Abstract. Colorectal cancer (CRC) is one of the prominent causes of cancer related deaths because, in part, there is not an early, non-invasive, effective detection strategy. Circulating microRNAs (miRNAs) have been proposed as potential non-invasive biomarkers for CRC. In this study, we evaluated the miRNA profile in sixteen CRC tissues by Next-Generation-Sequencing and compared the circulating expression levels of 22 miRNAs among 45 CRC, 14 hyperplastic polyps, 11 advanced adenoma patients and 45 control subjects, by reverse transcription-quantitative PCR, to search for miRNAs which could be potential biomarkers. In total, nine of them represented 70% of total read counts (miR-10a-5p, miR-192-5p, miR-10b-5p, miR-22-3p, miR-26a-5p, miR-148a-3p, miR-181a-5p, miR-92a-3p and miR-143-5p).

In silico analysis found eight candidates to mature miRNAs. With respect to circulating miRNA, we found higher serum expression levels of miR-143-3p, miR-141-3p and miR-200c-3p in the CRC and adenoma groups compared with controls (P<0.002), and we also found significant higher levels of miR-141-3p and miR-200c-3p in serum of adenoma patients compared with the CRC group. In conclusion, the measurement of miRNAs in the blood could complement current screening methods for CRC and might provide new insights into mechanisms of tumorigenesis. miR-143-3p, miR-141-3p and miR-200c-3p could be interesting miRNAs to study as potential biomarkers for CRC.

Introduction

Globally, colorectal cancer (CRC) is the second most common cancer in women and the third most common in men (1). Although the treatments are improved for CRC, increasing the 5-year survival rate, the overall estimated death rate is still 50-60% (2). There is not a good method to diagnose for this cancer, because although tumor markers greatly improve it, the invasive nature of current procedures, as colonoscopy, limits their application. Identification of useful non-invasive biomarkers in order to facilitate the correct diagnosis and treatment is critical to improve patient survival.

MicroRNAs (miRNAs) are a type of small RNA (18-22 nucleotides, nt) that mediates post-transcriptional gene silencing by binding to mRNAs. The role of miRNA in carcinogenesis has been increasingly recognized; miRNAs affect many oncogenes and tumor suppressor genes. miRNA-induced deregulation in CRC has been well documented and for this reason, those could be exploited as biomarkers in CRC due to its high tissue specificity, stability and the differences in the expression level between normal and tumor tissues (3-6). The detection of miRNAs in serum samples has raised the possibility that they could be used as non-invasive biomarkers for different types of cancer (7-9).

In our study, we evaluated the miRNA profile in 16 samples of tumor tissue from patients with CRC by next-generation sequencing (NGS) and compared the expression levels of 22 miRNAs between CRC, hyperplastic polyps and adenoma patients with control subjects to search differences that could be useful for a better understanding of CRC carcinogenesis and could be potential biomarkers.

Patients and methods

Subjects. We included 45 CRC cases (24 colon cancer and 21 rectum cancer), 11 advanced adenomas, 14 hyperplastic polyps and 48 controls from a multicenter hospital-based case-control study conducted in Colombia. All cases were incident and
confirmed by histopathology, while controls were individuals without gastrointestinal symptoms attending the outpatient services of primary care units. Advanced adenomas were adenomas with size ≥1 cm, tubulovillous or villous adenomas or with high-grade dysplasia. Subjects were unrelated and their age ranged between 30 and 76 years. Neither cases nor controls had a personal history of other cancers and received neither chemotherapy nor radiotherapy. Trained health professionals collected blood samples and administered structured questionnaires on socio-economic characteristics and other risk factors, once each participant gave written informed consent. Tissues were collected during colonoscopy. This study was approved by the Ethics Committee of the Instituto Nacional de Cancerología, Bogotá, Colombia and by all the other Ethical Boards from participant health institutions upon request.

miRNAs isolation and quantification. Total RNA was extracted from 10 mg of tumor tissue using Trizol (Thermo Fisher Scientific, Whaltan, USA) and after that, miRCURY™ RNA Isolation Kit–Tissue (Exiqon, Copenhagen, Denmark) was used for miRNA isolation. Serum miRNAs were extracted from 200 µl using miRCURY RNA Isolation Kit-Biofluids (Exiqon, Copenhagen, Denmark). Concentration and quality of the samples was assessed using Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, California, USA) and Qubit 3.0 Fluorometer (Thermo Fisher Scientific, Whaltan, USA), following the manufacturer's instructions. Isolated RNA was stored at -80°C until use.

Libraries construction and sequencing. Libraries were constructed from 1 µg of extracted miRNAs using Trueq Small RNA kit (Illumina, San Diego, USA), following manufacturer's instructions. A pool of 16 libraries was obtained with an adjusted concentration of 10 nM in a final volume of 100 µl. The final concentration of the pooled samples was 2 nM. The sequencing was performed in a MiSeq, using MiSeq Reagent Kit v2 (Illumina, San Diego, USA), following the manufacturer's instructions. Isolated RNA was stored at -80°C until use.

Bioinformatics analysis of detected miRNAs. Bioinformatics analysis was made following the pipeline from Hackenberg et al and the tool miRanalyzer (10). The reads shorter than 17 nt and longer than 26 nt were discard. Selected reads were aligned against different databases, including the RefSeq to detect mRNA, the miRBase database (11) to detect mature miRNAs, and the GRCh37 human reference genome assembly to predict possible new miRNAs. The miRanalyzer tool also does a prediction over candidate miRNAs. We also took into account the following high confidence criteria to select novel miRNA sequence defined by miRBase (11).

Expression levels of selected miRNA in serum and tumor tissue. The Universal cDNA synthesis kit II (Exiqon®, Copenhagen, Denmark) was used for reverse transcription (RT), according to the manufacturer's instructions. The UniSp6 RNA Spike-in control was added during the cDNA synthesis. For quantitative PCR, the obtained cDNA was diluted 1:100 in nuclease free water and then 5 µl of cDNA along with 5 µl of Exilent SYBR Green master mix were transferred to a custom-made Pick & Mix microRNA PCR panel that included primers for 22 target miRNAs (Exiqon®, Copenhagen, Denmark). These miRNAs were selected based on the abundance of reads found in the tumor tissue libraries and others by literature review of the most abundant in serum from CRC patients. The hsa-miR were: 10a-5p, 10b-5p, 193-5p, 22-3p, 26a-5p, 148a-3p, 29a-3p, 143-3p, 486-5p, 141-3p, 27b-3p, 92a-3p, 22-3p, 200c-3p, 145-5p, 423-3p, 155-5p, 223-3p, 320a, 21-5p, 20a-5p and 221-3p. Amplification was performed in duplicates in a Light Cycler 480 Real-Time PCR System (Roche, Basel, Switzerland). Amplification curves were analyzed using the Roche LC software, both for determination of Ct values and for melting curve analysis. Normality of miRNA levels was assessed by Shapiro-Wilk test. Correlations between serum and tissue levels were made using Pearson test.

Data analysis and normalization of RT-PCR. Amplification efficiency was calculated by using Exiqon GenEx software specifically adapted to miRCURY LNA™ Universal RT microRNA PCR products, following the manufacturer's instructions. First, we made a pre-processing and normalization of our data in GenEx. Only miRNAs detected with Ct <37 were included for analysis. During quality control steps, samples with a >50% of missing data and miRNAs with <40% of valid data were excluded. The software selected the hsa-miR-92a-3p as reference gene for normalization. This normalization step correspond to the first delta Ct, namely delta to the normalization factor, of the 2^(-ΔΔCt) method (12). After that, the serum Ct from polyps, adenomas and CRC patients were converted to relative quantities, comparing to control group, and by this step the data was expressed completely as N=2^(-ΔΔCt) method (12). Expression data was converted to log2 scale for further analysis. Comparisons between groups in serum samples were done by Welch's ANOVA method adjusted by sex and differential expressed genes were identified based on a Bonferroni corrected P-value of <0.002 (alpha of 0.05/22 tests). Finally, we used Pearson correlation analysis of the miRNAs expression values found in serum and tumor of CRC samples.

Bioinformatics analysis of target genes for detected miRNAs and related biological pathways. In order to determine target genes of the identified miRNAs, we used DIANA-TarBase v7.0 (13), that predict molecular targets of miRNAs in coding sequences 3'UTR. Related biological pathways associated with target genes and miRNAs were made using the Kyoto Encyclopedia of Genes and Genomes (KEEG) (14).

Results

Libraries in tumor CRC

Expression pattern of known miRNAs. Sixteen tumor samples were assessed by NGS, five correspond to colon cancer and eleven to rectal cancer, 60% were from males and the mean age was 59.1 years. 763 known mature miRNAs were detected in the sixteen libraries by at least one alignment in miRBase (11). The read counts of the mature miRNAs from sixteen libraries were pooled. The known mature miRNAs showed a wide range of expression values spanning from 1 to 222455 read counts. 176 of 763 known miRNAs detected had only one read
count and 167 had more than 100 read counts. Nine miRNAs had expression levels above 2% and it represents 70.4% of the total read counts (hsa-miR: 10a-5p, 192-5p, 10b-5p, 22-3p, 26a-5p, 148a-3p, 181a-5p, 92a-3p and 143-5p) (Table IA).

**Prediction and expression levels of potential novel miRNAs.** In total, eight potential novel miRNAs with fuzzy Dicer pattern were identified in the libraries; no potential novel miRNA was detected with a perfect Dicer pattern. Seven candidates were present in four or more libraries. *In silico* analysis of these sequences against miRBase, led to identify that each of these candidates had a partial or total complementarity with mature miRNAs (Table II).

**Pathways related with most common found miRNAs.** We did a search of gene targets and pathways related of the nine most common miRNAs found by sequencing in the 'Colorectal cancer pathway' (hsa05210) in TarBase 7.0/KEGG (13). We found that all of these miRNAs had gene targets involved in different pathways related with CRC. The most common pathways involved are WNT, MAPK, PI3K/Akt, TGF-β, DCC, p53 and microsatellite instability (MSI).

**miRNAs levels in serum.** From the most abundant miRNAs detected by sequencing in tumor tissue, along with others thirteen differentially expressed in serum from CRC patients according to literature (Table IB), we selected 22 miRNAs to be analysed by RT-qPCR in the serum of patients including 45 CRC, 11 advanced adenomas, 14 hyperplastic polyps and 48 controls were enrolled in this study. Table III shows the distribution of age and gender according to phenotype. Pre-processing data, using Exiqon GenEx software, excluded six serum samples because they had >50% of missing data. Table I shows the percentage of total read counts, by NGS in sixteen samples, of the thirteen miRNAs selected according to literature.

From the twenty-two miRNAs selected to evaluate differences in their levels in serum between groups, pre-processing data excluded two miRNAs (miR-10a-5p and miR-221-5p) because they had <40% of valid data. The data was expressed completely as N=2-ΔΔCt method, miR-92-3p was selected by GenEx as reference gene for normalization. Among the remaining 19 miRNAs, we found significant higher serum expression levels of miR-143-3p, miR-141-3p and miR-200c-3p in the CRC and adenoma groups compared to controls by Mann-Whitney test with Bonferroni corrected P-value (P<0.002; Fig. 1). In addition, we also found significant higher levels of miR-141-3p and miR-200c-3p in serum of adenoma patients compared to CRC group (P<0.002).

Other miRNAs did not show statistical significant differences between CRC patients and controls. Serum miRNA levels between polyps patients and controls were very similar and their behavior were the same. We also assessed levels of miRNAs in the available twenty-two tumor tissues of CRC patients by RT-PCR. None of the correlations in levels of miR-143-3p, miR-141-3p and miR-200c-3p in CRC and adenoma groups compared to controls by Mann-Whitney test with Bonferroni corrected P-value (P<0.002) were significant (Fig. 1).

Discussion

In the present work, deep sequencing and RT-qPCR were used to analyze the expression levels of miRNAs in tumor tissue and serum from patients with CRC. By deep sequencing, this study detects 763 mature miRNAs in CRC tissues from sixteen patients. Of the nine most expressed miRNA in our samples, three, miR-10a-5p, -26a-5p and -92a-3p, have been reported that can act as oncomiRs (6,16-28), three, miR-192-5p, miR-148a-3p and miR-143-5p behave like anti-oncomiRs (6,17,29-37). With respect to miR-10a-5p, miR-221-3p and miR-181a-5p, their tumorigenesis role is inconsistent. These miRNAs with dual roles in carcinogenesis prove that many targets from many pathways can be
regulated by one miRNA and their effect on expression is very complex at cellular and tissue levels.

One advantage of miRNA studies by deep sequencing is that this technique allows the detection of novel miRNAs. Our analysis found eight new miRNA candidates. All candidates showed partial or total complementarity with mature miRNAs (scores 90-105) based on miRBase analysis. These sequences with some grade or total complementary could be produced by miRNAs bidirectional transcription and processing (38,39). It is possible that miRNA:miRNA duplex can be formed in the cell, operating in competition with each other. Further experimental studies are needed in order to assess the role of these miRNA candidates in colorectal carcinogenesis, before register them into public databases such as miRBase.

We found three miRNAs with significantly higher expression in serum of CRC patients vs. controls (i.e. miR‑143‑3p, miR‑141‑3p and miR‑200c‑3p) and two of them were more expressed in patients with adenomas compared those with CRC (i.e. miR‑141‑3p and miR‑200c‑3p). Interestingly, miR‑141‑3p and miR‑200c‑3p derive from the same precursor, miR‑8. Serum miR‑141‑3p and miR‑200c‑3p was found over expressed in CRC patients compared to controls, as previously reported (40‑47). We found that patients with adenomas had the highest serum levels of miR‑141‑3p and miR‑200c‑3p, compared to all the others (i.e., controls, polyps and CRC groups). These two miRNAs are good candidates for CRC screening and prevention, as they could be measured through minimally invasive procedures; nevertheless, further population‑based studies are needed for validation purposes.

The results seem to be contradictory. On the one hand, lower levels of miRNAs have been found in CRC (18,48‑53). On the other hand, our findings are consistent with Luo et al study (54) regarding higher levels of miR‑143‑3p in CRC patients. The role of miRNAs in cancer is very complex and depends of many particular factors and not alone cancer type. Differences found in various studies in circulating levels of miRNA can be related with ethnicity, gender and technical variance, but there are other confounding lifestyle factors, such as smoking, physical activity, etc., that are hardly verifiable and correctly taken into consideration (9).

Like Waters et al (55), we did not find any correlation of miRNA levels between serum and tissue of cancer patients.
The absence of this correlation could be attributed to the complex nature of the circulating miRNAs sources. It has been found that circulating tumor cells and exosomal release from tumor cells contribute to circulating miRNAs (56,57). Also, other factors such as the host immune response or inflammation, could modulate miRNAs circulation levels and cause these levels to be different from those of the tissues. Therefore, the levels of miRNA in circulation not only reflect what happens in the tumor tissue, but also show what happens in the whole human body.

Pathway analysis of the target genes of these miRNAs uncovered a significant number of genes involved in many CRC pathways, in accordance with reports highlighting that the hallmark feature of CRC is the hyperactivation of the WNT pathway, usually caused by mutations in the tumor suppressor gene APC (~75% of all tumors) (58), mutations in CTNNB1 (β-catenin), or in other Wnt signaling activators (59-61).

In conclusion, this study found 763 miRNAs in tissue from CRC and eight candidates to novel miRNAs. In serum, we found that three miRNAs, miR-141-3p, miR-143-3p and miR-200c-3p, were significantly higher in CRC vs. controls, and that two of them, miR-141-3p and miR-200c, were also significantly lower in CRC vs. adenomas. The measurement of miRNAs in the blood could complement current screening methods for CRC and might provide new insights into mechanisms of tumorigenesis and metastasis. However, the differences between studies highlight the necessity to perform further investigation.

Acknowledgements
The authors would like to thank Dr Ivan Lesende (Universidad de la Coruña-Spain) for his technical support with the library construction, Gustavo Hernández (Instituto Nacional de...
Circulating microRNAs as stable blood-based biomarkers: DIANA-TarBase v7.0

4.

3.

2.

1.


