

MicroRNA Classifier and Nomogram for Metastasis Prediction in Colon Cancer

Inès J. Goossens-Beumer¹, Remco S. Derr¹, Henk P.J. Buermans², Jelle J. Goeman³, Stefan Böhringer⁴, Hans Morreau⁵, Ulrich Nitsche⁶, Klaus-Peter Janssen⁶, Cornelis J.H. van de Velde¹, and Peter J.K. Kuppen¹

Abstract

Background: Colon cancer prognosis and treatment are currently based on a classification system still showing large heterogeneity in clinical outcome, especially in TNM stages II and III. Prognostic biomarkers for metastasis risk are warranted as development of distant recurrent disease mainly accounts for the high lethality rates of colon cancer. miRNAs have been proposed as potential biomarkers for cancer. Furthermore, a verified standard for normalization of the amount of input material in PCR-based relative quantification of miRNA expression is lacking.

Methods: A selection of frozen tumor specimens from two independent patient cohorts with TNM stage II–III microsatellite stable primary adenocarcinomas was used for laser capture microdissection. Next-generation sequencing was performed on small RNAs isolated from colorectal tumors from the Dutch cohort ($N = 50$). Differential expression analysis, comparing in metastasized and nonmetastasized tumors, identified prognostic miRNAs.

Validation was performed on colon tumors from the German cohort ($N = 43$) using quantitative PCR (qPCR).

Results: miR25-3p and miR339-5p were identified and validated as independent prognostic markers and used to construct a multivariate nomogram for metastasis risk prediction. The nomogram showed good probability prediction in validation. In addition, we recommend combination of miR16-5p and miR26a-5p as standard for normalization in qPCR of colon cancer tissue-derived miRNA expression.

Conclusions: In this international study, we identified and validated a miRNA classifier in primary cancers, and propose a nomogram capable of predicting metastasis risk in microsatellite stable TNM stage II–III colon cancer.

Impact: In conjunction with TNM staging, by means of a nomogram, this miRNA classifier may allow for personalized treatment decisions based on individual tumor characteristics. *Cancer Epidemiol Biomarkers Prev*; 24(1); 187–97. ©2014 AACR.

Introduction

miRNAs are endogenous noncoding RNAs with a sequence length of 19 to 25 nucleotides (miRBase v19), that are key posttranscriptional regulators of gene expression (1), and highly conserved between species (2). They have important regulatory functions in basic cellular processes, thereby affecting major biologic systems such as stemness and immunity (3). Depending on the level of complementarity of the miRNA to its target mRNA, the particular transcripts are targeted for transcriptional repression

or cleaved for degradation. miRNAs are predicted to be involved in regulation of more than 60% of the protein-coding RNAs (4). With their frequent location at fragile sites and genomic regions involved in cancer (5), it is not surprising that deregulation of miRNA expression, by both genetic and epigenetic mechanisms, is linked to tumorigenesis (6, 7), including the formation of metastasis (8).

Treatment modalities of patients with colon cancer are still essentially based on the anatomic extent of disease at diagnosis (TNM staging method of the American Joint Committee on Cancer/UICC). Even under the current classification, patient groups still show large differences in clinical outcome (9–11), especially for patients with locally advanced, but not distantly metastasized tumors (TNM stage II–III). Clinically, current criteria for identification of high-risk patients in this group of colon cancer, which would require adjuvant therapy, include clinical and histologic criteria T4 stage, poor differentiation, tumor perforation, or insufficient lymphadenectomy (9, 12). Biomarkers that can predict the development of distant recurrent disease would help identifying those patients with cancer that may benefit from adjuvant therapy and/or close follow-up in addition to surgery. The use of such biomarkers in conjunction with the TNM staging method, by means of a nomogram, may contribute to an approach of personalized treatment based on individual tumor characteristics. miRNA profiles, identified by comparing expression of miRNAs between patients with and without metastases in the follow-up, have been constructed for several cancer types (13).

¹Department of Surgery, Leiden University Medical Center, Leiden, the Netherlands. ²Department of Human Genetics/Leiden Genome Technology Center, Leiden University Medical Center, Leiden, the Netherlands. ³Biostatistics, Department for Health Evidence, Radboud University Medical Center, Nijmegen, the Netherlands. ⁴Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the Netherlands. ⁵Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands. ⁶Department of Surgery, Klinikum rechts der Isar, Technical University München, Munich, Germany.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

Corresponding Author: Dr. Peter J.K. Kuppen, Department of Surgery, K6-R, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, the Netherlands. Phone: 0031-7152-62309; Fax: 0031-7152-66750; E-mail: p.j.k.kuppen@lumc.nl

doi: 10.1158/1055-9965.EPI-14-0544-T

©2014 American Association for Cancer Research.

As discussed by de Krijger and colleagues (14), a similar profile for patients with colon cancer is warranted as the development of distant metastases mainly accounts for the high lethality rates of colon cancer.

Single miRNAs have multiple target genes (15), and target genes can be regulated by different miRNA transcripts. The tissue-specific (16) and cell type-specific (17) nature of miRNA expression, combined with the high degree of heterogeneity of colon tumors with regard to the amount of stroma and immune infiltrate, requires selection of target cells by microdissection (18), like laser capture microdissection (LCM). To be able to investigate expression of the whole microneome, and the possibility of detecting novel miRNAs, a next-generation sequencing (NGS) approach is recommended (19). Previous studies reporting on miRNA classifier profiles for metastasis prediction in colon or colorectal cancer (20–24) did not use the optimal approach of identification of miRNAs by comparison of their expression patterns between patients with and without metachronous metastasis using NGS in microdissected tumors. For PCR-based relative quantification of miRNA expression, a verified standard for normalization of the amount of input material is lacking. We therefore also aimed to assess which miRNA(s) would be suitable for usage as reference in miRNA-specific real-time PCR.

In this bicentric international study, we studied miRNA expression in two independent retrospective patient cohorts, both containing microsatellite stable tumors. Tumor tissue from the discovery cohort (Dutch colorectal cancer cohort, $N = 50$) and the validation cohort (German colon cancer cohort, $N = 43$) were subjected to LCM. High-depth NGS was used to study the predictive value of miRNAs for metachronous metastasis risk in the

discovery cohort. NGS data were technically validated with low-depth NGS and miRNA-specific stem–loop real-time quantitative PCRs (qRT-PCR). Results were validated in an independent cohort, using qRT-PCR. Multivariate survival analyses for distant recurrence-free period showed a HR of nearly 11 in the validation cohort. With our study, we identified and validated a miRNA expression classifier in the primary tumor, and proposed a nomogram capable of predicting metastasis risk in microsatellite stable colon cancer. In addition, we recommend a new standard for normalization of the amount of input material in PCR-based relative quantification of colon tissue–derived miRNA expression.

Materials and Methods

Study cohort

Patients for the discovery phase were selected from a colorectal cancer patient population ($N = 200$) who underwent surgery between 2001 and 2012 at the Leiden University Medical Center (Leiden, the Netherlands), with available snap-frozen primary tumor material. Microsatellite stability status was determined as described previously (25, 26). For the discovery study cohort, we included only histologically proven TNM stage II–III first primary colorectal carcinomas of the microsatellite stable phenotype, of patients who did not receive preoperative treatment. We selected 24 patients with distant metachronous metastases and matched these with patients without metastases. Matching criteria were TNM stage, age at surgery, gender, and location of the tumor. The discovery study cohort consisted of 50 patients (Table 1). All patients were under clinical follow-up surveillance according to the current Dutch National Guidelines.

Table 1. Associations of the patient cohorts with clinicopathologic parameters

	Discovery cohort				Validation cohort				<i>P</i>
	Total <i>N</i> = 50 (%)	M– <i>n</i> = 26 (%)	M+ <i>n</i> = 24 (%)	<i>P</i>	Total <i>N</i> = 43 (%)	M– <i>n</i> = 21 (%)	M+ <i>n</i> = 22 (%)	<i>P</i>	
Age at surgery, y				0.981				0.973	0.002
<50	6 (12.0)	3 (11.5)	3 (12.5)		2 (4.7)	1 (4.8)	1 (4.5)		
50–75	34 (68.0)	18 (69.2)	16 (66.7)		41 (95.3)	20 (95.2)	21 (95.5)		
≥75	10 (20.0)	5 (19.3)	5 (20.8)		0 (0.0)	0 (0.0)	0 (0.0)		
Gender				0.555				0.817	0.021
Female	27 (54.0)	13 (50.0)	14 (58.3)		13 (30.2)	6 (28.6)	7 (31.8)		
Male	23 (46.0)	13 (50.0)	10 (41.7)		30 (69.8)	15 (71.4)	15 (68.2)		
TNM stage				0.769				0.729	0.000
II	7 (14.0)	4 (15.4)	3 (12.5)		36 (83.7)	18 (85.7)	18 (81.8)		
III	43 (86.0)	22 (84.6)	21 (87.5)		7 (16.3)	3 (14.3)	4 (18.2)		
Tumor location				0.686				0.760	0.307
Right sided	24 (48.0)	11 (42.3)	13 (54.2)		18 (41.9)	8 (38.1)	10 (45.5)		
Left sided	24 (48.0)	14 (53.9)	10 (41.7)		25 (58.1)	13 (61.9)	12 (54.5)		
Rectum	2 (4.0)	1 (3.8)	1 (4.1)		0 (0.0)	0 (0.0)	0 (0.0)		
Tumor diameter				0.090				0.876	0.147
<5 cm	33 (66.0)	20 (76.9)	13 (54.2)		22 (51.2)	11 (52.4)	11 (50.0)		
≥5 cm	17 (34.0)	6 (23.1)	11 (45.8)		21 (48.8)	10 (47.6)	11 (50.0)		
Deceased				0.000				0.000	0.452
No	24 (48.0)	21 (80.8)	3 (12.5)		24 (55.8)	19 (90.5)	5 (22.7)		
Yes	26 (52.0)	5 (19.2)	21 (87.5)		19 (44.2)	2 (9.5)	17 (77.3)		
Adjuvant therapy				0.749				0.072	0.001
No	22 (44.0)	12 (46.2)	10 (41.7)		34 (79.1)	19 (90.5)	15 (68.2)		
Yes	28 (56.0)	14 (53.8)	14 (58.3)		9 (20.9)	2 (9.5)	7 (31.8)		
FU years									
Range	0.14–10.86				0.31–12.47				
Mean	4.33				6.96				

NOTE: Shown are data from associations of the discovery and validation patient cohort with clinicopathologic parameters. Cohorts are divided into patients with and without metastasis in the follow-up. Significant associations are indicated in bold.

Abbreviations: N, number of patients; *P*, *P* value; M–, no metastasis; M+, metachronous metastasis.

For validation, patients were selected from a colon cancer patient population ($N = 242$), in part described previously (27), who underwent surgery between 1987 and 2006 at the Department of Surgery, Klinikum rechts der Isar, Technische Universität München (TUM, Munich, Germany). All patients gave written, informed consent, with approval of the Ethics committee of the Medical Faculty of TUM (#1926/07). Included tumors were histologically proven TNM stage II–III first primary colon carcinomas with microsatellite stable phenotype, with available snap-frozen primary tumor material. We selected the patients with metachronous distant metastases ($N = 21$) and matched these with patients without metastases. Matching criteria were similar as described above. The validation study population consisted of 43 patients (Table 1). All patients received the recommended postoperative follow-up care, according to the national German guidelines (S3 guidelines by the German Cancer Society DGK/AWMF).

All samples were coded, according to National ethical guidelines ("Code for Proper Secondary Use of Human Tissue," Dutch Federation of Medical Scientific Societies). This study was performed according to the REMARK guidelines (NCI-EORTC; ref. 28).

LCM and preparation of RNA

Snap-frozen tumor tissues were collected for each patient. LCM was performed to select cancerous epithelial cells. Hematoxylin staining, LCM, and subsequent sample lysis was completed within 1 hour. Sections (8 μm thick) mounted on LCM-suitable slides (Zeiss) were processed on the PALM MicroBeam (Zeiss). A minimum of 3-mm² cancer cell area of the sections was collected. RNA was extracted using the Just-a-Tube sample total RNA/microRNA purification kit (Charm Biotech) for the discovery cohort or TRIzol Reagent (Life Technologies) for the validation cohort, according to manufacturers' recommendations. Total RNA quantity and miRNA quality were assessed with the NanoDrop 1000 (Thermo Scientific) and RNA 6000 Pico total RNA-Chip (Agilent Technologies), respectively.

Preparation for NGS

Tumor tissue libraries for NGS were prepared as described by Buermans and colleagues (29). One universal forward primer and 14 reverse primers, with different barcodes, were available for the amplification PCR. A unique barcode was assigned to each sample within a sample pool. The presence of cDNA was assessed with a high-sensitivity DNA Chip (HS-DNA Chip; Agilent Technologies). In addition, a set of qRT-PCRs served to show whether miRNAs were present. Samples with at least 90% of desirable sequence lengths, with calculated insert sizes of 17–25 nucleotides, were randomly pooled into 4 pools in equimolar volumes of 2 nmol/L and subsequently reassessed with HS-DNA Chip.

NGS with HiSeq2000 and Ion Torrent and data extraction

High-depth NGS was performed on the four miRNA pools, loaded into 4 different lanes of a Illumina HiSeq2000 v3-flowcell with v3-reagents (Illumina). The sequence run type was a single-end read of 50 cycles with an added sample-index read for pooled-sample identification. In addition, low-depth NGS was performed using the Ion Torrent platform with an Ion314 chip (Life Technologies) to sequence a pooled subset of 13 of the 50 samples, and used for technical validation. Data from both NGS runs were

extracted and processed into raw count data per miRNA as described previously (29, 30), with minor adaptations. In summary, the 3'-library adapter was removed from the sequences, and remaining reads of at least 15 bases were aligned to the human Hg19 reference genome using the Bowtie short-read aligner allowing for up to three loci with two mismatches. Crossmapping correction was applied to the alignments as described by de Hoon and colleagues (31). Aligned sequence tags were annotated on the basis of the miRBase v19 release resulting in separate expression values for the 5' and 3' mature miRNAs from the hairpin precursors. Data are publicly available (GEO number: GSE63119).

Data analysis NGS

Data analysis was performed in R version 3.0.1. Library sizes were defined as the total number of annotated miRNA reads per sample. Normalization started with transforming raw miRNA count data to counts per million (cpm) using the library size consisting of only annotated miRNA isoforms (miRBase v19). Subsequently, quantile normalization (R-package limma) was performed, according to the recommendations of Garmire and colleagues (32, 33). All libraries passed quality control (R-package arrayQualityMetrics) before and after normalization. Low-expression miRNAs, which are those without expression of at least 1 cpm in at least 10 of 50 samples, were excluded from analysis. Subsequently library sizes were reset to the remaining number of miRNA reads.

Differential expression analysis was performed using the nonparametric Wilcoxon rank sum test (R-package stats) to restrict the influence of expression outliers on the results, and the number of false positive miRNAs. The number of normalized reads is representative for the expression of a miRNA in that sequence library. Multiple testing correction of P values was performed with false discovery rate (fdr).

Technical validation of the HiSeq NGS data with the Ion Torrent NGS data was performed by determining three different correlation coefficients. The intraclass correlation coefficient (ICC; absolute agreement) as a measure of intra- and intersamples variability and the Pearson product-moment (Pearson r) correlation coefficient as a measure of the linear correlation between the two datasets per sample were determined on cpm data of all detected miRNAs. In addition, the Spearman rank correlation coefficient (Spearman ρ) on rank data was calculated as a measure of statistical dependence between the two datasets per sample. Included were ranked data of miRNAs with more than 2 raw counts in both NGS datasets. A significant P value ($P < 0.05$) indicated correlation for the miRNA expression profiles per patient for the two datasets.

Identification of reference miRNAs for technical validation and validation of results by qPCR were determined using the coefficient of variation (CV) per miRNA, with a low CV indicating the miRNAs with the best linear relation between the number of isoform-specific miRNA reads and the total number of reads across the samples. To calculate the CV per miRNA, the SD was divided by the average number of reads.

Stem-loop real-time PCR and data analysis

Eight different genes from the top ten list of the differential expression analysis and two reference miRNAs were validated by qRT-PCR. miRNA-specific cDNA was prepared with the TaqMan miRNA Reverse Transcription Kit (Applied Biosystems) combined with miRNA assays (Applied Biosystems) in

96-well plates. The real-time PCR of the different TaqMan miRNA assays (Applied Biosystems) were run on the CFX-384 machine (Bio-Rad), using the TaqMan Universal Master Mix II (Applied Biosystems). Thresholds for calculation of the cycle threshold (C_t) were set separately for each miRNA. Data are available upon request. Triplicate measurements were averaged for further analyses. The average of miR16 C_t value and miR26a C_t value per sample served as reference for relative expression calculations (formula $2^{\Delta C_t}$). The Pearson's r , as a measure of the linear correlation between the two datasets per sample, was determined per miRNA on NGS and qRT-PCR data of all patients. In addition, the Spearman's ρ on rank data, as a measure of statistical dependence between the two datasets, were calculated per miRNA on NGS and qRT-PCR data of all patients.

Survival analyses

Relative expression of different miRNAs, as determined by qRT-PCR, was used as predictor in survival analyses with the statistical package SPSS 20.0 for Windows (SPSS Inc). Patients were dichotomized per miRNA based on the average of mean and median as determined in the discovery cohort. Receiver operating characteristic (ROC) curves for the individual miRNAs were used to assess the specificity and sensitivity of the chosen cut-off values. The relationship between dichotomized relative miRNA expression and known clinicopathologic parameters was investigated using the Pearson χ^2 test. The Cox proportional hazards model was used to analyze the association between relative miRNA expression with distant recurrence-free period (DRFP), defined as the time from surgery until the diagnosis of the first distant recurrence. Differences in DRFP between patient groups are presented as HRs. Cumulative risk curves were calculated, accounting for death by all causes as competing risk for metastasis, and served to visualize the associations (34). The most important covariates for colon cancer, being age at surgery, TNM stage, and gender, were included in multivariate analyses irrespective of statistical significance. The high expression group served as reference for all survival analyses. All tests were two-tailed and P values <0.05 were considered statistically significant. A nomogram for prediction of distant metastasis in colon cancer was constructed (R-package rms) using the model obtained from the multivariate survival analysis of the discovery cohort. Calibration plots (R-package rms) and ROC curves were used to assess the prediction performance and accuracy of the nomogram. For comparison, we performed the survival analyses for the validation cohort after building the survival model by using the dichotomization cutoff as determined on the discovery cohort.

Gene target prediction

miRNA gene target prediction was performed by combining results from three different target prediction databases and a database containing validated targets. Results from individual predicted miRNA target searches of the TargetScan, DIANA-microT-CDS, and the miRDB were combined with the results of the validated miRNA target database miRTarBase. Putative target genes were assessed for their associated Gene Ontology (GO) terms using the online functional annotation tool DAVID. In addition, a NCBI gene and PubMed search was performed for target gene associations with cancer (metastasis) using the HGNC symbol.

Results

NGS and annotation of miRNAs in colon cancer

High-depth multiplex-NGS of miRNAs from tumor cells of 50 patients in the discovery cohort, randomly divided into four sample pools, yielded variable raw library sizes with an average of >7 million reads. The total number of reads aligned to miRNA sequences is denoted the miRNA library size of a particular tumor specimen. Spreading of the amount of sequence reads (e.g., library sizes) across samples is shown in Supplementary File S1A. Analysis revealed expression of 2,707 from 7,160 annotated miRNA isoforms (miRBase v19) in at least 1 of 50 tumor specimens. After removal of low-expression miRNAs, 667 miRNAs remained eligible for statistical analyses. miR21-5p was most abundant in our dataset, representing an average of almost 10% of the miRNA sequence reads per sample. The spread of percentages of the total miRNA library sizes represented by the three most abundantly expressed miRNAs is shown as box plots in Supplementary File S1B, per miRNA.

Technical validation of miRNA expression by NGS

Low-depth NGS with the Ion Torrent was used to sequence 13 of the 50 tumor specimens for technical validation. Sequence depths ranged from 50 to 1,500 reads per sample. Analysis revealed expression of 325 annotated miRNA isoforms in at least one of the 13 tumor samples. All these miRNAs were also detected in the high-depth HiSeq NGS analysis. Correlation analysis of paired samples in the HiSeq and Ion Torrent datasets confirmed reproducibility of miRNA expression across the different NGS platforms, even with considerable differences in sequence depth between the two platforms (Supplementary File S2A).

Differential expression of miRNAs associated with distant metastasis

The expression of the 667 miRNAs detected by HiSeq NGS were assessed for differences between metastasized and nonmetastasized primary tumors. Differential expression analysis identified 32 miRNAs with nonadjusted P values of 0.02 or lower (Table 2, Supplementary File S3). From these miRNA isoforms, 19 (59%) showed lower expression and 13 (41%) showed higher expression in metastasized tumors compared with nonmetastasized tumors. This top list indicated that both the 3p and 5p variant of miR378 were differentially expressed. In addition, for miR1, miR135a, miR365, and miR550a two different transcripts were differentially expressed. Eight of the top ten listed miRNAs were selected for technical validation with PCR.

Identification of reference miRNAs for validation purposes

miRNA expression patterns were assessed to identify reference miRNAs for (technical) validation using qRT-PCR. Those miRNAs with the best linear relation between the number of isoform-specific miRNA reads (or miRNA expression) and the total number of miRNA reads (or miRNA library size) across the samples were identified by calculating the CV per miRNA. A low CV indicates low SD combined with high average expression. Calculated CVs and data of subsequent linear modeling of the two miRNAs with lowest CV, being miR16-1-5p and miR26a1-5p, are shown in Supplementary File S4A. The linear relation of miRNA expression and miRNA library size, and the residuals plots were plotted for graphical inspection (Supplementary File S4B). By selecting two miRNAs as combined reference, we

Table 2. List of most differentially expressed miRNAs

	miRNA	chr	P	FC	Validation
1	miR378a-3p	5	0.00032 ^a	-1.95	002243
2	miR135a-1-5p	3	0.0011	2.67	000460
3	miR135a-2-5p	12	0.0011	2.67	000460
4	miR29b-2-5p	1	0.0013	2.00	002166
5	miR589-5p	7	0.0015	-7.09	002409
6	miR140-5p	16	0.0036	1.72	001187
7	miR25-3p	7	0.0036	-1.70	000403
8	miR339-5p	7	0.0038	-1.62	002257
9	miR378a-5p	5	0.0050	-1.61	
10	miR550a-1-5p	7	0.0060	-1.38	002410
11	miR1-1-3p	20	0.0060	1.98	
12	miR550a-2-5p	7	0.0064	-1.36	
13	miR31-5p	9	0.0068	2.21	
14	miR1-2-3p	18	0.0068	1.92	
15	miR1226-3p	3	0.0069	3.21	
16	miR1977-3p	1	0.0076	3.19	
17	miR188-5p	X	0.0077	-1.62	
18	miR200c-3p	12	0.0082	-1.60	
19	miR33b-3p	17	0.0085	-4.36	
20	miR365b-3p	17	0.011	1.28	
21	miR365a-3p	16	0.012	1.27	
22	AL133472.1-5p	6	0.012	-2.18	
23	miR363-3p	X	0.013	8.67	
24	miR592-5p	7	0.015	-1.94	
25	miR299-5p	14	0.015	-2.01	
26	miR205-5p	1	0.016	9.60	
27	miR7-2-5p	15	0.016	-1.71	
28	miR194-2-3p	11	0.016	-1.66	
29	miR425-5p	3	0.016	-1.42	
30	miR106b-3p	7	0.018	-1.46	
31	miR147b-3p	15	0.019	-4.49	
32	miR552-3p	1	0.019	-1.71	
ref	miR16-5p	13	NA	NA	000391
ref	miR26a-5p	3	NA	NA	000405

NOTE: Shown is the top list of differentially expressed miRNAs, with *P* values below 0.02, when comparing normalized miRNA expression in patients with nonmetastatic and metastatic tumors. The chromosome (chr) on which the miRNAs is located is listed. *P* values were computed using the Wilcoxon rank sum test. The fold change (FC) is computed for the nonmetastasized tumors versus the metastasized tumors. The most right column indicates which miRNA isoforms were validated. Given is the assay number of the qRT-PCR (Applied Biosystems).

Abbreviations: ref, reference miRNA for qRT-PCR; NA, not applicable.

^aSignificant after *fd*.

aimed to produce a more robust reference, as compared with a single miRNA. For normalization of the amount of input material in the validation phase, we identified the combination of miR16-1-5p and miR26a1-5p as the most suitable reference. Validation of this combination of miRNAs was performed using online available miRNA sequencing data from the cancer genome atlas (TCGA) data portal (Supplementary File S4C). Results endorse our findings. In addition, normalization of our HiSeq NGS data with the combination of these two miRNAs resulted in similar results in differential expression analysis compared with the selected normalization of NGS data as described in the Materials in Methods section (data not shown). These results indicate the capability/suitability of this combination of miRNAs for normalization purposes of quantitative PCR of colon tissue-derived miRNA expression.

Assessment of miRNA expression by stem-loop real-time PCR

For technical validation in the discovery cohort and validation of results in the validation cohort, qRT-PCR was performed

for 8 top ten listed miRNAs identified in the differential expression analysis of HiSeq NGS data (Table 2). Expression was calculated relative to the average expression of miR16-1-5p and miR26a1-5p. Tested miRNAs were technically validated with Pearson's *r* and Spearman's ρ correlation analysis (Supplementary File S2B). Selection of variables that can provide robust prediction is an important aspect for survival model building. miRNAs varying strongly between cohorts can by definition not serve as such robust predictors. Therefore, selection of miRNAs for survival model building was based on miRNAs showing both similar expression patterns between NGS and qRT-PCR (discovery cohort), and similar relative expression ranges in the discovery and validation cohort as assessed by qRT-PCR. Four of the assessed miRNAs, namely miR25-3p, miR339-5p, miR378-3p, and miR550a-1-5p, showed similar expression patterns between the NGS and qRT-PCR data in the discovery cohort. For miR25-3p and miR339-5p relative expression ranges were similar between the tumors of the discovery and the validation cohort. Differential expression results, as identified by HiSeq NGS, were supported by qRT-PCR results (Fig. 1). In both patient cohorts, a lower expression of either miRNA was associated with the occurrence of a metachronous metastasis.

Prognostic value of miRNAs for distant recurrence-free period

Survival model building was performed on the qRT-PCR data of the discovery cohort using univariate and multivariate Cox regression analyses on miR25-3p and miR339-5p together with predetermined covariates. Afterwards, for comparison, we performed the survival analyses in the validation cohort. Except for survival status and distant metastasis status, neither individual expression of miR25-3p nor miR339-5p expression, or their combination was associated with any other clinicopathologic parameter (Supplementary File S5). Survival analyses were performed to assess differences in DRFP between low and high relative expression groups for the two individual miRNAs. Cut-off values for group classification were determined in the discovery cohort at 0.279 and 0.0143 for miR25-3p and miR339-5p, respectively. ROC curves for the individual miRNAs were used to assess the specificity and sensitivity of the chosen cut-off values in the discovery set, and confirmed them to be optimal (Supplementary File S6). Univariate analyses showed an association between high relative miRNA expression and shorter DRFP for either miRNA in both the discovery and the validation cohort (Table 3). Histograms showing the relation between height of miRNA expression and number of patients with a distant metastasis at the time point 5 years from surgery (Fig. 2A), and cumulative incidence curves for DRFP (Fig. 2B) were plotted. Multivariate analyses identified both miR25-3p and miR339-5p as independent prognostic factor for DRFP (Table 3). HRs were consistently more pronounced in the validation cohort compared with the discovery cohort.

Combination of biomarkers may better reflect tumor aggressiveness and might increase the clinical discriminative and prognostic value. Therefore, expression of miR25-3p and miR339-5p were combined into two groups, being high expression of both miRNAs and low expression of one or both miRNAs. Cumulative incidence curves illustrating univariate survival analyses for DRFP were plotted (Fig. 2C). The 5-year survival percentage for DRFP in the discovery cohort and the validation cohort was, respectively, 70% and 93% for patients

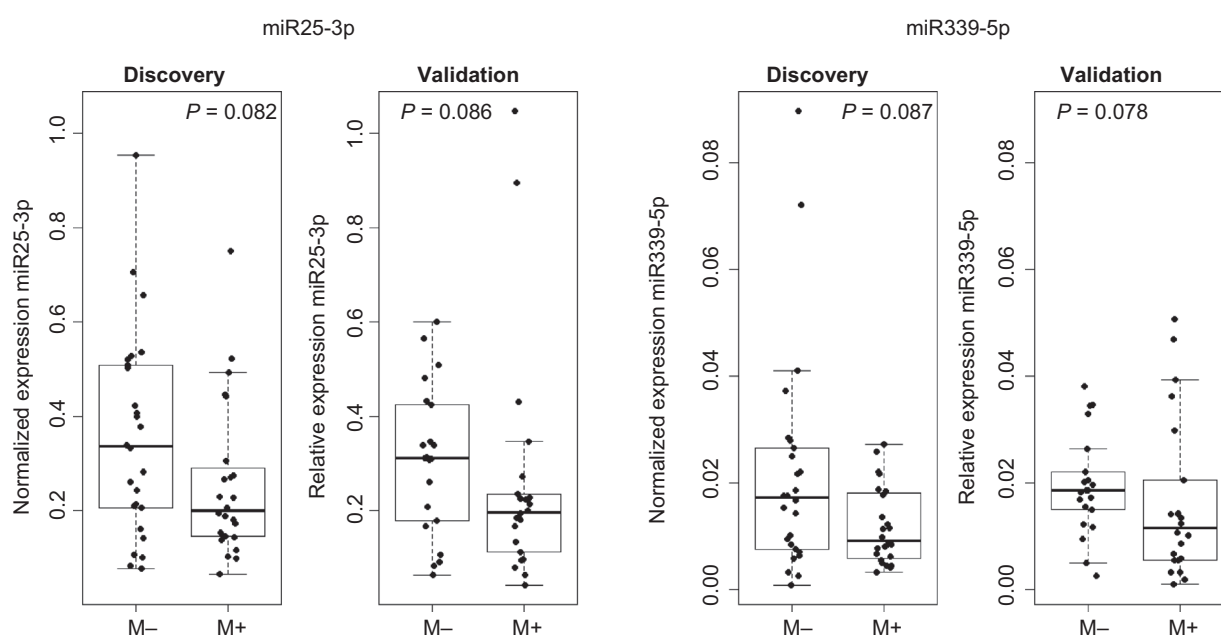


Figure 1.

Validation miRNA expression profiles. Shown are miRNA expression profiles per miRNA of the discovery cohort ($N = 50$) and the validation cohort ($N = 43$), for miR25-3p and miR339-5p, as assessed by qRT-PCR. Patients are divided on the basis of the occurrence of a metachronous metastasis. Expression is given relative to the average expression of miR16-5p and miR26a-5p, who together serve as normalization for the amount of input material. For indicative purposes, the P values using the Wilcoxon rank sum test were specified. Abbreviations: M-, no metastasis in the follow-up; M+, metastasis in the follow-up; P , P value.

with a high miRNA expression for both miRNAs compared with, respectively, 39% and 31% for patients with low expression for one or two of the miRNAs. Multivariate analyses identified combined expression levels of miR25-3p and miR339-5p as independent prognostic factor for DRFP (Table 3). Patients with a high expression of both miRNAs were shown to have an 11 times lower change of developing a distant metastasis in the validation cohort based on multivariate survival analyses.

Nomogram for prediction of distant recurrence risk

Using the multivariate survival model combining miR25-3p and miR339-5p for DRFP estimated on the discovery cohort (Table 3), we constructed a nomogram to quantitatively predict the probability of metachronous distant metastasis in new cancer patients (Fig. 3A), at 3 years and 5 years after surgery. The nomogram was based on primary tumors of TNM stage II–III colorectal cancer patients with a microsatellite stable phenotype. The miRNA-based classifier and the predetermined most

Table 3. Results of survival analyses for DRFP of individual and combined miRNA expression

	miR25-3p		miR339-5p		miRNA combination	
	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P
Discovery cohort						
Univariate						
miRNA	2.43 (0.96–6.13)	0.06	2.41 (1.00–5.84)	0.05	3.00 (1.02–8.80)	0.05
Multivariate						
miRNA	2.37 (0.93–6.03)	0.07	2.29 (0.90–5.82)	0.08	2.92 (0.98–8.64)	0.05
Age at surgery	1.12 (0.77–1.62)	0.56	1.09 (0.76–1.56)	0.63	1.11 (0.77–1.61)	0.57
TNM stage	1.42 (0.42–4.85)	0.58	1.06 (0.30–3.73)	0.93	1.39 (0.41–4.73)	0.60
Gender	0.78 (0.34–1.79)	0.57	0.86 (0.37–1.98)	0.72	0.81 (0.35–1.84)	0.61
Validation cohort						
Univariate						
miRNA	4.21 (1.42–12.53)	0.01	3.92 (1.51–10.17)	0.005	8.92 (2.07–38.43)	0.003
Multivariate						
miRNA	5.76 (1.82–18.20)	0.003	4.11 (1.57–10.75)	0.004	11.01 (2.50–48.51)	0.002
Age at surgery	1.67 (0.83–3.37)	0.15	1.30 (0.65–2.61)	0.45	1.55 (0.79–3.02)	0.20
TNM stage	1.88 (0.56–6.28)	0.31	1.56 (0.48–5.11)	0.46	1.96 (0.57–6.72)	0.28
Gender	0.67 (0.24–1.88)	0.45	0.84 (0.32–2.21)	0.72	0.79 (0.29–2.15)	0.64

NOTE: Shown are the data from univariate and multivariate survival analyses of individual and combined miRNA expression. Data are shown for the patients in the discovery ($N = 50$) and validation cohort ($N = 43$) separately. HRs are given for high expression versus low expression or high expression of both miRNAs versus low expression of one or both miRNAs, for individual and combined miRNA analyses, respectively. Significant associations are indicated in bold. Abbreviations: N, numbers at risk; CI, confidence interval; P , P value; DRFP, distant recurrence-free period.

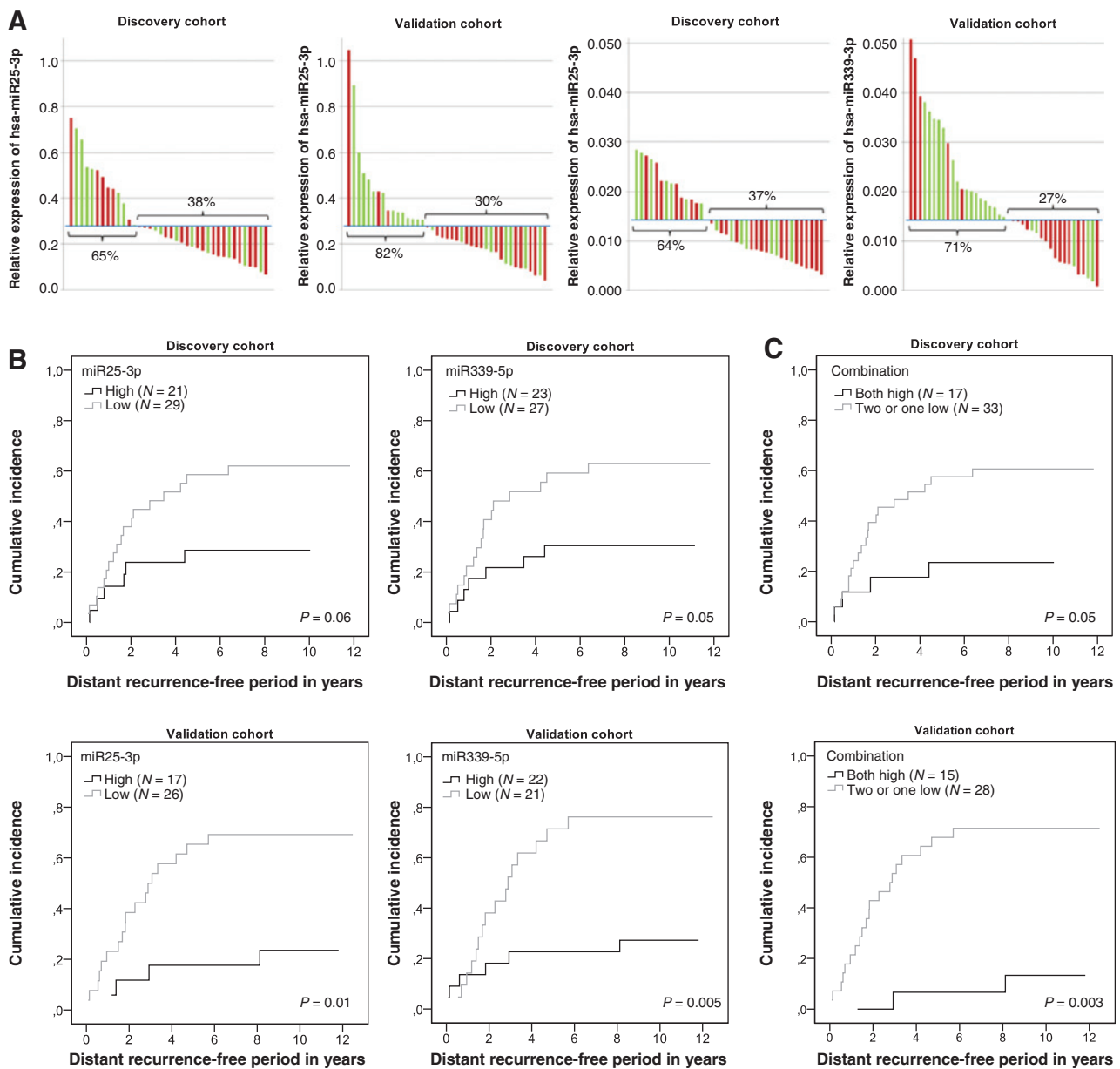
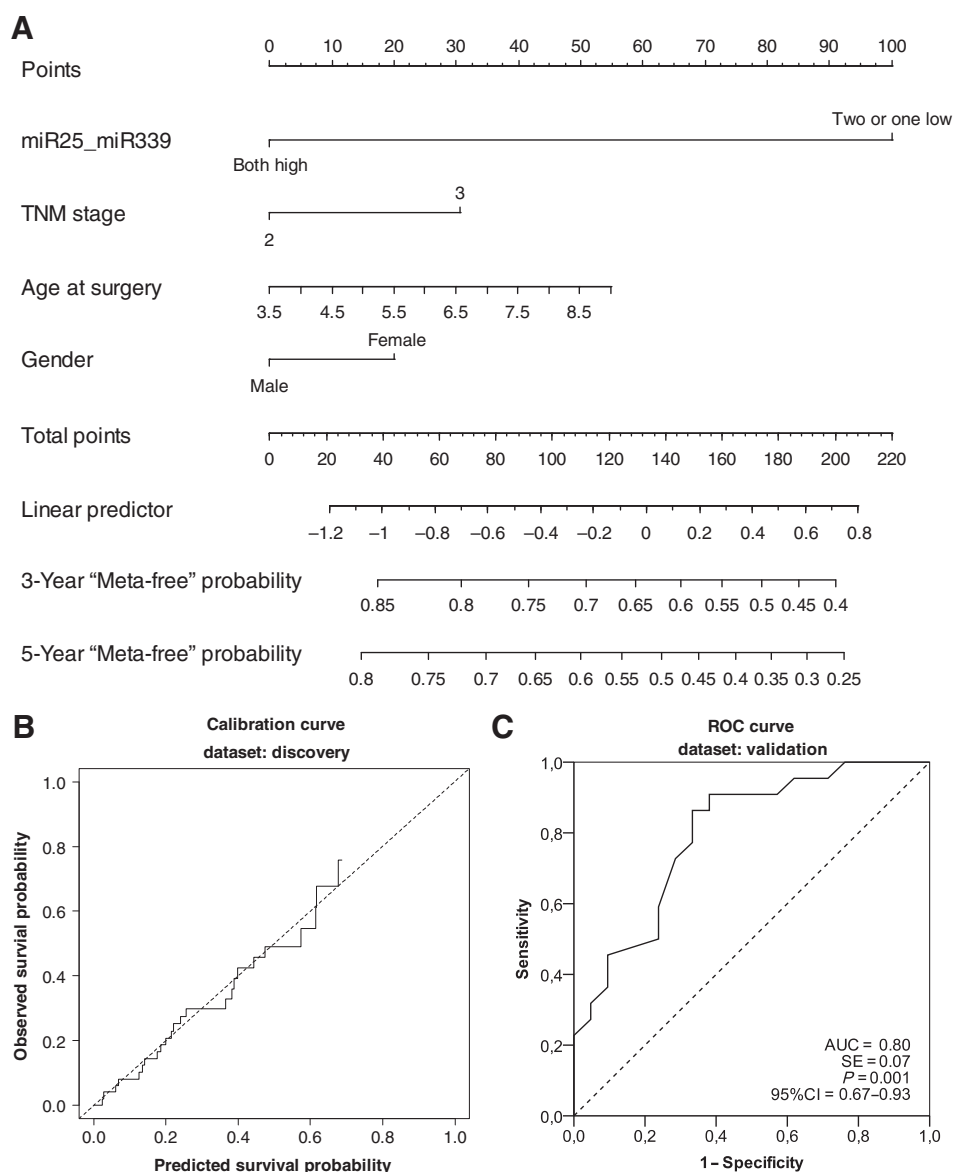


Figure 2. Graphical presentations of distant metastasis incidence and risk in patient groups. Data are shown for the discovery cohort ($N = 50$) and the validation cohort ($N = 43$) separately. A, presented are histograms showing individual miRNA expression in relation to the metastasis status of the different patients, at 5 years after surgery. The bar colors indicate if patients presented with (red) or without (green) a metastasis at the 5-year time point. Patients that were censored in the first 5 years after surgery were not shown. The cut-off values for division of patients into high and low expression groups for miR25-3p (0.279) and miR339-5p (0.0143) are presented as blue horizontal lines. The percentages of patients without a metastasis are indicated for high and low expression groups. B, the competing risk for the development of metachronous metastasis is death by any cause. Cumulative incidence curves that account for competing risks illustrate univariate survival analyses of individual miRNA expression. P values of the univariate Cox proportional hazard analyses are presented in the graphs. Patients were divided into two groups: high expression (■) and low expression (▬). C, cumulative incidence curves that account for competing risks illustrating univariate survival analyses of combined miRNA expression of miR25-3p and miR339-5p. P values of the univariate Cox proportional hazard analyses are presented in the graphs. Patients were divided into two groups: high expression of both miR25-3p and miR339-5p (■) and low expression of one or both miRNAs (▬).

important nonsubjective clinicopathologic risk factors were integrated in the nomogram. The calibration plot showed that the prediction performance of the nomogram was good compared with the ideal model (Fig. 3B). The predictive accuracy of the nomogram was assessed in an independent

patient cohort with a ROC analysis (Fig. 3C). The concordance index (Fig. 3C) for the multivariate survival model (Table 3), indicated good probability prediction for the prognostic multivariate survival model in the independent validation cohort.

**Figure 3.**

Nomogram for prediction of metastasis probability in colon cancer. A, shown is the nomogram for the prediction of metastasis in colon carcinoma based on the multivariate survival model of the discovery cohort. Age at surgery is in units of 10 years. The total points were converted to the Linear Predictor and then to predicted probabilities. 3-Year "metastasis-free" probability and 5-Year "metastasis-free" probability denotes the probability of not having a metastasis at 3 or 5 years after surgery. B, the calibration plot gives the prediction performance of the proposed nomogram in the discovery cohort, with the closer to the 45° line the better the performance. C, plotted is a ROC curve for the independent validation cohort of the survival model used in the nomogram. As a measure of the accuracy of the probability prediction of the prognostic model, the AUC is denoted, with the closer to 1 the better the model.

miRNA gene targets

miRNAs have numerous target genes and exert tissue-specific transcriptional regulation. To assess possible target genes for miR25-3p and miR339-5p, gene target prediction was performed (Supplementary File S7A) by combining results from three well-established prediction tools (TargetScan, DIANA-microT-CDS, and miRDB) and a database containing validated targets (miRTarBase). For miR25-3p and miR339-5p we found, respectively, 38 and 36 overlapping target genes predicted in all three target prediction databases, indicating poor agreement between the target prediction databases. GO term enrichment was performed for these overlapping targets (Supplementary File S7B). In addition, only 1 or 2 of the overlapping predicted target genes for miR25-3p or miR339-5p was validated according to the miRTarBase. The number of common target genes for miR25-3p and miR339-5p, predicted by the same database(s), was 66 (Supplementary File S7C). We assessed the associated GO terms of the

putative target genes using online annotation tool DAVID, and performed a NCBI gene and PubMed search using the HGNC symbols (Supplementary File S7). Annotation analysis of putative target genes suggests involvement of these genes in cellular metabolic processes, indicating a possible role for the identified miRNAs in cellular metabolism. In conclusion, functional studies are needed to unravel the actual target genes of miR25-3p and miR339-5p in colon cancer cells.

Discussion

In this study, we identified the combination of miR25-3p and miR339-5p expression levels in tumor cells as independent prognostic factor for the occurrence of distant metastasis in TNM stage II-III colon cancer with a microsatellite stable phenotype. We proposed, and retrospectively validated, in an independent patient cohort, a nomogram for prediction of distant

metastasis for this group of patients. In addition, we recommend combination of miR16-5p and miR26a-5p as standard for normalization in qPCR of colon cancer tissue-derived miRNA expression.

Multiple different processes are involved in carcinogenesis, with many key genes/proteins. The hallmarks of cancer (35) represent properties that a cell needs to attain to become and sustain itself as cancer cell. These hallmarks direct the cellular pathways involved in carcinogenesis, and particularly metastasis formation. Using expression of miRNAs for prognostication of clinical outcome for cancer has advantage over mRNAs, as miRNAs are considered the key posttranscriptional regulators of gene expression, with numerous simultaneous target genes (15) and involvement in various cellular pathways. In contrast to mRNA, these posttranscriptional master regulators, are highly conserved between species (2).

Several miRNA profiles for prediction of metastasis risk in patients with colon cancer have been proposed before (20–24). Recently, a six miRNA-based classifier, differing from the combination of miRNAs reported here, was published (24). Zhang and colleagues identified a six miRNA-based classifier as independent prognostic factor for DRFP in TNM stage II colon cancer patients. However, the discriminative value of these miRNAs could not be confirmed with our NGS data, as individual miRNAs were neither differentially expressed between patients with and without metachronous metastasis nor prognostic for distant metastasis risk. Differences in identified miRNAs might be due to the substantial dissimilarities of the respective study designs and patient cohorts. Given the large heterogeneity of colon cancers, with regard to the amount of stroma and immune infiltrate, it is essential to use microdissected cancer cells and high-depth sequencing of the microneome. A combination of these approaches with the specific assessment of differential expression between patients presenting with or without metachronous metastasis allows prediction of an accurate prognostic (multivariate) model for metastasis risk. Moreover, it is crucial to subsequently validate the model in an independent patient cohort. Exclusion of patients with low percentages of cancerous epithelial cells leads to biased patient cohorts, as high stroma content of tumors is itself associated with worse clinical outcome (36). Inclusion of these patients requires selection of target cells by microdissection (18), due to the tissue-type and cell-specific nature of miRNA expression (16, 17). To obtain the most uniform cell population, we used LCM, which is a laborious technique most likely not applicable for routine clinical use. Alternatively, manual microdissection can be used, which is already standard procedure for assessment of PCR-based biomarkers in many diagnostics laboratories. In addition, introduction of automated microdissection using robotics is upcoming. To the best of our knowledge, this is the first study to combine LCM with high-depth sequencing of miRNAs, aiming to address differential expression between patients presenting with and without distant metachronous metastasis for accurate prognostication.

A relatively small number of miRNAs account for the majority of raw sequence reads in the miRNA libraries, as seen before (22). We hypothesized that these miRNAs, being miR21-5p, miR192-5p, and miR19b-2-3p, are important in colon cells, and probably in colon carcinogenesis, but not necessary in the process of metastasis. miRNA-21-5p was previously reported to show increased expression in tumor compared with normal colon mucosa (37), and was linked to early onset carcinogenesis (38). Two independent manuscripts reported that increased

expression of miR21, predominantly in the stroma compartment and not the cancer cells, is associated with clinical outcome in TNM stage II colon cancer patients (39, 40). These findings might explain why we did not see an association between miR21 expression and metastasis risk, unlike others (24). Decreased expression of miR192 was identified in tumor compared with normal colon tissue (41) and linked to early onset colon carcinogenesis (38). Upregulation of miR19b was associated with colon cancer (42) and with 5-fluorouracil resistance in colon cancer cells (43).

On the basis of the linear relation between miRNA expression and miRNA library size, we identified combination of miR16-5p and miR26a-5p as most suitable reference for qPCR. Both miRNAs have been recognized before as suitable reference miRNAs (44) in a TaqMan microarray-based approach. While our analysis was limited to TNM stage II–III colon cancer patients, Chang and colleagues validated the miRNAs in a colorectal cancer cohort of TNM stage I–IV patients. These according results confirm the utility of these miRNAs as reference for qPCR in cancer tissues of colorectal origin.

As discussed previously (45), function of miR25 can be oncogenic or tumor repressive depending on the tissue type. Generally, expression of miR25 is higher in cancer cells of epithelial origin, and in tumor stroma, compared with normal colon mucosa (45, 46). We showed that expression of miR25-3p is lower in patients with colon cancer with distant metachronous metastases compared with those without distant metastases. In prostate cancer cell lines, a similar expression pattern was observed (47), with increase in expression of carcinoma cell lines compared with immortalized normal cells, and subsequently lower expression in cell lines from prostate carcinoma metastasized to distant locations compared with those of primary tumors (47). Taken together, these data suggest that in colon cancer, expression of miR25-3p, after initial increase, might need to be decreased to some extent to enable the initiation of the distant metastasis process. Indeed, Li and colleagues showed that overexpression of miR25 suppressed migratory ability of colon cancer cells *in vitro* and decreased tumorigenic potential in a colon cancer xenografts (45). In addition, Baffa and colleagues reported decreased expression of miR25 in metastases compared with their associated primary colon tumor (48). Expression of miR339-5p has been linked to the process of distant metastases in breast cancer (49). Similar to the association we observed for miR339-5p, higher expression was associated with lower metastasis rates, and in addition reduced cancer cell growth and better patient survival in breast cancer (49). Also for prostate cancer, lower expression of miR339-5p was observed in metastatic cancer (50).

To date, treatment allocation of patients with colon cancer is based on the TNM classification system. Under current classification, patient groups still show large differences in clinical outcome (9–11). Moreover, clinically defined risk criteria for adjuvant therapy in locally restricted colon cancer are currently considered to be of insufficient quality (9, 12). The development of distant recurrent disease mainly accounts for colon cancer-related mortality. The use of the miRNA combination reported here, capable of predicting metastasis risk in colon cancer, in conjunction with the TNM staging method, may contribute to an approach of personalized treatment based on individual tumor characteristics. In conclusion, in this bicentric international study, we retrospectively identified and validated a miRNA classifier in the primary tumor using two independent patient sets, and

propose a nomogram capable of predicting metastasis risk in microsatellite stable TNM stage II–III colon cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: I.J. Goossens-Beumer, C.J.H. van de Velde, P.J.K. Kuppen

Development of methodology: I.J. Goossens-Beumer, R.S. Derr, H.P.J. Buermans, P.J.K. Kuppen

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): I.J. Goossens-Beumer, R.S. Derr, H.P.J. Buermans, H. Morreau, U. Nitsche, K.-P. Janssen, C.J.H. van de Velde, P.J.K. Kuppen

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): I.J. Goossens-Beumer, H.P.J. Buermans, J. Goeman, S. Böhringer, U. Nitsche

Writing, review, and/or revision of the manuscript: I.J. Goossens-Beumer, R.S. Derr, H.P.J. Buermans, S. Böhringer, H. Morreau, U. Nitsche, K.-P. Janssen, C.J.H. van de Velde, P.J.K. Kuppen

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): I.J. Goossens-Beumer, H. Morreau, P.J.K. Kuppen

Study supervision: C.J.H. van de Velde, P.J.K. Kuppen

Acknowledgments

The authors thank Sabine Leis and Anne Benard for their help in the collection and processing of frozen tissue samples of the validation cohort.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 4, 2014; revised October 9, 2014; accepted October 9, 2014; published OnlineFirst October 14, 2014.

References

- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004;116:281–97.
- Moss EG. MicroRNAs: hidden in the genome. *Curr Biol* 2002;12:R138–R140.
- Sassen S, Miska EA, Caldas C. MicroRNA: implications for cancer. *Virchows Arch* 2008;452:1–10.
- Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009;19:92–105.
- Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A* 2004;101:2999–3004.
- McManus MT. MicroRNAs and cancer. *Semin Cancer Biol* 2003;13:253–8.
- Croce CM, Calin GA. miRNAs, cancer, and stem cell division. *Cell* 2005;122:6–7.
- Ma L, Weinberg RA. MicroRNAs in malignant progression. *Cell Cycle* 2008;7:570–2.
- Benson AB III, Schrag D, Somerfield MR, Cohen AM, Figueredo AT, Flynn PJ, et al. American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. *J Clin Oncol* 2004;22:3408–19.
- Nitsche U, Maak M, Schuster T, Kunzli B, Langer R, Slotta-Huspenina J, et al. Prediction of prognosis is not improved by the seventh and latest edition of the TNM classification for colorectal cancer in a single-center collective. *Ann Surg* 2011;254:793–800.
- Quirke P, Cuvelier C, Ensari A, Glimelius B, Laurberg S, Ortiz H, et al. Evidence-based medicine: the time has come to set standards for staging. *J Pathol* 2010;221:357–60.
- van Cutsem E, Oliveira J. Primary colon cancer: ESMO clinical recommendations for diagnosis, adjuvant treatment and follow-up. *Ann Oncol* 2009;20 Suppl 4:49–50.
- Nicoloso MS, Spizzo R, Shimizu M, Rossi S, Calin GA. MicroRNAs—the micro steering wheel of tumour metastases. *Nat Rev Cancer* 2009;9:293–302.
- de Krijger I, Mekenkamp LJ, Punt CJ, Nagtegaal ID. MicroRNAs in colorectal cancer metastasis. *J Pathol* 2011;224:438–47.
- Lim LP, Lau NC, Garrett-Engle P, Grimson A, Schelter JM, Castle J, et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 2005;433:769–73.
- Sood P, Krek A, Zavolan M, Macino G, Rajewsky N. Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc Natl Acad Sci U S A* 2006;103:2746–51.
- Bartels CL, Tsongalis GJ. MicroRNAs: novel biomarkers for human cancer. *Clin Chem* 2009;55:623–31.
- Lechpammer M, Sgroi DC. Laser Capture Microdissection: a rising tool in genetic profiling of cancer. *Expert Rev Mol Diagn* 2004;4:429–30.
- 't Hoen PA, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH, de Menezes RX, et al. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Res* 2008;36:e141.
- Lin M, Chen W, Huang J, Gao H, Ye Y, Song Z, et al. MicroRNA expression profiles in human colorectal cancers with liver metastases. *Oncol Rep* 2011;25:739–47.
- Rohr C, Kerick M, Fischer A, Kuhn A, Kashofer K, Timmermann B, et al. High-throughput miRNA and mRNA sequencing of paired colorectal normal, tumor and metastasis tissues and bioinformatic modeling of miRNA-1 therapeutic applications. *PLoS ONE* 2013;8:e67461.
- Schee K, Lorenz S, Worren MM, Gunther CC, Holden M, Hovig E, et al. Deep sequencing the MicroRNA transcriptome in colorectal cancer. *PLoS ONE* 2013;8:e66165.
- Schepeler T, Reinert JT, Ostefeld MS, Christensen LL, Silaharoglu AN, Dyrskjot L, et al. Diagnostic and prognostic microRNAs in stage II colon cancer. *Cancer Res* 2008;68:6416–24.
- Zhang JX, Song W, Chen ZH, Wei JH, Liao YJ, Lei J, et al. Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. *Lancet Oncol* 2013;14:1295–306.
- Maak M, Simon I, Nitsche U, Roepman P, Snel M, Glas AM, et al. Independent validation of a prognostic genomic signature (ColoPrint) for patients with stage II colon cancer. *Ann Surg* 2013;257:1053–8.
- Zeestraten EC, Reimers MS, Saadatmand S, Dekker JW, Liefers GJ, van den Elsen PJ, et al. Combined analysis of HLA class I, HLA-E and HLA-G predicts prognosis in colon cancer patients. *Br J Cancer* 2014;110:459–68.
- Zeestraten EC, Maak M, Shibayama M, Schuster T, Nitsche U, Matsushima T, et al. Specific activity of cyclin-dependent kinase 1 is a new potential predictor of tumour recurrence in stage II colon cancer. *Br J Cancer* 2012;106:133–40.
- McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer* 2005;93:387–91.
- Buermans HP, Ariyurek Y, van OG, den Dunnen JT, 't Hoen PA. New methods for next generation sequencing based microRNA expression profiling. *BMC Genomics* 2010;11:716.
- Nolte-'t Hoen EN, Buermans HP, Waasdorp M, Stoorvogel W, Wauben MH, 't Hoen PA. Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. *Nucleic Acids Res* 2012;40:9272–85.
- de Hoon MJ, Taft RJ, Hashimoto T, Kanamori-Katayama M, Kawaji H, Kawano M, et al. Cross-mapping and the identification of editing sites in mature microRNAs in high-throughput sequencing libraries. *Genome Res* 2010;20:257–64.
- Garmire LX, Subramaniam S. Evaluation of normalization methods in mammalian microRNA-Seq data. *RNA* 2012;18:1279–88.
- Garmire LX, Subramaniam S. The poor performance of TMM on microRNA-Seq. *RNA* 2013;19:735–6.

34. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007;26:2389–430.
35. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–74.
36. Huijbers A, Tollenaar RA, Pelt GW, Zeestraten EC, Dutton S, McConkey CC, et al. The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. *Ann Oncol* 2013;24:179–85.
37. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaiharu N, et al. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA* 2008;299:425–36.
38. Yantiss RK, Goodarzi M, Zhou XK, Rennert H, Pirog EC, Banner BF, et al. Clinical, pathologic, and molecular features of early-onset colorectal carcinoma. *Am J Surg Pathol* 2009;33:572–82.
39. Kjaer-Frifeldt S, Hansen TF, Nielsen BS, Joergensen S, Lindebjerg J, Soerensen FB, et al. The prognostic importance of miR-21 in stage II colon cancer: a population-based study. *Br J Cancer* 2012;107:1169–74.
40. Nielsen BS, Jorgensen S, Fog JU, Sokilde R, Christensen IJ, Hansen U, et al. High levels of microRNA-21 in the stroma of colorectal cancers predict short disease-free survival in stage II colon cancer patients. *Clin Exp Metastasis* 2011;28:27–38.
41. Karaayvaz M, Pal T, Song B, Zhang C, Georgakopoulos P, Mehmood S, et al. Prognostic significance of miR-215 in colon cancer. *Clin Colorectal Cancer* 2011;10:340–7.
42. Caruso S, Bazan V, Rolfo C, Insalaco L, Fanale D, Bronte G, et al. MicroRNAs in colorectal cancer stem cells: new regulators of cancer stemness? *Oncogenesis* 2012;1:e32.
43. Kurokawa K, Tanahashi T, Iima T, Yamamoto Y, Akaike Y, Nishida K, et al. Role of miR-19b and its target mRNAs in 5-fluorouracil resistance in colon cancer cells. *J Gastroenterol* 2012;47:883–95.
44. Chang KH, Mestdagh P, Vandesompele J, Kerin MJ, Miller N. MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. *BMC Cancer* 2010;10:173.
45. Li Q, Zou C, Zou C, Han Z, Xiao H, Wei H, et al. MicroRNA-25 functions as a potential tumor suppressor in colon cancer by targeting Smad7. *Cancer Lett* 2013;335:168–74.
46. Nishida N, Nagahara M, Sato T, Mimori K, Sudo T, Tanaka F, et al. Microarray analysis of colorectal cancer stromal tissue reveals upregulation of two oncogenic miRNA clusters. *Clin Cancer Res* 2012;18:3054–70.
47. Polisenio L, Salmena L, Riccardi L, Fornari A, Song MS, Hobbs RM, et al. Identification of the miR-106b~25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation. *Sci Signal* 2010;3:ra29.
48. Baffa R, Fassan M, Volinia S, O'Hara B, Liu CG, Palazzo JP, et al. MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. *J Pathol* 2009;219:214–21.
49. Wu ZS, Wu Q, Wang CQ, Wang XN, Wang Y, Zhao JJ, et al. MiR-339-5p inhibits breast cancer cell migration and invasion in vitro and may be a potential biomarker for breast cancer prognosis. *BMC Cancer* 2010;10:542.
50. Watahiki A, Wang Y, Morris J, Dennis K, O'Dwyer HM, Gleave M, et al. MicroRNAs associated with metastatic prostate cancer. *PLoS ONE* 2011;6:e24950.