, Navarro A, Moreno I, Monzó M, *et al*: Identification by real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. Mol Cancer 5: 29-38, 2006.
15. Michael MZ, O' Connor SM, van Holst Pellekaan NG, Young GP and James RJ: Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol Cancer Res 1: 882-891, 2003.
16. Peng J, Xie Z, Cheng L, Zhang Y, Chen J, Yu H, Li Z and Kang H: Paired design study by real-time PCR: miR-378* and miR-145 are potent early diagnostic biomarkers of human colorectal cancer. BMC Cancer 15: 158-165, 2015.
17. Schepeler T, Reinert JT, Ostenfeld MS, Christensen LL, Silahtaroglu AN, Dyrskjøt L, Wiuf C, Sørensen FJ, Kruhøffer M, Laurberg S, *et al*: Diagnostic and prognostic microRNAs in stage II colon cancer. Cancer Res 68: 6416-6424, 2008.
18. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, Yuen ST, Chan TL, Kwong DL, Au GK, *et al*: MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. JAMA 299: 425-436, 2008.
19. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, *et al*: MicroRNA expression profiles classify human cancers. Nature 435: 834-838, 2005.
20. Waldman SA and Terzic A: MicroRNA signatures as diagnostic and therapeutic targets. Clin Chem 54: 943-944, 2008.
21. Tsang JC and Lo YM: Circulating nucleic acids in plasma/serum. Pathology 39: 197-207, 2007.
22. Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K, Guo J, Zhang Y, Chen J, Guo X, *et al*: Characterization of microRNAs in serum: A novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res 18: 997-1006, 2008.
23. Huang Z, Huang D, Ni S, Peng Z, Sheng W and Du X: Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. Int J Cancer 127: 118-126, 2010.
24. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, *et al*: Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci USA 105: 10513-10518, 2008.
25. Liu X, Xu X, Pan B, He B, Chen X, Zeng K, Xu M, Pan Y, Sun H, Xu T, *et al*: Circulating miR-1290 and miR-320d as novel diagnostic biomarkers of human colorectal cancer. J Cancer 10: 43-50, 2019.
26. Motoyama K, Inoue H, Takatsuno Y, Tanaka F, Mimori K, Uetake H, Sugihara K and Mori M: Over- and under-expressed microRNAs in human colorectal cancer. Int J Oncol 34: 1069-1075, 2009.
27. Wang X, Chen L, Jin H, Wang S, Zhang Y, Tang X and Tang G: Screening miRNAs for early diagnosis of colorectal cancer by small RNA deep sequencing and evaluation in a Chinese patient population. Onco Targets Ther 9: 1159-1166, 2016.
28. UICC: TNM Classification of Malignant Tumours. Sobin LH and Wittekind Ch (eds). 6th edition, 2006.
29. Kirschner MB, Kao SC, Edelman JJ, Armstrong NJ, Vallely MP, van Zandwijk N and Reid G: Haemolysis during sample preparation alters microRNA content of plasma. PLoS One 6: e24145, 2011.
30. Jeong SY, Chessin DB, Schrag D, Riedel E, Wong WD and Guillem JG: Re: Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. J Natl Cancer Inst 97: 1705-1707, 2005.
31. Launoy G, Smith TC, Duffy SW and Bouvier V: Colorectal cancer mass-screening: Estimation of faecal occult blood test sensitivity, taking into account cancer mean sojourn time. Int J Cancer 73: 220-224, 1997.
32. Hassan C, Pickhardt PJ, Laghi A, Kim DH, Zullo A, Iafrate F, Di Giulio L and Morini S: Computed tomographic colonography to screen for colorectal cancer, extracolonic cancer, and aortic aneurysm: Model simulation with cost-effectiveness analysis. Arch Intern Med 168: 696-705, 2008.

33. Carpelan-Holmström M, Louhimo J, Stenman UH, Alfthan H and Haglund C: CEA, CA 19-9 and CA 72-4 improve the diagnostic accuracy in gastrointestinal cancers. Anticancer Res 22: 2311-2316, 2002.
34. Wang J, Li J, Shen J, Wang C, Yang L and Zhang X: MicroRNA-182 downregulates metastasis suppressor 1 and contributes to metastasis of hepatocellular carcinoma. BMC Cancer 12: 227, 2012.
35. Xu X, Ayub B, Liu Z, Serna VA, Qiang W, Liu Y, Hernando E, Zabludoff S, Kurita T, Kong B, *et al*: Anti-miR182 reduces ovarian cancer burden, invasion, and metastasis: An in vivo study in orthotopic xenografts of nude mice. Mol Cancer Ther 13: 1729-1739, 2014.
36. Li P, Sheng C, Huang L, Zhang H, Huang L, Cheng Z and Zhu Q: MiR-183/-96/-182 cluster is up-regulated in most breast cancers and increases cell proliferation and migration. Breast Cancer Res 16: 473, 2014.
37. Segura MF, Hanniford D, Menendez S, Reavie L, Zou X, Alvarez-Diaz S, Zakrzewski J, Blochin E, Rose A, Bogunovic D, *et al*: Aberrant miR-182 expression promotes melanoma metastasis by repressing FOXO3 and microphthalmia-associated transcription factor. Proc Natl Acad Sci USA 106: 1814-1819, 2009.
38. Xu X, Wu J, Li S, Hu Z, Xu X, Zhu Y, Liang Z, Wang X, Lin Y, Mao Y, *et al*: Downregulation of microRNA-182-5p contributes to renal cell carcinoma proliferation via activating the AKT/FOXO3a signaling pathway. Mol Cancer 13: 109, 2014.
39. Wang X, Li H, Cui L, Feng J and Fan Q: MicroRNA-182 suppresses clear cell renal cell carcinoma migration and invasion by targeting IGF1R. Neoplasma 63: 717-725, 2016.
40. Kouri FM, Ritner C and Stegh AH: miRNA-182 and the regulation of the glioblastoma phenotype - toward miRNA-based precision therapeutics. Cell Cycle 14: 3794-3800, 2015.
41. Kouri FM, Hurley LA, Daniel WL, Day ES, Hua Y, Hao L, Peng CY, Merkel TJ, Queisser MA, Ritner C, *et al*: miR-182 integrates apoptosis, growth, and differentiation programs in glioblastoma. Genes Dev 29: 732-745, 2015.
42. Qin J, Luo M, Qian H and Chen W: Upregulated miR-182 increases drug resistance in cisplatin-treated HCC cell by regulating TP53INP1. Gene 538: 342-347, 2014.
43. Zhan T, Rindtorff N and Boutros M: Wnt signaling in cancer. Oncogene 36: 1461-1473, 2017.
44. Knight T and Irving JA: Ras/Raf/MEK/ERK pathway activation in childhood acute lymphoblastic leukemia and its therapeutic targeting. Front Oncol 4: 160, 2014.
45. Santarpia L, Lippman SL and El-Naggar AK: Targeting the MAPK-RAS-RAF signaling pathway in cancer therapy. Expert Opin Ther Targets 16: 103-119, 2012.
46. Chappell WH, Steelman LS, Long JM, Kempf RC, Abrams SL, Franklin RA, Bäsecke J, Stivala F, Donia M, Fagone P, *et al*: Ras/Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR inhibitors: Rationale and importance to inhibiting these pathways in human health. Oncotarget 2: 135-164, 2011.
47. Slattery ML, Mullany LE, Sakoda LC, Wolff RK, Samowitz WS and Herrick JS: The MAPK-signaling pathway in colorectal cancer: Dysregulated genes and their association with microRNAs. Cancer Inform 17: 1176935118766522, 2018.