

# Prognostic biomarkers in colorectal cancer with emphasis on microRNA

**Kristina Schee**

Department of Tumor Biology

Institute for Cancer Research

Oslo University Hospital

The Norwegian Radium Hospital

Faculty of Medicine, University of Oslo



© **Kristina Victorija Zarins Schee, 2012**

*Series of dissertations submitted to the  
Faculty of Medicine, University of Oslo  
No. 1457*

ISBN 978-82-8264-206-4

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.  
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika publishing.  
The thesis is produced by Akademika publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.



# TABLE OF CONTENTS

ACKNOWLEDGEMENTS	4
ABBREVIATIONS	6
LIST OF PUBLICATIONS	7
AIMS OF THE STUDY	8
GENERAL INTRODUCTION	
1. CANCER	9
2. COLORECTAL CANCER	11
a. Epidemiology	12
b. Classification	14
c. Metastasis	15
d. Treatment	17
e. Early detection and screening	20
f. Prognostic biomarkers	26
g. Predictive biomarkers	31
3. CEA	33
4. MICRORNA	35
a. MiRNA biogenesis and function	35
b. MiRNA and cancer	38
c. MiRNA in colorectal cancer	39
d. MiRNA as early detection biomarkers in serum, plasma and feces	42
e. MiRNAs as prognostic biomarkers	43
f. MiRNAs as predictive biomarkers	45
g. MiRNAs as therapeutic targets	47
SUMMARY OF PAPERS	
Paper I	49
Paper II	50
Paper III	51
METHOLOGICAL CONSIDERATIONS	
a. Clinical samples-sample size and origin.	53

b. Antibody specificity	57
c. MiRNA quantitative RT-PCR and normalization	63
d. Deep sequencing and genome mapping	64
RESULTS AND DISCUSSION	66
CONCLUDING REMARKS AND FUTURE PROSPECTS	72
REFERENCES	74
PAPERS	

## ACKNOWLEDGEMENTS

This work was carried out at the department of Tumor biology, Oslo University Hospital, Radiumhospitalet in the years 2008-2012.

First and foremost I would like to express my gratitude to my supervisor Kjersti Flatmark. Your broad knowledge of colorectal cancer and systematic approach to science has been an inspiration. Thank you for always being available, for your super fast replies to all kinds of questions, for your optimism, and for your understanding and support during difficult times.

I would also like to thank my co-supervisor and group leader, Øystein Fodstad, for providing an excellent research environment and for introducing me to research abroad.

Thanks to all the co-authors of my papers. I have learned a lot from our discussions and the work we have done together. Thank you, Kjetil Boye, for guiding me through SPSS and the world of statistics. Without the large set of RNA samples I would not have gotten far, so a warm thanks to Torveig Weum Abrahamsen whose help on preparing the patient samples was invaluable.

To my colleagues at the department of Tumor Biology; you are a great bunch. Thank you for creating an outstanding scientific and social environment. It is a pleasure to work with you and I am grateful for all the help with theoretical questions and practical lab work. The mornings would not be as pleasurable without the early intake of coffee and “serious” scientific

discussions on the latest news on blood diets, how eggs are as unhealthy as smoking and other life altering matters. So a warm thanks Ane Sager Longva and Karen-Marie Heintz.

To my family, thank you for your support. I love you all. To the love of my life, Henning, thank you for your patience and understanding and for all the hugs and kisses when things didn't go my way.

I would like to dedicate this thesis to my father, Peter Vilhelm Zarins. His love and curiosity for biology and biochemistry was what got me interested in the field. He himself was a force of nature and I miss him dearly.

-“Papsen, this one is for you”

*Kristina Schee*

Oslo, August 2012

## ABBREVIATIONS

5-FU	5-fluorouracil	mRNA	messenger RNA
APC	Adenomatous polyposis coli	MiRNA	microRNA
CDK	cyclin-dependant kinase	MS	mass spectrometry
CEA	carcinoembryonic antigen	MSI	microsatellite instability
CIN	chromosomal instability	MSI-H	microsatellite instable-high
CIMP	CpG island methylator phenotype	MSS	microsatellite stable
CRC	colorectal cancer	NCA-2	nonspecific cross-reacting antigen-2
DCC	deleted in colorectal cancer	NPV	negative predictive value
DHFR	dihydrofolate reductase	OS	overall survival
DFS	disease free survival	PSG	pregnancy-specific glycoprotein
DNA	deoxyribonucleic acid	PPV	positive predictive value
EGFR	endothelial growth factor receptor	RISC	RNA-induced silencing complex
FOBT	fecal Occult Blood Test	RNA	ribonucleic acid
FAP	familial adenomatous polyposis	RT-PCR	reverse transcription polymerase chain reaction
HNPCC	hereditary Nonpolyposis Colorectal Cancer	TNM	tumor-node-metastasis
IFMA	immunofluorometric assay	TGF- $\beta$	transforming-growth-factor- $\beta$
KRAS	Kirsten rat sarcoma gene	TRBP	transcription binding protein
MAb	monoclonal antibody	TS	thymidylate synthase
MAPK	mitogen-activated protein kinase	UTR	untranslated region
MMR	mismatch repair	VEGF	vascular endothelial growth factor
		WB	western immunoblot



## LIST OF PUBLICATIONS

- I. **Schee K**, Flatmark K, Holm R, Boye K, Paus E. Tumour Biol. 2012 Feb;33(1):73-83. Epub 2011 Oct 18.  
  
Investigation of nonspecific cross-reacting antigen 2 as a prognostic biomarker in bone marrow plasma from colorectal cancer patients. Tumour Biol. 2012 Feb;33(1):73-83
- i. **Schee K**, Flatmark K, Holm R, Boye K, Paus E. Tumour Biol. 2012 Aug;33(4):1263-Erratum to: Investigation of nonspecific cross-reacting antigen 2 as a prognostic biomarker in bone marrow plasma from colorectal cancer patients.
- II. **Schee K**, Boye K, Abrahamsen TW, Flatmark K.  
  
Clinical relevance of miR-21, miR-31, miR92a, miR-101, miR-106a and miR-145 in colorectal cancer. (Under revision, BMC Cancer)
- III. **Schee K**, Lorenz S, Molton Worren M, Günther C-C, Holden M, Hovig E, Fodstad Ø, Meza-Zepeda L, Flatmark K.  
  
Deep sequencing the miRNA transcriptome in colorectal cancer. (Manuscript)

### ***Related publications not included in the thesis***

- I. **Schee K**, Fodstad Ø, Flatmark K. Am J Pathol. 2010 Oct;177(4):1592-9. MicroRNAs as biomarkers in colorectal cancer.

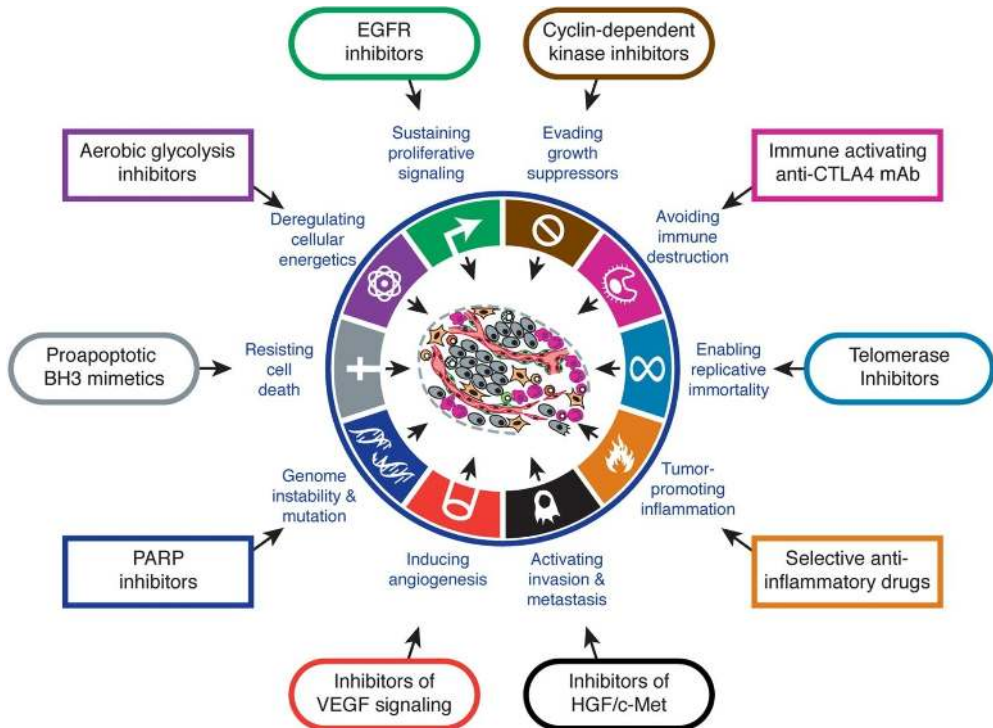
## **AIMS OF THE STUDY**

The overall objective of this work was to search for prognostic biomarkers in colorectal cancer:

1. To assess whether CEA and NCA-2 in bone marrow plasma were prognostic biomarkers in CRC
2. To investigate miRNA expression in our CRC tumor panel and assess whether any miRNAs presented themselves as potential prognostic markers in CRC.

# 1. CANCER

The first documented case of cancer was recorded on a papyrus roll in Egypt around 1500 B.C., but the term cancer appeared much later used by the Greek physician Hippocrates at around 300 B.C. There have been many different theories through the ages as to what causes cancer; the Egyptians believed it was caused by the Gods, while the Greek physicians believed it was caused by black bile, one of the body's four fluids. It wasn't until the 17<sup>th</sup> century, after the discovery of the lymphatic system, that Hippocrates black bile theory was challenged, and other theories surfaced such as cancer being the result of trauma or parasites. In the late 19<sup>th</sup> century Rudolph Virchow recognized that cancerous cells were derived from other cells. Today the leading theory is that cancer is the result of both hereditary and environmental influences that change gene expression leading to uncontrolled cell growth. While the sequence in which a cancer cell acquires the necessary traits for tumor formation can vary, there are a number of mechanisms that are common to most, if not all, cancers. These are deemed "The hallmarks of cancer"; proliferative signaling, evading growth suppressors, avoiding immune destruction, enabling replicative immortality, tumor promoting inflammation, activating invasion and metastasis, inducing angiogenesis, genome instability and mutation, resisting cell death and deregulating cellular energetics (Figure 1) [1]. A single change is inadequate to upset the balance in the cell, so multiple processes need to be deregulated in order to cause cancer.



**Figure 1.** The Hallmarks of cancer, published with permission from Elsevier [1]. The outer boxes describe possibilities for therapeutic targeting.

## 2. COLORECTAL CANCER

Cancers which arise in the colon or rectum have many features in common and are therefore referred together as colorectal cancer (CRC). The development of CRC is a multistep process that involves accumulation of changes in expression levels in both oncogenes and tumor suppressor genes. These changes can be mutations in the genes themselves or silencing of gene expression by epigenetic mechanisms. Most (~75%) CRCs arise sporadically, while the remaining 25% of the patients have a family history, and only 5-6% of CRC cases are due to inherited conditions (Lynch syndrome previously Hereditary Nonpolyposis Colorectal Cancer; HNPCC) [2]. It is therefore suggested that the complex processes leading to CRC occur through interactions between genes, environmental factors and lifestyle [2-4]. The adenoma-carcinoma sequence refers to a traditional view on the progression of CRC from adenomatous polyps to invasive cancer. The stepwise progression of CRC was already suggested in 1974 [5], and some years later Vogelstein et al [6] described genetic alterations in several genes accompanying the stepwise progression from benign adenoma to malignant carcinoma where certain mutations, like APC, BRAF, KRAS, TP53, TGF- $\beta$  and PI3K among others [7], were directly related to distinct stages of tumor development [8] (Figure 2). Although there is evidence that cancer arises from adenomas, it is widely recognized that only a small proportion of adenomas progress to invasive CRC, and in large they appear to stabilize their progression and even regress [9]. It would be preferential to be able to distinguish the adenomas with a potential towards malignancies but there are no biomarkers available today to distinguish them from one another [10].

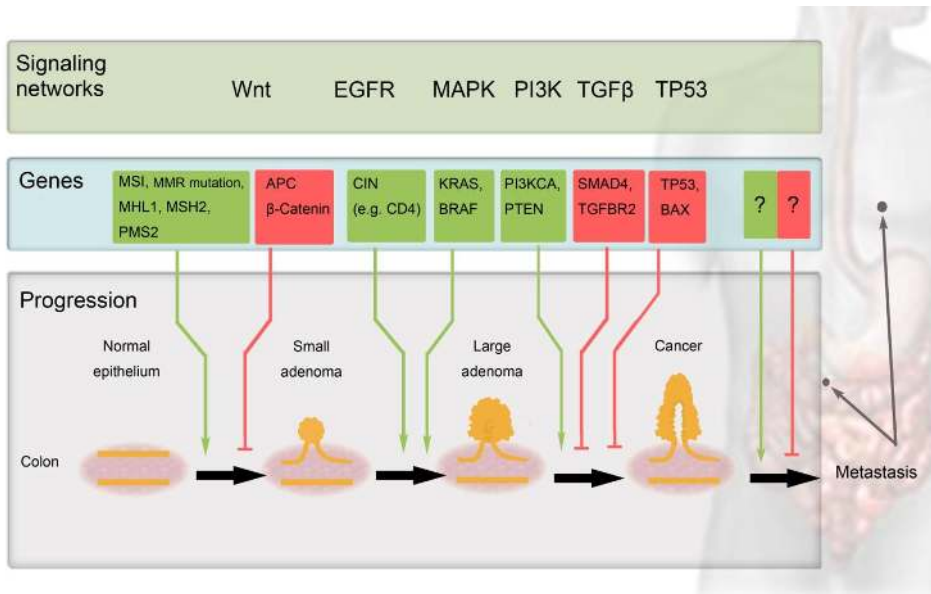
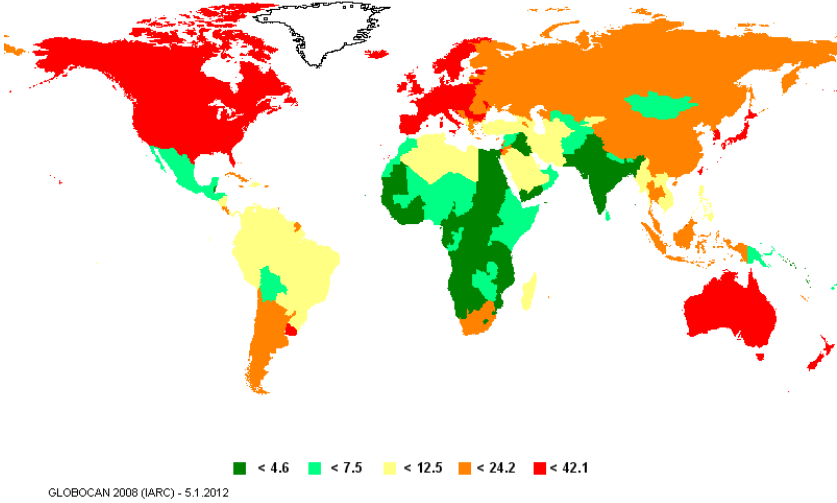


Figure 2. Cancer progression adapted from Markowitz et al (2009) [7].

### a) *Epidemiology*

CRC is one of the leading causes of cancer morbidity and mortality in the Western world and is equally distributed between the sexes. It is a disease that increases with age and about 90% are diagnosed after the age of 55. The incidence rate and mortality varies greatly between the different continents of the world due to difference in dietary factors, lifestyles [11] and ethnicity (Figure 3). In Norway the incidence rate has steadily increased in the past 50 years (Figure 4), but the mortality rate has subsided probably due to earlier detection and better treatments [12]. Intensified diagnostic surveillance, along with improved treatment strategies involving surgery, chemo- and radiotherapy have contributed to earlier detection and improved survival

from CRC. However, almost 50% of patients diagnosed with CRC will die of the disease, mainly because of metastasis development, most commonly in the liver and lungs.



**Figure 3.** World incidence rate for CRC age-standardized incidence rate per 100.000 both sexes, all ages.

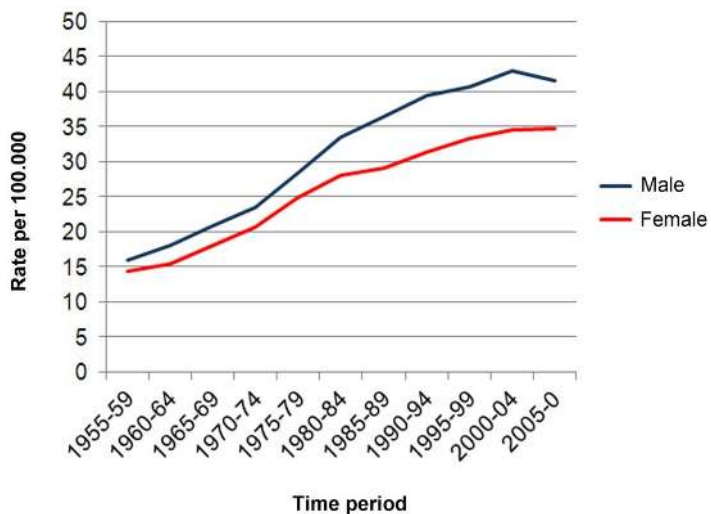


Figure 4. Incidence of CRC in Norway in both sexes (age adjusted incidence rates per 100.000) [12].

## b) Classification

CRC is commonly classified according to either Dukes' [13] or to American Joint Committee on Cancer tumor-node-metastasis (TNM) staging system based on tumor growth in the bowel wall at diagnosis (T), the presence of regional lymph node metastasis (N) and the presence of distant metastasis (M) (Table 1) [14]. Either classification system can be based on clinical information such as physical examination, radiologic and other specialized imaging, as well as surgical and histopathological findings. The purpose of classification is to categorize tumors with the aim to predict prognosis and aid in the choice of treatment. However, utilizing only clinical,



radiological and pathological evaluation results in failure to identify high- and low-risk patients, and patients may receive either insufficient or unnecessary treatment [15].

The staging system was in 2009 revised adding more subgroups. Whether the new additions are an asset or only make the classification for daily clinical use more complex is still debated [16].

TNM stage	T	N	M	Dukes Stage
I	T1, T2	N0	M0	A
II A	T3	N0	M0	B
II B	T4	N0	M0	B
III A	T1, T2	N1	M0	C
III B	T3, T4	N1	M0	C
III C	All T	N2	M0	C
IV	All T	All N	M1	D

**Table 1.** A table of the TNM stages with the corresponding Duke’s stages. Table is modified from AJCC [14]. T=tumor stage, N=lymph node metastasis, M=metastasis.

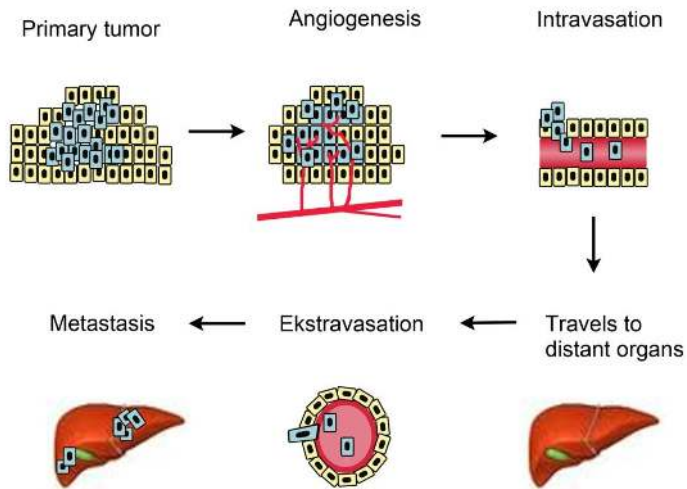
### c) *Metastasis*

Metastasis is the process by which cancer cells spreads from the primary tumor to invade and grow in a distant organ. These metastatic tumor cells may progress at once or lay dormant for years after treatment for the original lesion. The cause of death for the vast majority of cancer

patients is the development of metastasis, and in Norway ~90% die of CRC metastasis [12]. CRC metastasizes mainly to the liver and lung and the metastatic tumors are heterogeneous, consisting of multiple subpopulations of cells making targeted treatment difficult resulting in poor treatment response.

Already in 1889, Paget postulated the “seed and soil” theory based on his examination of 735 autopsy records of women with breast cancer where he concluded that the metastatic process was not random but that the tumor cells (seeds) needed good growing conditions (soil) [17]. Later it was postulated that metastatic end-points were based on the vascular system and total blood flow [18] but this did not explain the reason why some tumors preferably metastasize to specific organs not in the vicinity of the vasculatory system [19]. Today the belief is that the metastatic location is more likely an interaction between the genetics of the patient, the tumor cells and the local microenvironment at the secondary site [20].

There are several criteria that must be fulfilled in order for metastasis to occur. First the tumor requires nutrients as it grows and therefore it needs to establish new blood vessels (angiogenesis). As the tumor grows it invades nearby tissue. Thereafter, the tumor cells must gain the ability to detach and enter the bloodstream (intravasation), where they must survive before exiting the blood vessels (extravasation) to enter and proliferate in the secondary organ and form metastasis (Figure 5). The ability of a tumor cell to metastasize is dependent on a combination of growth factors, motility and angiogenic factors produced by the tumor cell itself and/or by the host cells in the distant organ [21].



**Figure 5.** The metastatic process from primary tumor to the development of angiogenesis, intravasation, traveling to distant organs, ekstravasation before developing metastasis.

## d) *Treatment*

The main curative treatment modality for colon and rectal cancer is surgery. Treatment decisions are based on the extent of the disease at the time of diagnosis where staging plays an essential role. In addition to surgery rectal cancer patients receive a combination of chemotherapy and radiation (chemoradiotherapy), while selected colon cancer patients may receive chemotherapy. Chemotherapy can be given either before surgery (neo-adjuvant) or after surgery (adjuvant). The reason for giving chemotherapy before surgery may be to shrink the tumor to facilitate complete removal, but also in the hope of preventing metastasis

development. In some cases, rectal cancer patients receiving neoadjuvant chemoradiation achieve a complete clinical response where no tumor can be found on the clinical assessment before surgery. Some studies have therefore included a “watch and wait” policy with close clinical follow up where surgery is only given if local recurrence manifests, however no difference in 5-year survival between “waiting” and surgery has been found [22]. Also, chemoradiotherapy exhibits reduction in local recurrence rate and further clinical trials are ongoing [23, 24].

After surgery, there is a potential risk for microscopic residual tumor cells remaining in target organs, and therefore some patients receive adjuvant chemotherapy with either 5 fluorouracil (5-FU) or 5-FU in combinations with Leucovorin (FLV regimen) or Oxaliplatin (FLOX regimen). 5-FU is an inhibitor of thymidylate synthase (and therefore DNA synthesis) and since cancer cells are faster growing this drug affects cancer cells greater than normal cells. Several studies have shown that stage III colon cancer patients benefit from adjuvant chemotherapy; but this has not been proven for stage II patients, and is controversial in rectal cancer [14, 25-27]. Therefore only stage III patients receive adjuvant chemotherapy, although there are some stage II patients that get recurrence and probably would benefit from this treatment. Combinations of different drugs like 5-FU and leucovorin or oxaliplatin have been shown to improve disease free-survival (DFS) by ~3-4% in stage II and ~10% in stage III patients [27-29]. In metastatic CRC patients, different combinations of chemotherapeutic drugs have prolonged median survival from about 10-12 months to > 20 months [30, 31].

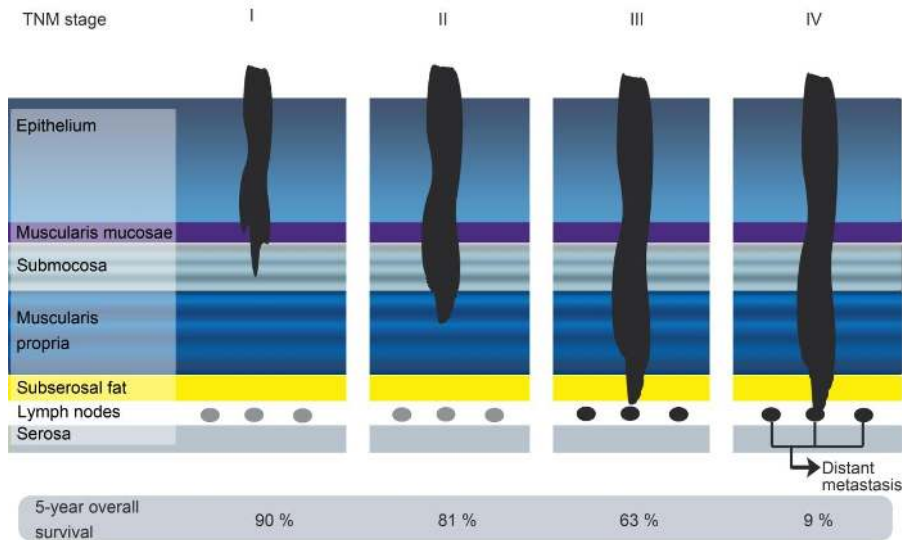
Targeted therapy involves drugs designed to inhibit specific genes and disrupt certain biological pathways. A recent attempt to control malignant proliferation and spread included the inhibition of neoangiogenesis, the development of new blood vessels. The most successful anti-angiogenic drug has focused on inhibiting vascular endothelial growth factor (VEGF), a protein which stimulates blood vessel formation. This drug is called bevacizumab and given in combination with drugs mentioned above increased median overall survival in metastatic CRC patients from 15.6 months to 20.3 months [32]. But like with any other drug, anti-VEGF treatment has severe side effects which include hypertension, bleeding, and vomiting, and several questions remain unanswered regarding the duration, which chemotherapy regimens to be used in combination and whether bevacizumab should be used as maintenance therapy, and/or in patients who are progressing on treatment.

Epidermal growth factor (EGFR) is a transmembrane glycoprotein involved in signaling pathways, affecting cellular growth, proliferation, differentiation and programmed cell death. The receptor is present on the cell surface in up to 80% of tumors and overexpression has been associated with poorer prognosis in CRC [33, 34]. Cetuximab and panitumumab are drugs which inhibit EGFR, but only a subset of patients treated respond to the treatment, thus identification and characterization of molecular markers to predict tumor response have been an area of interest. CRC patients with *KRAS* mutations appear to be relative resistant to treatment with cetuximab and panitumumab, with lower response rates and poorer survival [12, 35, 36]. *KRAS* mutation status is used as a predictive biomarker in the clinic today and helps define a subset of patients who will derive benefit from treatment with EGFR.

It has become essential to identify high-risk subgroups of CRC patients who may benefit from adjuvant therapy to avoid a potentially toxic over-treatment and an unprofitable financial burden for the health care system [24]. The optimal sequence of administration of the above mentioned drugs remains under investigation in patients with metastatic CRC, but the success in prolonged survival of patients with metastatic disease is also translating to improve cure rates among patients with stage III disease. The goal of ongoing investigations of adjuvant trials is to further improve survival rates by the use of chemotherapeutic drugs.

### **e) *Early detection and screening***

The survival rate of patients decreases with increasing CRC stage pointing out the need to detect cancer at an early stage (Figure 6). However, the life-time risk of developing CRC in the Western world is only 5%, meaning that 95% will never develop the disease, which poses particular challenges regarding early detection and screening. Early detection can be achieved by assessing risk for developing the disease (in CRC the risk increases with increased TNM stage see Figure 6), recognition of the disease by increasing clinical awareness in the population (what signs to look for) and screening programs.



**Figure 6. Associations between TNM stage and overall survival.** TNM stage marked with survival in percentages for the individual stages. 5-year overall survival decreases with increased TNM stage. Survival numbers are taken from the Norwegian Cancer Registry [12].

Screening refers to the use of diagnostic tests across a healthy population in order to identify individuals who have the disease, but do not yet have symptoms. The strength of a screening test is based on its sensitivity and specificity [37]. The sensitivity of a clinical test refers to the ability of the test to correctly identify those patients with the disease, and this can be calculated as a percentage (Figure 7). A test with 80% sensitivity would detect 80% with the disease (true positives) and while 20% with the disease would go undetected (false negatives).

$$\text{Sensitivity} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

$$\text{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{False positives}}$$

**Figure 7. Calculation of sensitivity and specificity.** True positives = the patient has the disease and tests positive, False negative = the patient has the disease but tests negative, True negative = the patient does not have the disease and tests negative, False positive = the patient tests positive but does not have the disease.

A high sensitivity is clearly important when it comes to screening as it is preferable to detect as many positive patients as possible. The specificity of a clinical test refers to the ability of the test to correctly identify patients without the disease. A test with 80% specificity correctly reports 80% of patients without the disease as test negative (true negatives) but 20% patients without the disease are incorrectly identified as test positive (false positives) (Figure 7). The sensitivity and specificity are not dependent on the prevalence of the disease in the given population. To identify the strength of a screening test in a specific population, the positive predictive value and the negative predicted values are calculated. A positive predictive value determines how likely it is for a positive test result to actually be positive, while a negative predictive value determines how likely it is for a negative test result to be negative. Both positive and negative predictive values are dependent on the prevalence of the disease in the population of interest (Figure 8) [38].



		Patients with metastasis		
		Condition Positive	Condition Negative	
CEA measurements in bone marrow plasma at time of surgery	Test Outcome Positive ( $\geq 5\mu\text{g/l}$ )	<b>True Positive</b> (TP) = 33	<b>False Positive</b> (FP) = 33	<b>Positive predictive value</b> $= \text{TP} / (\text{TP} + \text{FP})$ $0,5$ $= 50\%$
	Test Outcome Negative ( $< 5\mu\text{g/l}$ )	<b>False Negative</b> (FN) = 50	<b>True Negative</b> (TN) = 149	<b>Negative predictive value</b> $= \text{TN} / (\text{FN} + \text{TN})$ $0,75$ $= 75\%$
		<b>Sensitivity</b> $= \text{TP} / (\text{TP} + \text{FN})$ $0,38$ $= 38\%$	<b>Specificity</b> $= \text{TN} / (\text{FP} + \text{TN})$ $0,81$ $= 81\%$	

**Figure 8.** An example of how to calculate sensitivity, specificity, positive- and negative predictive value. Values are taken from Paper I. CEA measurements have been taken at time of diagnosis and the endpoint is metastasis. Total number of patients in the cohort = 265. Figure adapted from [37]

The goal of screening for CRC is to detect early stage cancer thereby enabling early treatment and increasing survival [39]. The two most common forms of screening today are colonoscopy and Fecal Occult Blood Test (FOBT). FOBT is an over-the-counter test which detects occult blood in feces. The use of FOBT has been shown in different randomized trials to reduce cancer mortality [39] but their application in screening is still debated [40] as test from different

manufacturers has a variable sensitivity (40%-80%) resulting in many false negatives [41, 42].

The advantages and disadvantages for colonoscopy and FOBT are listed in Table 2.

	Advantages	Disadvantages
<b>Fecal occult blood test</b>	Relatively simple to perform Inexpensive Confirmed reduced mortality	Low sensitivity Certain foods give false positives Not efficient at detecting precursor lesions.
<b>Colonoscopy</b>	Sensitive and Specific Able to detect precursor lesions Possible to remove polyps	Expensive Invasive; requires bowel Requires educated personnel Complications like bleeding in the bowel may occur

**Table 2. Advantages and disadvantages of FOBT and colonoscopy.**

There are two types of FOBT tests; one that measures guaic fecal occult test (gFOBT) or immunohistochemical test (iFOBT) [43]. iFOBTs have shown to have a higher sensitivity compared to gFOBT, but they are more expensive and no randomized trials have investigated the effect on incidence or mortality.

Other fecal tests include fecal DNA tests which detect mutant or abnormal DNA shed from neoplastic CRC lesions excreted in feces [43]. Examining two DNA markers; hypermethylated vimentin gene (hV) and a two site DNA integrity assay (DY), the training set demonstrated high

sensitivity (88%) for CRC with a specificity of 82% [44]. Other DNA markers include methylations in *APC*, *KRAS* and *TP53* among others [45]. The downside with these methods is that DNA is easily degradable in feces and requires buffer application for stabilization [46].

Mutations in the tumor suppressor gene *APC* is considered to be one of the most common early events in development of CRC and is present in most sporadic forms of CRC [47]. When occurring in germline cells *APC* mutations give rise to Familial Adenomatous Polyposis (FAP), an inheritable cancer-predisposition where hundreds of polyps may develop and the risk of CRC by the age of 40 is almost 100% [48]. The protein encoded by *APC* is part of a degradation complex that among other things regulates  $\beta$ -catenin, a protein that initiates transcription in the Wnt signaling pathway [49]. Without the regulation of  $\beta$ -catenin Wnt target genes are constantly activated leading to increased proliferation [47]. The high frequency and inability of segregating patients into subgroups makes *APC* mutations unsuitable as a prognostic marker, but the high frequency in CRC coupled with the absence of *APC* mutations found in a study on fecal markers in healthy people implicate this as a potential early detection marker [50].

Since FOBT is not specific for CRC, a positive test is always followed by bowel visualization technique like colonoscopy. Colonoscopy is a complete endoscopic examination of the colon with the ability to remove detected polyps and obtain biopsy samples. This method may require sedation of the patient and bowel preparation [51]. In case-control studies, colonoscopy is associated with reductions in the incidence of and mortality from CRC [52-54] with a 10-year follow up.

CRC is a good candidate for a screening program as the disease is quite common, it is mainly contained within an age group and early detection, and therefore earlier treatment, is associated with increased survival.

## **f) *Prognostic biomarkers***

Poor survival is associated with increased TNM stage in CRC, and the TNM staging system remains the gold standard of prognostic factors in CRC (Figure 6). However, patients within the same TNM subgroups may have different prognosis and respond differently to therapy, making it a suboptimal system. New biomarkers which can distinguish molecular subtypes and predict the patient prognosis are needed. Although in the recent years a huge amount of research has been devoted to finding prognostic biomarkers, none have emerged that are clearly better than existing ones and the search continues. Nevertheless, there have been many promising biomarker candidates within mutational phenotypes like chromosomal instability (CIN) and microsatellite instability (MSI), methylation in CpG island promoters (CIMP) mutations in single genes, and dysregulation of proteins like EGFR, RAS/RAF, PI3K and TP53. A short overview of a selection of promising biomarker candidates is given below.

### ***Chromosomal instability (CIN)***

CIN is the most common type of genomic instability observed in cancer and occurs in 80%-85% of CRC. It causes numerous changes in chromosomal copy number either by chromosomal gains or losses, substitutions and rearrangements, but the cause of this form of genetic instability

remains unknown [55]. It is not the mere accumulation of events that causes the progression to cancer, but rather the specific combinations of chromosomal abnormalities that increase with tumor progression [56, 57]. A promising prognostic biomarker is the loss of chromosome 18q, where among other genes the tumor suppressor gene “deleted in colorectal carcinoma” (DCC) is located. DCC has structural features in common with certain types of cell-adhesion molecules and may participate with other proteins in cell-cell and cell-matrix interactions, and loss of this gene might contribute to tumor growth and invasion [58]. Loss of the long arm on chromosome 18q has been associated with poor prognosis [59]. However, numerous methods have been used among the different studies and although they agree that loss of 18q is associated with poor survival the estimates of the prognostic value have varied considerably [60, 61] therefore improved consistency within methodology is needed to precisely quantify its effect and role in patients with stage II-III disease.

### ***Microsatellite instability (MSI)***

Nucleotide repeats (microsatellites) are distributed all over the DNA and are prone to DNA slippage during replication, requiring the mismatch repair system (MMR) to correct such errors. Microsatellite instability (MSI) is associated with the loss of function of the DNA MMR (by methylation) and occurs in about 15%-20% of patients with sporadic CRC and >95% of patients with Lynch syndrome [62]. Patients with MSI have a generally favorable prognosis, and MSI could be used as a marker to identify patients who might benefit from surgery alone, but the

low abundance of MSI patients leaves an uncertainty of the usefulness of MSI as a prognostic biomarker [60].

### ***CpG island methylator phenotype (CIMP)***

Methylation of gene promoters can mimic genetic deletions by silencing expression of tumor suppressor genes thereby promoting malignant progression. The CpG island methylator phenotype (CIMP) [63] has been implicated as a possible prognostic biomarker, but there remains controversy whether this is due to the phenotype itself or because this phenotype is associated with other factors known to affect prognosis in CRC [64-66]. Different subgroups of CIMP phenotype also show MSI, *BRAF*, *KRAS* and *TP53* mutations and there is speculation if the relationship between CIMP and prognosis is actually a relationship between *KRAS/BRAF* mutation and prognosis [67].

### ***Dysregulation of gene and protein expression***

#### ***Gene signatures***

There have been attempts at generating prognostic DNA gene- signatures based on microarray gene expression measures using mRNA. Studies with gene signatures ranging from 23-43 genes resulted in 78%-90% accuracy [68-71], but these gene signatures need validation in large CRC cohorts. Another study based the prognostic stratification test on a smaller gene signature of only 7 genes, called ColoGuidePro, could seemingly predict prognosis, tumor stage and other

clinicopathological parameters [72], but there is still some way to go before this can be applied in the clinic.

### **EGFR**

One important pathway affected in CRC is the endothelial growth factor receptor (EGFR) pathway. EGFR is a transmembrane protein that phosphorylates and activates other intracellular proteins mediating cellular growth, differentiation, and proliferation through the activation of two major pathways; the RAS/RAF mitogen activated protein kinase (MAPK) pathway and the phosphatidylinositol 3-kinase (PI3K) pathway [73]. EGFR has been implicated in CRC tumorigenesis, tumor progression and metastasis but the prognostic relevance remains controversial [74, 75].

### ***RAS/BRAF***

The RAS-RAF-MAPK signaling pathway lies downstream of EGFR and is a major pathway for tumor cell proliferation in CRC. *BRAF* and *KRAS* mutations tend to be mutually exclusive events in tumors [76]. Oncogenic mutations in *RAS* are seen in a large array of human cancers, but absolute specificity towards a single type of cancer has not been observed. Overexpression of *KRAS* is present in about 40%-50% of CRC cases and is believed to be an early event, but its presence also plays a role for the ability of cells to metastasize [35]. *KRAS* mutations are also responsible for lack of response to EGFR targeted treatment, but the prognostic value of *KRAS* mutations remains inconclusive [36, 77]. *BRAF* mutations occur more frequently in

Microsatellite instable (MSI) than in Microsatellite stable (MSS) tumors [78] and is also tightly associated with the CIMP phenotype[79], however *BRAF* mutations are only associated with a worse survival in MSS [80].

### ***PI3K***

PI3K is a family of enzymes involved in cellular functions such as cell growth, proliferation, differentiation, motility, survival and intracellular trafficking. *PIK3CA*, which encodes the p110 $\alpha$  subunit of PI3K, is often mutated in CRC and by binding with KRAS it activates the AKT pathway which in turn enhances cell proliferation and survival. The prognostic role of *PI3KCA* is still under investigation but mutation in *PI3KCA* has been proposed as a marker for poor overall survival [81].

### ***TP53***

*TP53* is a tumor suppressor gene which codes for the protein p53 and is located at chromosome 17p which is frequently deleted in CRC [82]. P53 induces apoptosis or programmed cell death when damage to DNA occurs, but when p53 is altered or mutated apoptosis does not occur and this may lead to unregulated cell growth and accumulation of mutated cells. Both alleles are usually inactivated by a combination of mutation and chromosome deletion. *TP53* mutation has been investigated as both prognostic factor and predictor of response to therapy, but despite multiple studies the prognostic value of *TP53* mutation remains undetermined [83, 84].



With an increasing understanding of tumor biology more prognostic marker candidates emerge, and validation in large cohorts will follow. Validation of new biomarkers may enable clinicians in the future to select patients with a worse prognosis for further tailored specific treatment, maximizing the drug efficacy and minimizing the severe side effects.

### **g) Predictive biomarkers**

The improvement of therapeutics in the last 20 years has improved survival, however these new therapeutics are expensive and have potential toxic side effects. This has instigated an interest in elucidation of predictive biomarkers with the intention to improve outcome by predicting response or resistance to specific therapies and to identify patients who are likely to develop severe toxic side effects. Patients with high levels of microsatellite instability (MSI-H) have in many studies been associated with a favorable prognosis, but it is however not clear if this is because MSI-H tumors are less aggressive or if they are more sensitive to chemotherapy (5-FU) treatment [85, 86]. *TP53* mutations have shown to reduce therapeutic response in 5-fluorouracil (5-FU) studies, but the potential predictive value of *TP53* mutations needs to be further investigated [83, 87].

A predictive biomarker which is used in the clinic today is the presence or absence of *KRAS* gene mutations which can determine the benefit of using epidermal growth factor receptor (EGFR) inhibiting antibodies [73, 75, 76, 88-93]. The inhibition of EGFRs was found to have antitumor

activity and a synergy with chemotherapy and radiotherapy [73, 75, 92, 93]. However, mutations in *KRAS* resulted in failure of anti-EGFR treatment [76, 89-91]. Therefore, mutant *KRAS* has been considered a predictive marker for anti-EGFR therapy resistance in metastatic CRC and implemented into clinical use [88].

In the future, there will probably be an increase in individual targeted therapy dependent on the patient's tumor characteristics, and not based upon the generality of the disease. Although there are immense potential implications, clinicians are currently unable to use these data in clinical practice for decision making because of a lack of definition, adequate validation, and easy implementation.

### 3. CEA

Carcinoembryonic antigen (CEA), a 180 kDa glycoprotein, remains the oldest and most widely used biomarker in patients with CRC [94]. It is a member of the immunoglobulin super-family and comprised of 29 genes/pseudogenes which are divided into two branches, defined as CEACAM and PSG (pregnancy-specific glycoprotein) [95]. CEACAM is further divided into subclasses CEACAM 1, 3, 4 which are generally anchored to the cell surface by cytoplasmic domains, and CEACAM 5-8 which are anchored through glycoposphatidylinositol lipid moieties [96]. The definite role of CEA in normal colon is still unclear but it is suggested that CEA may play a role in protecting the colon from microbial infection [97].

CEA is present at elevated levels in tumor tissue and in addition it is released into the blood stream and increased levels are detectable in serum or plasma of many patients with CRC. Serum levels of CEA are often low in early stages of CRC making it an insensitive test for screening and early detection [98, 99]. Even though many CRC patients do not have elevated levels of CEA at diagnosis, there exists a correlation between increased CEA levels and CRC stage [100]. Today serial serum measurements are used in the clinic to detect CRC recurrence with a specificity and sensitivity of 80% and 70%, respectively [101-104]. Its specificity as a biomarker is limited as elevated levels of CEA are also found in smokers, benign diseases (perforated ulcer and liver disease) as well as in patients with epithelial tumors of non-intestinal origin [94, 97], and sensitivity is affected by heterogeneous CEA tumor expression, as for instance some poorly differentiated tumors may not express CEA [99, 105, 106]. Several proteins have been tested

together with CEA in an attempt to improve the prognostic accuracy, but this has not proven useful [107, 108].

CEA assays have been established by several groups and it became evident that the presence of cross-reacting antigens led to discrepancies between the results when using different antibodies [109]. The discovery of one such molecule, the non-specific cross-reacting antigen 2 (NCA-2, a truncated version of CEACAM 5) resulted in several studies examining the cross reactivity to CEA. NCA-2 is a 160 kDa homologue of CEA which is transcribed from the same gene as CEA. It was first isolated from meconium in 1973 [110] and has since then been described as cross-reacting to CEA and never on its own merit [111]. NCA-2 and CEA are identical in the amino terminal sequence [112] and both are highly glycosylated, but they differ in glycosylation patterns [113, 114] which might influence conformational stability and protease resistance [113, 114]. Studies on the cross reactivity of NCA-2 has resulted in improved CEA assays that do not recognize NCA-2 [115] but still little is known about the specificity of NCA-2 antibodies in different experimental settings, and it has as far as we know not been explored for clinical associations in a large patient cohort [116-118].

## 4. MICRORNA

MicroRNAs (miRNAs) are small (20-22 nucleotides long) non-coding RNA sequences, evolutionary conserved, which post-transcriptionally regulate gene expression by binding to a target messenger RNA (mRNA) [119, 120]. MiRNAs were first discovered in genes that control timing of larval development in *C.Elegans* and have later been found to be involved in fine tuning many biological processes, like cell proliferation, differentiation, apoptosis and metabolism, but it is their involvement in cancer that has sparked their increased study [121]. There are today ~1500 human miRNA annotated in the miRNA registry called miRBase [122, 123], and although discovery of new miRNAs has increased exponentially in recent years, mRNA target identification has been a slow process.

### ***a) MiRNA biogenesis and function***

MiRNA genes are located within exons and intergenic regions but are mainly found within introns of coding or noncoding genes [124, 125]. MiRNAs are initially transcribed as longer precursors by RNA polymerase II. MiRNAs primary transcripts, called pri-miRNA, are processed by a ribonucleoprotein complex, Drosha/DGCR8, into a 70-100 nucleotide long sequence called pre-miRNA, with a hairpin stem-loop structure (Figure 9). This pre-miRNA structure is recognized by the nuclear export protein, Exportin 5, and exported in to the cytosol [126, 127] where it is further processed by the RNase III enzyme Dicer [128]. The pre-miRNA consists of two miRNA strands complementary to one-another, where one is called 3-p and the other one

5-p. The strand most abundant (which can be either 3-p or 5-p) is called the mature strand and is transferred into an Ago protein, whereas the other strand (determined the star (\*)-strand) is usually degraded [129]. The miRNA strand selected, similar to single stranded RNA (siRNA) is defined by the base pair stabilities at the 5' end of the duplex [130]. The strand selection and RNA-induced silencing complex (RISC) assembly is accomplished by a complex that contains Dicer, Ago and Transcription binding protein (TRBP) [128]. Not all precursor miRNA are processed to mature miRNA and the explanation for this still remains unclear. Most likely this event is an accumulation of several processes like reduced nuclear transport, processing by Drosha and Dicer or other specific factors which might promote or interfere with transport or processing [131, 132].

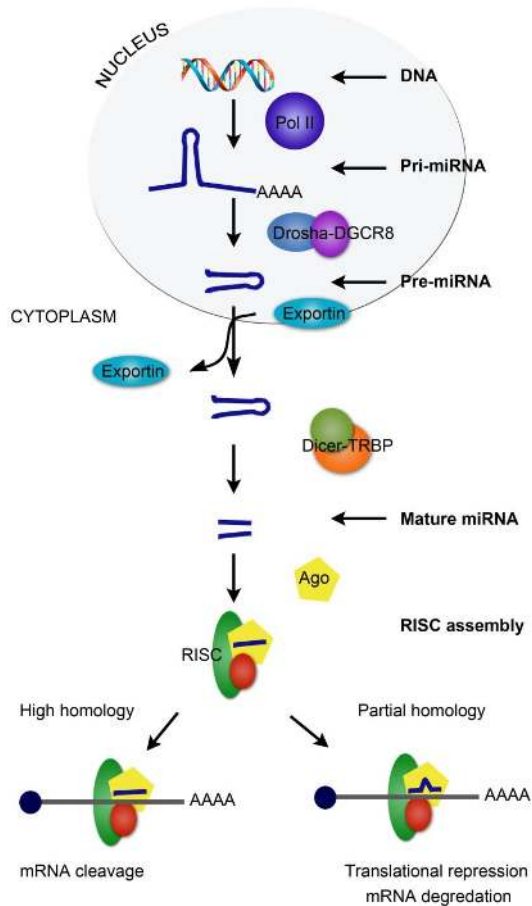


Figure 9. MiRNA biogenesis.

MiRNAs can bind to mRNA with partial complementarity, therefore one miRNA may downregulate several mRNAs making the study of downstream applications challenging. Nucleotides 2-8 of the miRNA are particularly important for pairing with the target mRNA. This

motif is referred to as the miRNA seed sequence [129]. Tools for predicting mRNA targets use different lengths of this seed sequence in their predictions. Most of the miRNA target sites reside in the 3'UTR of mRNA but functional binding sites in the 5'UTR as well as in the open reading frame have also been reported [133, 134]. Depending on the recognition site, the binding of miRNA complex to mRNA can have two different outcomes; target silencing or target degradation [135]. MiRNAs may degrade the mRNA using different mechanisms. One is by the miRISC which cleaves and degrades the mRNA transcript, another is removal of the mRNA poly-A tail by a deadenylase complex such as CCR4-NOT [136-138]. This tail shortening leads to the removal of the mRNA 5' cap. Uncapped mRNAs are rapidly removed from the cell by 5' to 3' exoribonucleases such as Xrn1 [139]. It looks as if target degradation is the predominant mode of regulation by miRNAs in mammalian cell cultures [140, 141]. Mechanistic details of miRNA-mediated repression are starting to emerge, but a comprehensive picture of the inhibition, and particularly the effects on mRNA translation, is still lacking.

## ***b) MiRNA and cancer***

MiRNA genes are frequently located at fragile sites, as well as in minimal regions of loss of heterozygosity, minimal regions of amplification or common breakpoint regions causing deletions and mutations leading to dysregulation of miRNAs [142, 143]. The first miRNA discovered in humans was let-7 [144]. This miRNA has tumor suppressor activity and among its targets is the oncogene *RAS*. Initially, changes in miRNA expression levels in cancer was discovered in B-cell chronic lymphocytic leukemia (CLL) and since then large profiling studies



have been published showing that miRNA have different profiles in cancer cells compared to normal cells in various tumor types [145-147]. Alterations seen in cancer cells that express miRNAs consist of both overexpressed and downregulated miRNAs, and it is postulated that over 60% of all genes may be regulated by miRNAs [148].

MiRNA expression patterns have also been shown to be associated with diagnosis and prognosis in several tumor types [149-154]. An interesting study showed that miRNA expression levels could accurately identify cancer tissue origin from metastatic tumors based on a miRNA expression library consisting of 48 miRNAs from 22 different cancer types thereby implicating that miRNAs are tissue specific [155].

### ***c) MiRNA in colorectal cancer***

Since the discovery that miRNA play a role in human pathogenesis, several groups have used a variety of methods to show that miRNA expression patterns are indeed altered in CRC compared to normal tissue [146, 147, 156-158] (Table 3).

The first miRNAs discovered in CRC were miR-143 and miR-145 which showed reduced expression levels in tumors compared to normal tissue [159], and these miRNAs were later found to have tumor suppressor activity in CRC [157, 160, 161]. Spectrums of dysregulated miRNAs have since then been identified to be associated with CRC genesis, progression and therapeutic response [162-164]. Nevertheless there are some discrepancies between studies that may be explained by different factors like tumor location or genetic background including

different MSI and CIMP tumors [161, 165]. MiRNAs and their potential value as biomarkers in CRC early detection, prognosis and therapy will be discussed below.

Author (year)	Sample size and type	miRNA	Tumor levels	Quantification technique	Clinical outcome	Ref
Michael et al.	16 CRC and matched normals (FFPE)	miR-143, miR-145	Down	Northern blot analysis	ND	[159]
Volinia et al.	46 CRCs and eight normals (FFPE)	miR-9-3	Down	miRNA microarray	ND	[147]
		miR-21, miR-17-5p, miR-191, miR-29b-2, miR-223, miR-128b, miR-24-1, miR-24-2, miR-155, miR-20a, miR-107, miR-32, miR-30c, miR-221, miR-106a,	Up	miRNA microarray	ND	
Bandres et al.	12 CRCs and matched normals (FFPE)	miR-133, miR-145, miR-129, miR-124a, miR-30-3p, miR-328	Down	Real-Time qRT-PCR	ND	[157]
		miR-19a, miR-20, miR-21, miR-183, miR-96, miR-31, miR-135b	Up	Real-Time qRT-PCR	miR-31 stage IV> stage II	
Slaby et al.	29 CRCs and six normals (FFPE)	miR-143, miR-145	Down	Real-Time qRT-PCR	miR-143, miR-145 tumors > 50mm	[166]
		miR-21, miR-31	Up	Real-Time qRT-PCR	miR-21 node positivity and distant metastasis	
Schepeler et al.	49 CRCs (stage II) and 10 normals (FFPE)	miR-145, miR-320, miR-498	Down	Real-Time qRT-PCR and miRNA <i>in situ</i> hybridization	High miR-320 and miR-498 associated with recurrence free survival	[161]
Diaz et al.	110 CRCs and matched normals (frozen)	miR-17-5p, miR-106a	Down	Real-Time qRT-PCR	miR-106a associated with short disease-free survival	[167]
Motoyama et al.	69 CRCs and matched normals (frozen)	miR-143, miR-145	Down	miRNA microarray	ND	[168]
		miR-31, miR-183, miR-17-5p, miR-18a, miR-20a, miR-92	Up	miRNA microarray	miR-18a associated with poor clinical prognosis	
Xi et al.	24 CRCs and matched normals (frozen)	miR-15b, miR181b, miR-191, miR200c	Up	Real-Time qRT-PCR	miR-200c short median survival	[169]
Lanza et al.	39 CRCs (23 MSS and 16 MSI-H) (FFPE)	miR-17-92 family	Up	MIRNA microarray and northern blot analysis	miR-17-92 MSS MSI-H	[170]
Nakajima et al.	46 CRCs treated with 5-FU-based antimetabolite S-1 (FFPE)	Let-7g, miR181b, miR-200c	Up	Real-Time qRT-PCR	Let-7g, miR-181b associated with therapy response but not with survival	[171]
Monzo et al.	22 CRCs (six stage I and 16 stage II) and matched normals (FFPE)	miR-17-92 family	Up	Real-Time qRT-PCR and miRNA <i>in situ</i> hybridization	miR-17-92 family associated with CRC carcinogenesis	[162]
Schetter et al.	197 colon cancers (84 test and 113 validation) and matched normals (FFPE)	miR-20a, miR-21, miR-106a, miR-181b, miR-203	Up	miRNA microarray and <i>in situ</i> hybridization	miR-21 associated with poor survival in both cohorts	[158]
Svoboda et al.	35 rectal tumors before and after treatment with capecitabine (frozen)	miR-10a, miR-21, miR-31, miR-125b, miR-137, miR-145, miR-212, miR-339, miR-361	Up	Real-Time qRT-PCR	miR-125b and miR-137 associated with worse response to therapy	[172]
Ng et al.	95 plasma (five test and 90 validation) and 50 healthy controls	miR-17-3p, miR-92	Up	miRNA microarray	miR-92 differentiates CRCs from healthy subjects	[173]
Yamamichi et al.	73 tumors (includes 34 polyps and 39 CRCs) (FFPE)	miR-21	Up	Locked nucleic acid <i>in situ</i> hybridization	Precancerous adenomas and CRCs but not in nontumorigenic polyps	[174]
Huang et al.	157 plasma (120 CRCs a 37 advanced adenomas) and 59 healthy controls	miR-29a, miR-92a	Up	Real-time qRT-PCR	miR-29a and miR-92a differentiates adenomas from CRCs and healthy controls	[175]
Schee et al.	193 CRCs (frozen)	miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145	-	qRT-PCR	High expression of miR-31 associated with increased pT stage	(submitted)
Akcakaya et al.	50 CRCs	miR-133b, and miR-185	-	MIRNA microarray platform and qRT-PCR	High miR-185 and low miR-133b associated with poor OS	[176]
Sanver et al.	80 CRCs and 26 normals (frozen)	miR-31, miR-135, miR-96, miR-182, miR-182*, miR-183	Up	miRNA BeadArray TM and qRT-PCR	miR-31 increase was significantly greater in dMMR than in pMMR tumors	[177]
Shibuya et al.	156 CRC and matched normals (frozen)	miR-21and mir-155	Up	qRT-PCR	miR-21 associated with venous invasion, liver metastasis, tumor stage and OS. miR-155 was associated with OS	[178]

**Table 3. Differentially expressed miRNAs in CRC and their potential clinical relevance. 5-FU: 5-Fluorouracil, FFPE: Formalin-fixed paraffin embedded, MSI-H Microsatellite instable-high, MSS: Microsatellite stable, ND: Not determined, qRT-PCR: Quantitative reverse transcription-PCR, pT-stage: tumor invasion stage at diagnosis, dMMR: defective Mismatch Repair, pMMR: proficient Mismatch repair. Table modified from Manne et al. [179]**

## ***d) MiRNA as early detection biomarkers in serum, plasma and feces***

Although therapeutic strategies have greatly improved over the last years, surgery is still the only curative form of treatment in CRC. As discussed above, early detection improves survival and therefore, the search for early detection biomarkers is highly warranted. In addition to being found in tissue, miRNAs have also stably been detected in feces, serum, plasma and urine making them potential good biomarkers [180-185]. Therefore, researchers have been looking at miRNAs from either feces or blood as a possible minimal invasive strategy to improve early detection methods for CRC.

Mutations in the *APC* gene, leading to reduced expression, are defined as an early event in CRC. MiR-135a and miR-135b are both overexpressed in adenomas and adenocarcinomas and one of their target is APC, implying that the up regulation of miR-135 is an early event in CRC and might be used as an early detection biomarker [186].

Ng and colleagues examined miRNA in plasma from 25 CRC and 20 healthy controls and found that higher levels of miR-17-3p and miR-92 were detected in CRC patients compared to healthy controls [173]. By measuring miR-92 in plasma it was possible to distinguish CRC from healthy controls with 70% specificity and 89% sensitivity. This was the first evidence that miRNA measurement could be used as a minimally invasive test for detecting CRC. The presence of miR-92a together with miR-29 was in a different cohort proven to be able to distinguish CRC patients from healthy individuals with 83.0% sensitivity and 84.7% specificity [175].

MiRNAs have been found to be conserved in feces, most likely shed from tumors in the colon. The levels of miR-21, miR-92a, miR-106a, miR-144\* and miR-34b/c have been found to be elevated in fecal samples, indicating that miRNA can also be collected from this kind of material [183, 184, 187, 188] and might be an indicator of CRC. Additional studies will have to reexamine these findings and determine the specificity and sensitivity of such assays.

It was long speculated how miRNAs could be conserved in environments such as blood and feces. Microvesicles, or exosomes, are extracellular membrane-enclosed vesicles that have been shown to contain miRNA. They provide a protective membrane that in the harsh environments increases the stability of its contents [189, 190]. MiRNA are in this manner protected from RNases and can mediate communication between cells [191]. Explanations on the role of exosomes remains limited but the accumulating evidence suggests that they might be applied in cancer therapeutics in the future as a delivery system for miRNAs [192].

### ***e) MiRNA as prognostic biomarkers***

The search for prognostic biomarkers in CRC has in recent years been extended to evaluating the potential value of analyzing expression of single or multiple miRNAs in tumor tissue. There are several interesting candidates, but the varying results make it difficult to draw definitive conclusions.

Several studies have utilized large scale miRNA panels in the search for possible prognostic biomarkers. Some of these studies have found similar results while others have not, and the

reason for these discrepancies remains elusive. One example is the overexpression of miR-21 which has been associated with poor overall survival (OS) in three separate studies [158, 166, 178]. MiR-21 is among the most highly expressed miRNAs in CRC and is considered to be an oncogene [193-195]. One of the proposed mechanisms by which miR-21 might promote tumor invasion and metastasis is through downregulation of its target Pcd4, a 64kDa protein that inhibits tumor progression by interacting with translation initiation factors eIF4A and eIF4G [196-198]. At first glance, miR-21 shows great potential as a prognostic biomarker in CRC, but there are issues to address. First, several exploratory studies performed on CRC cohorts failed to identify miR-21 as a prognostic biomarker, but instead other miRNAs, such as miR-17, miR-18a, miR-133b, miR-185, miR-22, miR-200c, miR-141, miR-221 and miR-150 were associated with overall survival [168, 169, 176, 199-201] [202, 203]. Secondly, miR-21 is not tissue specific, as it is found to be overexpressed in several cancers including breast, lung and kidney [204-206].

Although overall survival is an important endpoint it is not the only useful endpoint in the search for biomarkers. Among clinicopathological features, increased expression of miR-31 has been associated with increased TNM- and pT stage [157, 160][Schee et al. submitted] and low expression of miR-106a was found to correlate with advanced disease stage [167]. MiR-21 overexpression has been associated with lymph node metastasis and distant metastasis [166]. MiRNAs have also been found to be differentially expressed in microsatellite stable (MSS) groups compared to microsatellite unstable (MSI) in addition to subgroups (MSI-high and low) in CRC [161, 207]. As stated earlier, MSI tumors have a favorable prognosis compared to MSS tumors and the identification of these patients might influence treatment decisions. Members of the 17-92 cluster were upregulated in CRCs with MSS compared to MSI-H, while miR-92, miR-

223, miR-155, miR-196a, miR-31 and miR-26b were differentially expressed in MSI subgroups [170, 208]. In another study, a different set of miRNA expression profiles, consisting of three miRNAs (miR-622, miR-362-5p and miR-486-5p) could classify the tumor based on MSI subgroups.

There is no question that miRNA are differentially expressed between tumor tissue and normal tissue, the challenge lies in the interpretation of miRNA expression data, particularly when trying to use the data to explain clinical tumor progression. There is no clear indication why these results vary and further validation is needed to determine their prognostic value and whether or not they can be incorporated into routine oncological practice.

### ***f) MiRNA as predictive biomarkers***

A major concern in the management of CRC patients is prediction of therapeutic efficacy. Chemotherapy is becoming an increasingly important treatment modality, but drugs have significant side effects and treatment is costly, making biomarkers an important tool in determining individualized treatment.

MiRNA expression patterns have been shown to change when tumor cells are exposed to chemotherapeutic agents. For example, in sub clones of CRC cell lines HT29 and HT116, 19 miRNAs were upregulated and 3 downregulated when exposed to 5-FU compared to untreated cells [209]. However, as expected, not all changes in miRNAs expression were consistent with the anti-tumor effects of the drug. For example, miR-21 remained upregulated after treatment.

On the other hand, miR-200b, which is normally upregulated in tumors, was downregulated after 5-FU treatment. MiR-200b suppresses a gene that codes for a protein tyrosine phosphatase, PTPN12 that inactivates products of oncogenes such as c-ABL, Src or RAS resulting in decreased proliferation. Thus, although miR-21 remained upregulated after 5-FU treatment, downregulation of miR-200b may have contributed to the anti-tumor effects of the drug.

Anti-EGFR treatment has been shown to be less efficacious when *KRAS* mutations are present. There are, however, patients harboring *KRAS* mutations that still benefit from anti-EGFR treatment, but the underlying mechanisms have not been determined. One of the let-7a targets is *KRAS* and patients with *KRAS* mutations and high let-7a expression have shown increased survival benefit from anti-EGFR therapy [210, 211]. Let-7a may therefore be a predictive biomarker for subgroups of patients receiving anti-EGFR therapy. MiR-143 has among its targets DNA methyl transferase 3A (DNMT3A), *KRAS* and extracellular signal-regulated kinase-5 (ERK5) [212-214]. As would be expected, the loss of miR-143 resulted in increased *KRAS* expression, while overexpression of miR-143 caused reduction of *KRAS* protein levels, thereby reducing cell proliferation, suggesting that loss of miR-143 expression is a critical step in CRC progression. Overexpression of miR-143 has also shown to impair CRC growth in xenografts with induction of apoptosis [215].

Several individual miRNAs have also been suggested to be of importance for drug sensitivity in CRC. MiR-192 has been shown to downregulate the anticancer target dihydrofolate reductase (DHFR) that participates in a cycle feedback loop together with thymidylate synthase (TS) and is



involved in DNA and RNA synthesis [216]. One of the mechanisms of 5-FU is that it blocks TS, and therefore miR-192 might indirectly sensitize cells for 5-FU treatment. Other miRNAs that may sensitize cells for 5-FU treatment are miR-140 and miR-143. MiR-143 is downregulated in CRC, and introduction of miR-143 in HCT116 cells resulted in reduced cell viability and increased apoptotic cell death upon 5-FU exposure [217]. MiR-140 has been shown to be upregulated in osteosarcoma xenografts, HCT 116 cells and stem-like populations of HCT116, and all were shown to be resistant to 5-FU treatment [218]. Therefore, it has been proposed that transfection of anti-miR-140 might be used to sensitize cells for 5-FU treatment. The 5-FU analog, S-1, is used as a component of fluoropyrimidine-based adjuvant chemotherapy. Studies on RNA isolated from formalin fixed paraffin embedded samples, from 46 patients that were treated with S-1 showed that reduced expression of let-7g and miR-181b was associated with favorable response but no association with patient survival was observed [171].

Several miRNA have emerged as predictive biomarker candidates, but the use of miRNA as predictive biomarkers is still in its infancy and verification of targets in clinical panels is necessary to assess the potential value of the above mentioned candidates.

### ***g) MiRNA as therapeutic targets***

One miRNA may target several mRNAs, thereby influencing a cascade of biological pathways. This makes miRNA interesting potential therapeutic targets, but herein also lies the challenge in determining the range of possible side effects. Although no miRNAs are used as therapeutic targets in CRC today, there are several promising candidates. MiRNA used in treatment would

either have to restore miRNA expression or silence overexpressed miRNAs. Restoring miRNA expression may be done by introducing miRNA in form of a miRNA oligo mimic (which comprises of the same nucleotide sequence as an endogenous miRNA, and is designed to target the same mRNAs as that miRNA), introducing the miRNA gene by a DNA vector or reverse epigenetically silenced miRNA [219]. Inhibition of miRNAs is performed with antagomirs (anti-miRs), which are oligonucleotides complementary to the target miRNA and has been shown to inhibit miRNA for several days [220]. The challenge is devising a specific, effective and stable delivery system, and of course in finding a miRNA with the desired downstream effects on the biological system. Advances have been made using small interfering RNA (siRNA) for cancer therapy through local or targeted administration in vivo, and the similarity between the miRNA and siRNA systems in silencing mRNA makes the findings applicable to both [221].

To summarize, several potential therapeutic miRNA candidates have been described in CRC, as mentioned in the preceding chapters. For instance, introduction of miR-143 might e.g. inhibit tumor proliferation, lowering the expression of miR-21 might counteract tumor progression and metastasis while lowering let-7g and miR-181b could improve response to S-1 treatment. There is still a lot to learn about miRNA targets and the downstream effects they have in CRC, but future experiments will continue to improve our understanding of this disease and hopefully the end result will provide novel therapeutic candidates to aid in the treatment of CRC.

# SUMMARY OF PAPERS

## *Paper I*

### **Investigation of nonspecific cross-reacting antigen 2 as a prognostic biomarker in bone marrow plasma from colorectal cancer patients.**

Schee K, Flatmark K, Holm R, Boye K, Paus E.

CEA is one of the most widely used biomarkers in the clinic today and is mainly used to monitor recurrence in CRC patients. Although CEA has been used for the past 30 years the clinical relevance is still debated, and its use as a biomarker is limited by poor sensitivity and specificity. Hence, the search for new biomarkers is warranted. NCA-2 is a truncated version of CEA, transcribed from the same gene and has been proposed as a biomarker. However, its relevance has not been investigated in a large cohort of CRC patients, and previous studies on NCA-2 have focused on its cross-reactivity with CEA. Examining 277 bone marrow plasma samples from CRC patients by immunofluorometric assay, we found that both CEA and NCA-2 were associated with advanced tumor stage at diagnosis and adverse patient outcome. Expression of CEA above 5µg/l exhibited, as expected, associations with poor outcome. NCA-2 was expressed at much lower concentrations than CEA, but apart from that NCA-2 showed the same trends as CEA, but no additional prognostic information was gained. Specific NCA-2 expression was observed in the immunofluorometric assay but when exposed to common fixation chemicals, cross reactivity with CEA was observed. This is most likely attributable to the presence of epitopes in CEA which are masked in its native form that are made accessible during conformational changes due to chemical processing by the different fixation chemicals.

The results from this study indicate that NCA-2 is probably not a useful prognostic biomarker in CRC and furthermore, underline the issue of antibody specificity when investigating CEA species molecules.

## ***Paper II***

### **Clinical relevance of miR-21, miR-31, miR92a, miR-101, miR-106a and miR-145 in colorectal cancer.**

Schee K, Boye K, Abrahamsen T.W, Fodstad Ø, Flatmark K.

MicroRNAs (miRNAs) are short non-coding RNAs (~22nt) that regulate gene expression post-transcriptionally by either inhibiting or degrading mRNA. They are chemically stable and can be detected in a broad range of clinical samples acting as oncogenes or tumor suppressors. Based on the literature we selected six miRNAs (miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145) and analyzed for associations between expression in tumor tissue and clinicopathological data and patient outcome. We performed qRT-PCR on the six selected miRNAs in our cohort consisting of 196 CRC curatively resected tumors. Our results revealed detectable and highly variable levels of five of the six miRNAs in the tumor panel. MiR-21 exhibited the highest expression while miR-101 was hardly expressed relative to the reference (RNU44). Increased expression of miR-31 was associated with increased pT stage and with poorly differentiated tumors. Both miR-92a and miR-106a were associated with differentiation, as higher median expression levels were found in intermediately differentiated tumors than in well and poorly

differentiated tumors. In addition, miR-31, miR-92a and miR-106a were associated with tumor localization as miR92a and miR-106a exhibited higher levels in rectal tumors while miR-31 had higher expression in colon tumors. Although both miR-21 and miR-106a had in previous studies showed association with overall survival, none of the selected miRNAs in our cohort showed any such associations. These results emphasize the importance of validating potential targets in individual large clinical panels, and indicate that the potential prognostic role of miRNAs in CRC is still unclear and needs further evaluation.

### ***Paper III***

#### **Deep sequencing the microRNA transcriptome in colorectal cancer.**

Schee K, Lorenz S, Molton Warren M, Günther C-C, Holden M, Hovig E, Fodstad Ø, Meza-Zepeda L, Flatmark K.

Colorectal cancer (CRC) is a leading cause of cancer related deaths in the Western world and the search for biomarkers which might aid in treatment decisions is warranted. MiRNAs are short non-coding RNAs which regulate gene expression and have previously been shown to be possible biomarkers in CRC. Deep sequencing is a new method allowing for detection of a global miRNA expression with a wide expression range. 90 CRC tumor samples were deep sequenced and associations with clinicopathological parameters and outcome were assessed. In our cohort, 523 known miRNAs were expressed. The five most highly expressed miRNAs (miR-10a-5p, miR-21-5p, miR-22-3p, miR-143-3p, and miR-192-5p) accounted for ~54% of the total amount of reads in the cohort. Pathway analysis was performed for the top five miRNA, and the predicted

targets included genes involved in cancer-related pathways, including the CRC pathway. In the CRC pathway, the identified genes included oncogenes (K-RAS), tumor suppressors (APC and TGFR $\beta$ II), and DNA repair genes (hMSH6), suggesting that the top five most highly expressed miRNAs may regulate the expression of highly cancer-relevant proteins.

When examining associations between miRNA expression and clinicopathological parameters, high expression of miR-10b was found to be associated with tumors located in the colon relative to the rectum and with poor tumor differentiation. Among the five most highly expressed miRNAs, low expression of miR-192 was associated with high pT stage and poorly differentiated tumors. No associations were found between miRNA expression and outcome. MiRNA expression, analyzed by hierarchical clustering, revealed relatively constant expression across the patient samples for most of the expressed miRNAs. Given the low variability observed between samples, it is not surprising that very few associations with clinical parameters was detected and that no miRNAs emerged as prognostic biomarker candidates.

## METHOLOGICAL CONSIDERATIONS

### *a) Clinical samples - sample size and origin*

A crucial component when planning a clinical study is defining the number of participants needed in order to reliably answer the proposed question. Including too many patients will waste time and money, while too few may not be enough to obtain the statistical power needed. In addition, the composition of the cohort should be representative of the patient group with respect to clinicopathological parameters such as tumor stage, age and localization. Our studies were based on a cohort of 316 patient samples collected from 5 different hospitals in the Oslo region in between 1998-2000. The cohort was collected with the aim of examining possible biomarkers in CRC. The patients were included at the time of primary surgery for assumed or verified colorectal cancer [222]. The samples collected were assessed by routine histopathology and additional tumor tissue was snap frozen in liquid nitrogen in the operating theatre. Thirty-one patients were excluded from further analysis for the following reasons: not invasive cancer (25), histology other than adenocarcinoma (5) and unknown stage of disease (1). In addition, sections were not obtainable in eight cases (Figure 8). Thus the cohort comprised of 277 patient samples of tumor tissue and bone marrow plasma with diverse clinicopathological parameters. As these samples were collected by the year 2000, long-term follow-up is available for the patients, which makes this a suitable cohort in the search for biomarkers. This cohort has previously been used in several studies examining tumor expression of protein markers such as S100A4, EMMPRIN and B7H3 and their association with different clinicopathological parameters and outcome in CRC [223-226].

For papers I, II and III different selections from the CRC cohort were used and the basis for this selection is depicted in Figure 10.

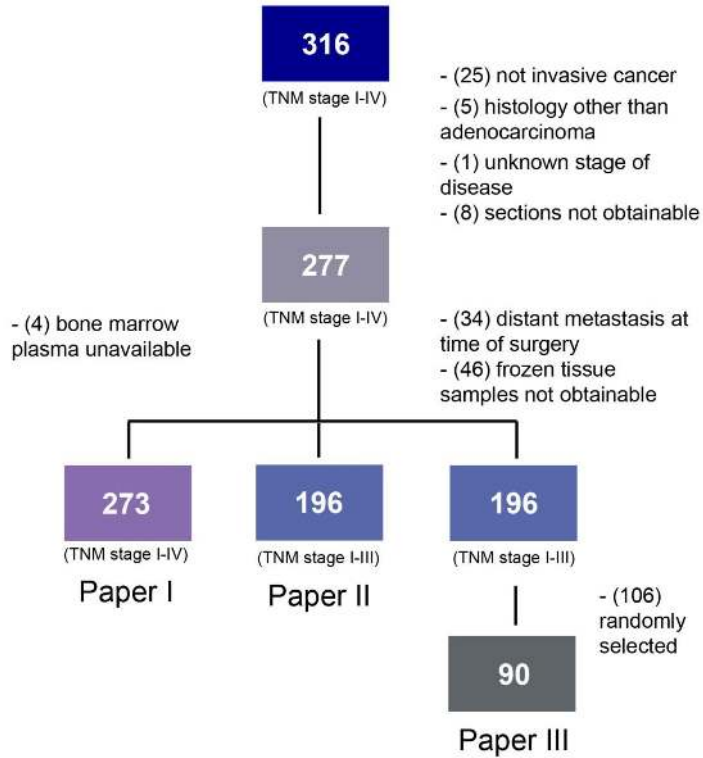
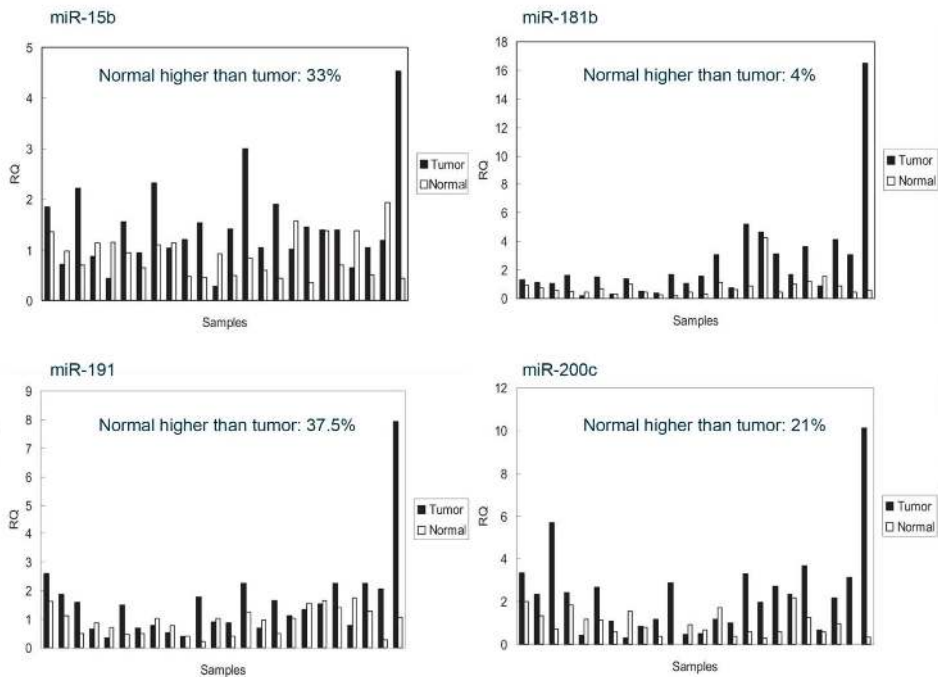


Figure 10. Patient cohorts analyzed in papers I, II and III.

The colorectum is a heterogeneous organ with many different cell types which are influenced by lifestyle and genetics, albeit, many miRNA studies utilize normal tissue taken from adjacent colon from CRC patients and use this as a normalization factor. It is difficult to assess how



miRNA expression in normal tissue might influence the miRNA expression in cancer tissue as very few studies report their normal miRNA expression profiles. A microarray study performed on 40 normal tissues revealed differentially expressed miRNAs within the tissues represented, however only one sample from colon was included making it difficult to draw any further conclusions as to the global miRNA expression pattern [227]. It can be argued that miRNA expression in colon from healthy individuals varies considerably and is therefore not ideal to normalize against. In addition, it is debatable whether colon taken adjacent from cancer can be denoted normal. Studies performed on normal colon alone to establish a baseline for miRNA expression are lacking and the few publications that show normal miRNA expression levels, relative to cancer, show great variation in expression of the individual miRNA [200, 212]. In a study performed on CRC tissue samples where miRNA expression was reported to be upregulated in tumor tissue as opposed to normal adjacent tissue, a significant proportion of normal samples actually had higher expression levels than in tumor tissue: 33% for miR-15b, 37.5% for miR-191, 4% for miR-181b and 21% for miR-200c (Figure 11) [169]. This implies that the expression varied greatly between normal tissue and tumor tissue among miRNAs. Moreover, the figure also illustrates that expression of individual miRNAs varied considerably between normal samples.



**Figure 11. miRNA expression in normal and tumor tissue adapted from Xi et al. [169].** Relative Quantities (performed with RT-PCR) of four selected miRNAs from 24 CRC tumor and adjacent normal tissues. The miRNAs have been described as overexpressed in tumor tissue relative to normal tissue. The “normal higher than tumor” describes the percentage of patients in the cohort with higher miRNA expression in normal as opposed to tumor tissue.

It would be interesting to explore what causes the differences between these matched normal/tumor samples. Some studies have instead of normalizing against matched normal tissue, pooled normal tissue from several different patient samples together. Normalizing against values derived from analyzing a pool of samples as opposed to corresponding adjacent

normal tissue will most likely result in different miRNA expression outcome. Exploring miRNA expression in the normal colorectum and determining a baseline is an important task that requires more research. Whether this may account for some of the discrepancies observed between different miRNA studies is only speculation at this time but the few examples shown do imply that normal tissue has highly variable miRNA expression patterns.

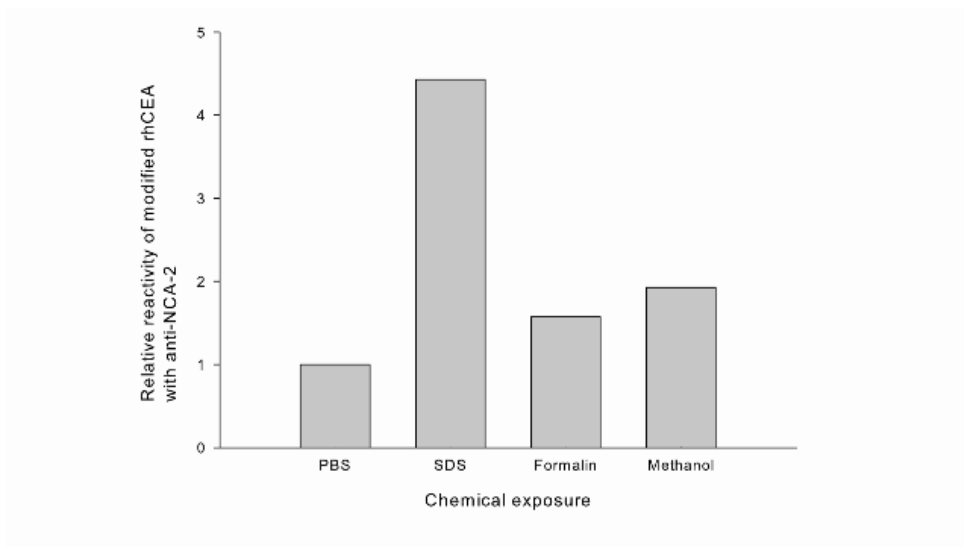
### ***b) Antibody specificity***

In the studies used to assess the expression of CEA and NCA-2 immunological detection methods were utilized, and consequently, antibody specificity is an essential issue. In paper I we used monoclonal antibodies (MAb) which consist of a homogenous pool of antibody molecules specific for a single epitope on the immunogen. The NCA-2 antibody (O-22) is not commercially available and was established at our institution [228], anti-CEA<sub>total</sub> (12-140-10) was also established in our lab, while anti-CEA<sub>spec</sub> (T84.66) was bought from ATCC. Antibodies, bought or self-established, should prior to utilization be properly examined with respect to sensitivity and specificity for the application planned. The anti-CEA<sub>total</sub> antibody has been extensively investigated in collaborative epitope screening projects [97, 229], and has been subject to studies of differential CEA antibody specificity, and cross-reactivity with other antigens like NCA, and NCA-2 has also been extensively examined [115, 230, 231]. The T84.66 antibody was shown to be specific for full-length CEA, not recognizing the shorter, C-terminally truncated versions [229]. No cross reactivity was observed between the anti-NCA-2 antibody and full-length CEA in the native form in our investigations. The proposed binding site for anti-CEA<sub>total</sub> antibody is in

the N-terminus which would explain why it recognizes both full-length CEA and NCA-2 as they are believed to share this part of the sequence. The binding site of the anti-CEA<sub>spec</sub> antibody is believed to be in the C-terminal end which is absent in NCA-2 and NAC-2 is therefore not recognized by this antibody.

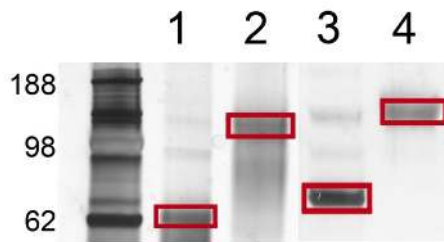
In our experiments we observed inconsistencies in NCA-2 expression when using different detection methods such as the immunofluorometric assay (IFMA), Western Immunoblot (WB) and immunofluorescence. NCA-2 was found with the same high expression as was found when staining with anti-CEA with respect to both signal intensity and distribution.

Immunohistochemical staining performed on gastric tissue in a previous study, where NCA-2 was pre-absorbed with NCA and CEA, did not exhibit equally high signal intensity and distribution and therefore this discrepancy of what was previously reported together with the discrepancies between our IFMA-and immunohistochemistry results led us to question the specificity of our NCA-2 antibody [232]. A hypothesis was that this could be caused by the presence of epitopes that were usually masked in the native form, that were made available for antibody binding when fixation chemicals conformational changes in the protein in the different detection methods [233, 234]. This theory was tested in paper I by exposing rhCEA to different chemicals and probing with anti-NCA-2 (Figure 12), which showed that the read-out of the IFMA changed depending on the fixation chemical the recombinant protein was exposed to.



**Figure 12.** Relative expression of rhCEA exposed to different fixation chemicals probed with anti-NCA-2. Figure from paper I [235].

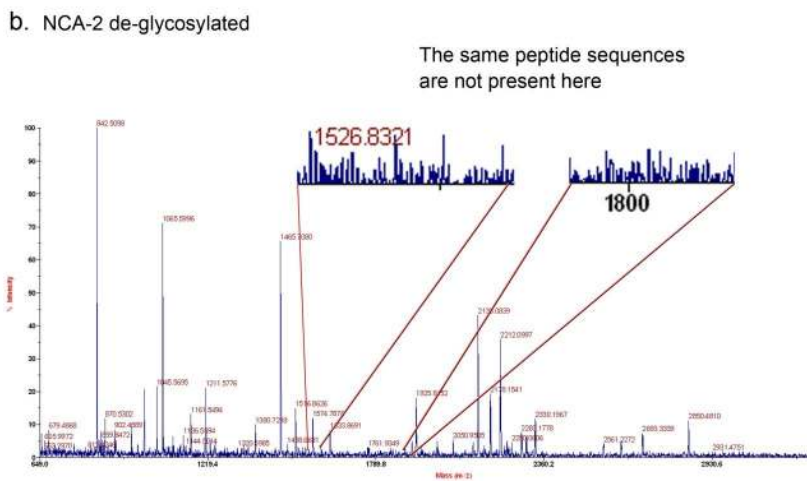
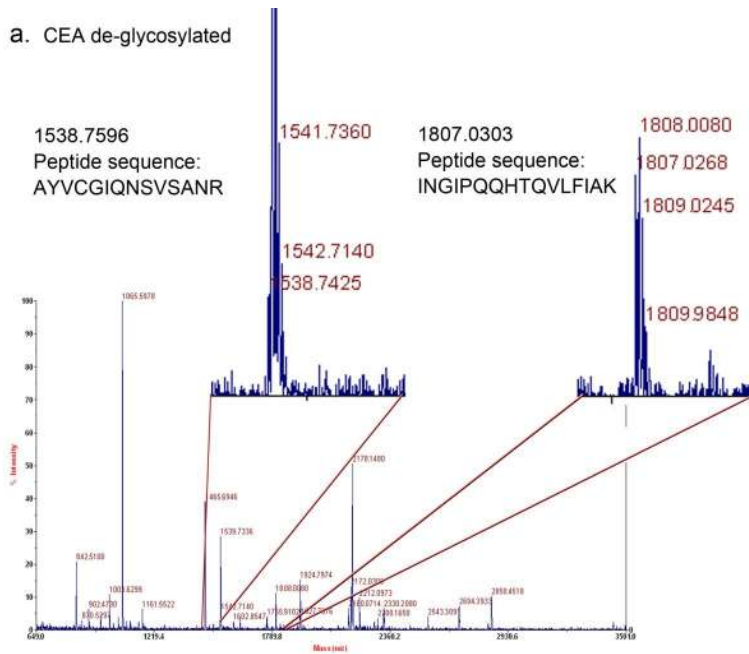
Although our results indicated that the apparent cross reactivity was related to differences in epitope availability dependent on protein conformation, questions were raised regarding the purity of the immunogen used to establish the NCA-2 antibody. To examine this, the NCA-2 immunogen (purified from human meconium) and rhCEA were separated by SDS-PAGE with and without prior deglycosylation (Glycoprofile II, enzymatic in-solution N-deglycosylation kit, Sigma-Aldrich), and silver staining was performed to visualize bands. Distinct bands were detected at appropriate levels (Figure 13) suggesting that successful deglycosylation was achieved, and confirming the slightly smaller size of NCA-2.



**Figure 13.** Silver staining gel of; 1: NCA-2 de-glycosylated, 2: NCA-2 native, 3: CEA de-glycosylated, 4: CEA native The red squares indicate which protein bands were excised from the gel for mass spectrometry analysis.

The bands were excised, in-gel digestion with trypsin was performed and peptides were extracted and subjected to mass spectrometric (MS) analysis. Detected peptide sequences were matched against the theoretical trypsination spectra of CEA. In the rhCEA sample, 10 peptides were detected that unequivocally identified CEA. Two of these peptides were observed as distinct peaks, and these were located in the C-terminal end of the CEA molecule (that according to theory should be absent from NCA-2). In the NCA-2 sample, 4 peptides corresponding to the N-terminal part of CEA were detected, but the two C-terminal peptides could not be detected in this sample (Figure 14). The results were relatively similar for native and de-glycosylated samples.

In conclusion, our results show that specific detection of CEA and NCA-2 was possible in native samples, such as bone marrow plasma, whereas the specificity of the anti-NCA-2 antibody was lost when the samples were exposed to any of the tested fixation chemicals. Antibody specificity should constitute an important consideration whenever an immunologic assay is being used.



**Figure 14.** Mass spectrometry of CEA and NCA-2. **a. CEA de-glycosylated.** Two peaks were detected in CEA located in the C-terminal end. **b. NCA-2 de-glycosylated** In NCA-2 the peaks corresponding to the C-terminal end of CEA could not be detected.



### ***c) MiRNA quantitative RT-PCR and normalization***

The most commonly used method for analyzing expression of individual miRNA is quantitative reverse transcription polymerase chain reaction (qRT-PCR) which allows for sensitive, specific and reproducible quantification of nucleic acids [236]. The limited length of the miRNA and the sequence homology to pre- and pri-miRNA pose several challenges for correct reverse transcription, a challenge manufacturers have solved differently [237], with no indications that one method is preferentially better. There are two methods primarily used for reverse transcription of miRNAs today; miRNA specific and universal reverse transcription. In the specific reaction each miRNA is individually reverse transcribed with the help of a stem-loop primer designed to be complementary to the known 3' end of the miRNA. The second method uses a universal primer to reverse transcribe all miRNAs in the sample. This approach is especially useful if many miRNAs need to be analyzed from a small sample amount. The specificity and sensitivity of qPCR assays are dependent upon primer design. The binding affinity of the primers determined by the sequence and the GC content of a miRNA determines the  $T_m$  (melting point) against complementary sequences [237].

To be able to compare miRNA expression across different samples normalization is needed. This is performed by normalizing the target gene value against values obtained for a reference gene in the same sample. There are, however, no universally accepted reference genes identified for normalization of miRNA expression. Several candidates of small non-coding and small nucleolar RNAs have been suggested, such as RNU44, RNU6B, U24 and U26. In our studies, we examined RNU44 and RNU6B in a pilot study composed of 30 cases, and analyzed the results using geNORM, which is an algorithm that determines the most stable reference in a given sample

panel. Both miRNA were stably expressed in our samples, and for further analyses we chose to use RNU44 as a reference gene [238].

#### ***d) Deep sequencing and genome mapping.***

Global expression analysis of miRNAs has most often been performed using microarrays. Microarrays utilize fluorescent probes to detect miRNA expression in a sample using signal intensities. Deep sequencing, or next generation sequencing as it is also called, is a novel approach to discover new miRNAs as well as characterizing expression levels of annotated miRNAs [239]. One advantage of using deep sequencing opposed to microarrays is the ability to get absolute quantitative measurements instead of relative signal intensity [240]. There is however some disagreement on what is the optimal method [241].

MiRNAs are processed by enzymes through multiple steps from a precursor sequence into mature sequences. These mature sequences are complementary to each other and are denoted -3p and -5p. Usually one of the strands is selected for the microRNA induced silencing complex (miRISC) while the other one is degraded, but in some instances both miRNA strands are equally utilized. Although thermodynamic asymmetry of the duplex ends appears to play a role in which strand is selected, the mechanisms behind the strand selection is still unclear [242, 243]. With microarrays, the probes have to be pre-selected, and the array is size limited, while with deep sequencing it is possible to detect both strands.

The result from the sequencing is a data file which contains all the detected sequences from the analyzed samples, called reads. In order to denote which miRNAs the reads represent in the different samples, the reads have to be mapped to the genome. In the mapping process it is

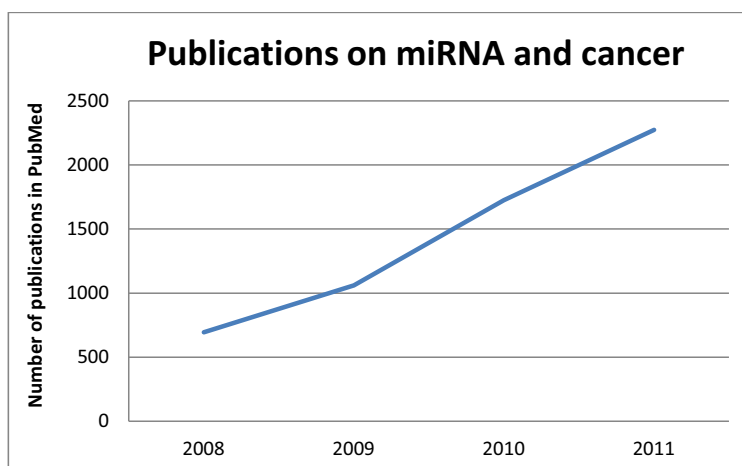
possible to set several different criteria beforehand, among them the number of mismatches tolerated, amount of overlap required and whether miRNAs are allowed to map multiple places in the genome [244]. There are both advantages and disadvantages by choosing different criteria. Not allowing for multiple mapping may result in losing miRNA expression represented by miRNA families, but allowing for multiple mapping may result in either increased or decreased miRNA expression depending on how multiple mappers are counted. MiRNAs that map multiple places in the genome can either be counted as one hit per place it maps to or it is possible to divide the expression among all the possible mapped locations. This will give different expression patterns of the expressed miRNA. As miRNAs are short sequences which may vary with only one nucleotide the different amounts of mismatches allowed when mapping may give rise to different miRNA results. In our deep sequencing study, multiple mapping was allowed where when two miRNAs shared a given number of multiple mapped reads, we identified the ratio of unique reads between these two miRNAs. This ratio was applied to divide the number of multiple mapped reads and assign them. If multiple hits were found to be perfectly mapped to one genomic region and mapped with mismatch to another one, only the perfect matches were considered.

These different settings all have an impact on the mapping results and because deep sequencing is a new method there are a multitude of challenges with finding standardized methods within genome mapping, normalization and the underlying statistic.

## RESULTS AND DISCUSSION

Bone marrow plasma from patients with CRC was utilized to measure the presence of CEA and NCA-2, and associations between plasma levels and clinicopathological parameters and outcome were analyzed, to investigate these proteins as possible prognostic biomarkers in CRC. As bone marrow plasma is not a conventional material to use, levels of CEA detected in bone marrow plasma were compared to preoperative serum measurements performed as part of the routine preoperative work-up. The levels seemed to correlate well (CEA linear  $R^2 = 0.838$ ,  $p < 0.001$ ) making this a suitable material for our study. Examining plasma samples from 273 CRC patients with immunofluorometric assay (IFMA) we found that NCA-2 behaved similar to CEA with respect to associations with clinicopathological parameters however no additional associations were found. NCA-2 concentrations were lower compared to CEA, and the cut-off level for statistical calculations was determined to be  $2\mu\text{g/l}$  (as opposed to  $5\mu\text{g/l}$  for CEA). CEA concentrations in plasma were, as expected, associated with poor overall and metastasis-free survival. Our analysis of the specificity and sensitivity for CEA with endpoint metastasis were 81% and 38%, respectively, while the calculated specificity and sensitivity for NCA-2 with endpoint metastasis was 82% and 36%, respectively, showing that these two proteins performed relatively similarly as prognostic biomarkers. This indicates that both CEA and NCA-2 are poor prognostic biomarkers as there are >60% of patients with metastases which would not be detected with this analysis. In addition, both positive predictive value (PPV) and negative predictive value (NPV) were relatively similar between the two proteins with PPV 50% for CEA and 45% for NCA-2, and NPV 75% for CEA and 74% for NCA-2 respectively.

MiRNA research on CRC has increased exponentially during the working years of this thesis (Figure 15). When this work was started in 2008, we decided to evaluate the research done on the potential prognostic relevance of miRNAs in CRC which resulted in a review article on the topic [245].



**Figure 15.** Number of publications in PubMed for miRNA and cancer in the years 2008-2011.

Based on this work, six miRNAs (miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145) were selected for their promising association with clinicopathological parameters and outcome in CRC and their expression in tumor tissue were explored with qRT-PCR in our CRC cohort (Paper II). All of the miRNAs, except miR-101, which was hardly detectable in our samples, exhibited variable expression in the tumor samples. Surprisingly, few associations were found between expression levels and age, gender, tumor stage, differentiation, localization, vascular invasion, perineural infiltration and lymphocytic infiltration.

Increased expression of miR-31 was associated with both poorer differentiated tumors and with increased pT-stage which is in concordance with previous results [157, 160, 166]. There are however discrepancies in the previous results reported on miR-31 and its associations with pT-stage and differentiation. In the study where miR-31 is associated with differentiation, no association with pT-stage was found, and in the studies where association with pT-stage was found, no association with differentiation was detected. This implicates that results may vary, and that it is important to examine such associations in different cohorts.

MiR-92a was included because of its potential as an early detection biomarker in plasma and feces and we wished to examine whether it could be associated with clinicopathological parameters or outcome [173, 183]. The expression levels of miR-92a did not increase with tumor stage nor show any other differential expression within other clinicopathological parameters as would be expected for a potential early detection biomarker. We did not detect any association with overall or metastasis-free survival, but miR-92a was associated with differentiation where median expression was higher in intermediately differentiated tumors compared to well or poorly differentiated tumors.

MiR-21 is a highly studied oncogenic factor that targets several tumor suppressor genes in various model systems and has been found to be overexpressed in at least 18 malignancies indicating that altered expression of miR-21 is a common mechanism in carcinogenesis [193, 204-206]. MiRNA-21 is the most studied miRNA in CRC and was included in our study for its proposed association with advanced tumor stage and survival [158, 166]. In our cohort miR-21

showed both high and varied expression, but we did not find any association with clinicopathological parameters or outcome. As with miR-31, there are discrepancies between studies on miR-21 expression among the ones that have found an association with outcome and the ones that have not [168, 169, 176, 199, 200]. What may cause the discrepancies is difficult to say whether they are caused by biological differences in the cohorts, sample size, normalization methods or other factors is only speculation at this time.

Downregulation of miR-145 has been identified as an early event in CRC, which might explain why it was not associated with any clinicopathological parameters in our advanced tumor panel [91, 246]. Overexpression of miR-145 has shown to suppress cancer growth and inhibit metastasis [247, 248] thereby implicating as a possible therapeutic target [161, 194, 249].

MiR-106a has previously been shown to be upregulated, relative to normal colon, and has in addition to tissue been found in fecal samples and therefore proposed as a possible early detection biomarker [91, 184, 250]. This miRNA has also been associated with survival although we did not find any such associations [158, 167]. MiR-106a, like miR-92a, was associated with differentiation where median expression was higher in intermediately differentiated tumors compared to well or poorly differentiated tumors.

Alternatively to searching for possible prognostic biomarkers among single miRNAs, a global miRNA expression search was undertaken. 90 samples from our CRC cohort were analyzed by deep sequencing, which revealed the global expression pattern of all known miRNAs annotated in miRBase to date (release 18, November 2011). A total of 523 mature miRNAs were detected

with expressions ranging from less than 10 reads up to ~500.000 reads in a single patient, showing the broad detection range of the deep sequencing technology.

The five most highly expressed miRNAs (miR-10a-5p, miR-21-5p, miR-22-3p, miR-143-3p and miR-192-5p) represented ~54% of the total miRNA counts in all the patient samples, indicating that some miRNAs are more abundantly expressed in the tumor compared to the rest. In a study on peripheral blood the expression of the let-7 family contributed to 77% of the global expression thereby supporting that few individual miRNAs account for a large part of the global expression [251]. It is difficult to say whether the abundance of a few miRNA is more important than the absence of others, as the biological pathways that these miRNAs affect might be equally important. Pathway analysis of the target genes from the five highly expressed miRNAs revealed targets relevant in cancer pathways, including colorectal cancer pathway. Targets within the CRC pathway included K-RAS, APC, TGFR $\beta$ II, and DNA repair genes which are involved in important signaling networks. Further evaluation of the proposed target genes could elucidate the implication of these highly expressed miRNAs.

Increased expression of miR-10b-5p was associated with tumors located in the colon relative to rectum and poorly differentiated tumors. MiR-10b-5p has previously been associated with increased pT stage in CRC [252]. Examining the top five highly expressed miRNAs for association with clinicopathological parameters revealed that reduced expression of miR-192-5p was associated with increased pT stage and poorly differentiated tumors. Reduced expression of miR-192 has previously been associated with MSI status in CRC while overexpression of miR-192 exhibited reduced cell proliferation by targeting cell cycle progression [207, 208, 253, 254].



The expression patterns represented by hierarchical clustering reveal that the miRNAs were relatively uniformly expressed across all the patient samples thereby supporting the lack of associations between miRNA expression and clinicopathological parameters or outcome. Although, LASSO and Cox regression analysis, with either endpoint metastasis or survival, revealed six miRNAs, adjusting for multiple testing none of these remained significantly associated with outcome.

Examining expression on both single miRNA candidates and global expression in CRC patients, and looking for associations with clinicopathological parameters or outcome revealed few associations. Our results, together with the varying results from previous studies on miRNAs as possible prognostic biomarkers in CRC, imply that there does not appear to be a definite miRNA prognostic biomarker candidate. However, the miRNA target prediction from our deep sequencing study, together with increasing evidence of validated targets with relevance in CRC, suggest that miRNA may be used as therapeutic targets in the future.

## CONCLUDING REMARKS AND FUTURE PROSPECTS

In this work, the prognostic biomarker potential of CEA and NCA-2 was investigated in plasma samples from 273 CRC patients. High levels of CEA and NCA-2 were associated with clinicopathological parameters and outcome, but although NCA-2 exhibited interesting prognostic potential, it was not superior to CEA which is used in the clinic today. Furthermore, important issues regarding the specificity of the NCA-2 antibody were explored, the results clearly underlining the importance of verifying antibody specificity for each experimental setting. The development of anti-NCA-2 variants which do not cross react with CEA during denaturing conditions would be necessary in order to further explore the biological importance of NCA-2.

Increasing evidence suggests that miRNAs play important roles in colorectal carcinogenesis, cancer progression and metastasis development, and these molecules are being explored as potential diagnostic, prognostic, and predictive biomarkers, as well as therapeutic targets [255-260]. In the present work, miRNA expression in primary tumors from patients with CRC was explored to examine their potential as prognostic biomarkers in this disease. The exploration of miRNA expression identified several miRNA that were of relevance in CRC but few associations were found with clinicopathological parameters and outcome. Taken together with the variable results reported in previous studies, no specific miRNAs can at present be identified as prognostic biomarker candidates.

Research technologies and approaches are still being established and perfected, and tools which were unavailable when this thesis was started are available today. MiRNA isolation, amplification and detection have become easier and faster with improved isolation kits and probes. In addition, complete qRT-PCR arrays with cancer pathways are easily accessible and can be run and compared to our miRNA expression levels to look for associations with known pathways within cancer.

Our deep sequencing results were based on mapping to known miRNAs, but we do have the opportunity to further examine these results and look for unknown miRNAs where new and possible important biomarkers might be found. As the technology becomes less expensive, deep sequencing will most likely be used to explore miRNA expression patterns in a similar fashion as microarrays are used today. An advantage of deep sequencing is the ability to re-analyze data at a later date when new versions of databases are released. In addition, these data may be explored in a multitude of different ways and promising single miRNAs candidates can be further examined to look for possible targets and cancer relevance. Predicted targets among the highly expressed miRNAs from our deep sequencing study included K-RAS which, when mutated, has been shown to influence EGFR therapy. It would be interesting to explore the K-RAS expression for our patient cohort and compare this to the miRNA expression levels from the deep sequencing.

We believe that the work presented in this thesis has shown the importance of validating possible findings in independent large cohorts and also laid the groundwork for future research.

# REFERENCES

1. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell 2011. **144**(5): p. 646-74.
2. Migliore, L., et al., *Genetics, cytogenetics, and epigenetics of colorectal cancer*. J Biomed Biotechnol. **2011**: p. 792362.
3. Giovannucci, E., *Diet, body weight, and colorectal cancer: a summary of the epidemiologic evidence*. J Womens Health (Larchmt), 2003. **12**(2): p. 173-82.
4. Millen, A.E., et al., *Fruit and vegetable intake and prevalence of colorectal adenoma in a cancer screening trial*. Am J Clin Nutr, 2007. **86**(6): p. 1754-64.
5. Muto, T., H.J. Bussey, and B.C. Morson, *The evolution of cancer of the colon and rectum*. Cancer, 1975. **36**(6): p. 2251-70.
6. Vogelstein, B., et al., *Genetic alterations during colorectal-tumor development*. N Engl J Med, 1988. **319**(9): p. 525-32.
7. Markowitz, S.D. and M.M. Bertagnolli, *Molecular origins of cancer: Molecular basis of colorectal cancer*. N Engl J Med, 2009. **361**(25): p. 2449-60.
8. Fearon, E.R. and B. Vogelstein, *A genetic model for colorectal tumorigenesis*. Cell, 1990. **61**(5): p. 759-67.
9. Risio, M., *The natural history of adenomas*. Best Pract Res Clin Gastroenterol 2010. **24**(3): p. 271-80.
10. Leslie, A., et al., *The colorectal adenoma-carcinoma sequence*. Br J Surg, 2002. **89**(7): p. 845-60.
11. Lin, O.S., *Acquired risk factors for colorectal cancer*. Methods Mol Biol, 2009. **472**: p. 361-72.
12. Cancer-Registry-of-Norway. <http://www.kreftregisteret.no/en/The-Registries/Cancer-Statistics/>. 2012 [cited].
13. Jass JR and Morson BC. *Reporting colorectal cancer*. J Clin Pathol 1987. **40**(9):1016-23.
14. O'Connell, J.B., M.A. Maggard, and C.Y. Ko, *Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging*. J Natl Cancer Inst, 2004. **96**(19): p. 1420-5.
15. Weiser, M.R., et al., *Individualized prediction of colon cancer recurrence using a nomogram*. J Clin Oncol, 2008. **26**(3): p. 380-5.
16. Nitsche, U., 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**(5): p. 793-800; discussion 800-1.
17. Paget, S., *The distribution of secondary growths in cancer of the breast. 1889*. Cancer Metastasis Rev, 1989. **8**(2): p. 98-101.
18. *Neoplastic Diseases: A Treatise on Tumours*. By James Ewing, A.M., M.D., Sc.D., Professor of Pathology at Cornell University Medical College, N.Y.; Pathologist to the Memorial Hospital. Third edition. Royal 8vo. Pp. 1127, with 546 illustrations. 1928. Philadelphia and London: W. B. Saunders Co. Ltd. 63s. net. British Journal of Surgery, 1928. **16**(61): p. 174-175.
19. Sugarbaker, E.V., *Cancer metastasis: a product of tumor-host interactions*. Curr Probl Cancer, 1979. **3**(7): p. 1-59.
20. Mendoza, M. and C. Khanna, *Revisiting the seed and soil in cancer metastasis*. Int J Biochem Cell Biol, 2009. **41**(7): p. 1452-62.
21. Fidler, I.J., *The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited*. Nat Rev Cancer, 2003. **3**(6): p. 453-8.

22. Habr-Gama, A., et al., *Complete clinical response after neoadjuvant chemoradiation therapy for distal rectal cancer: characterization of clinical and endoscopic findings for standardization*. Dis Colon Rectum 2010. **53**(12): p. 1692-8.
23. Folkvord, S., et al., *Prediction of response to preoperative chemoradiotherapy in rectal cancer by multiplex kinase activity profiling*. Int J Radiat Oncol Biol Phys 2010. **78**(2): p. 555-62.
24. Midgley, R.S., Y. Yanagisawa, and D.J. Kerr, *Evolution of nonsurgical therapy for colorectal cancer*. Nat Clin Pract Gastroenterol Hepatol, 2009. **6**(2): p. 108-20.
25. *Efficacy of adjuvant fluorouracil and folinic acid in colon cancer. International Multicentre Pooled Analysis of Colon Cancer Trials (IMPACT) investigators*. Lancet, 1995. **345**(8955): p. 939-44.
26. Wolmark, N., et al., *Clinical trial to assess the relative efficacy of fluorouracil and leucovorin, fluorouracil and levamisole, and fluorouracil, leucovorin, and levamisole in patients with Dukes' B and C carcinoma of the colon: results from National Surgical Adjuvant Breast and Bowel Project C-04*. J Clin Oncol, 1999. **17**(11): p. 3553-9.
27. Andre, T., et al., *Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial*. J Clin Oncol, 2009. **27**(19): p. 3109-16.
28. Francini, G., et al., *Folinic acid and 5-fluorouracil as adjuvant chemotherapy in colon cancer*. Gastroenterology, 1994. **106**(4): p. 899-906.
29. O'Connell, M.J., et al., *Controlled trial of fluorouracil and low-dose leucovorin given for 6 months as postoperative adjuvant therapy for colon cancer*. J Clin Oncol, 1997. **15**(1): p. 246-50.
30. Meyerhardt, J.A. and R.J. Mayer, *Systemic therapy for colorectal cancer*. N Engl J Med, 2005. **352**(5): p. 476-87.
31. Wolpin, B.M. and R.J. Mayer, *Systemic treatment of colorectal cancer*. Gastroenterology, 2008. **134**(5): p. 1296-310.
32. Giantonio, B.J., et al., *Bevacizumab in combination with oxaliplatin, fluorouracil, and leucovorin (FOLFOX4) for previously treated metastatic colorectal cancer: results from the Eastern Cooperative Oncology Group Study E3200*. J Clin Oncol, 2007. **25**(12): p. 1539-44.
33. Hemming, A.W., et al., *Prognostic markers of colorectal cancer: an evaluation of DNA content, epidermal growth factor receptor, and Ki-67*. J Surg Oncol, 1992. **51**(3): p. 147-52.
34. Mayer, A., et al., *The prognostic significance of proliferating cell nuclear antigen, epidermal growth factor receptor, and mdr gene expression in colorectal cancer*. Cancer, 1993. **71**(8): p. 2454-60.
35. Bos, J.L., et al., *Prevalence of ras gene mutations in human colorectal cancers*. Nature, 1987. **327**(6120): p. 293-7.
36. Losi, L., G. Luppi, and J. Benhattar, *Assessment of K-ras, Smad4 and p53 gene alterations in colorectal metastases and their role in the metastatic process*. Oncol Rep, 2004. **12**(6): p. 1221-5.
37. Herman, C.R., et al., *Screening for Preclinical Disease: Test and Disease Characteristics*. American Journal of Roentgenology, 2002. **179**(4): p. 825-831.
38. Loong, T.W., *Understanding sensitivity and specificity with the right side of the brain*. Bmj, 2003. **327**(7417): p. 716-9.
39. Mandel, J.S., et al., *Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study*. N Engl J Med, 1993. **328**(19): p. 1365-71.
40. Duffy, M.J., et al., *Use of faecal markers in screening for colorectal neoplasia: a European group on tumor markers position paper*. Int J Cancer 2011. **128**(1): p. 3-11.
41. Allison, J.E., et al., *A comparison of fecal occult-blood tests for colorectal-cancer screening*. N Engl J Med, 1996. **334**(3): p. 155-9.
42. Rabeneck, L., et al., *Cancer Care Ontario guaiac fecal occult blood test (FOBT) laboratory standards: evidentiary base and recommendations*. Clin Biochem, 2008. **41**(16-17): p. 1289-305.

43. van Dam, L., E.J. Kuipers, and M.E. van Leerdam, *Performance improvements of stool-based screening tests*. Best Pract Res Clin Gastroenterol 2010. **24**(4): p. 479-92.
44. Itzkowitz, S., et al., *A simplified, noninvasive stool DNA test for colorectal cancer detection*. Am J Gastroenterol, 2008. **103**(11): p. 2862-70.
45. Lin, J.S., et al. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012 Feb. Report No.: 12-EHC022-EF
46. Olson, J., et al., *DNA stabilization is critical for maximizing performance of fecal DNA-based colorectal cancer tests*. Diagn Mol Pathol, 2005. **14**(3): p. 183-91.
47. Lamlum, H., et al., *APC mutations are sufficient for the growth of early colorectal adenomas*. Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2225-8.
48. Laurent, S., et al., *Familial adenomatous polyposis: clinical presentation, detection and surveillance*. Acta Gastroenterol Belg 2011. **74**(3): p. 415-20.
49. Xue, B., A.K. Dunker, and V.N. Uversky, *The roles of intrinsic disorder in orchestrating the wnt-pathway*. J Biomol Struct Dyn 2012. **29**(5): p. 843-61.
50. Traverso, G., et al., *Detection of APC Mutations in Fecal DNA from Patients with Colorectal Tumors*. New England Journal of Medicine, 2002. **346**(5): p. 311-320.
51. Puckett, J. and M. Soop, *Optimizing colonoscopy preparation: the role of dosage, timing and diet*. Curr Opin Clin Nutr Metab Care 2012. **15**(5): p.499-504.
52. Baxter, N.N., et al., *Association of colonoscopy and death from colorectal cancer*. Ann Intern Med, 2009. **150**(1): p. 1-8.
53. Muller, A.D. and A. Sonnenberg, *Prevention of colorectal cancer by flexible endoscopy and polypectomy. A case-control study of 32,702 veterans*. Ann Intern Med, 1995. **123**(12): p. 904-10.
54. Singh, H., et al., *Risk of developing colorectal cancer following a negative colonoscopy examination: evidence for a 10-year interval between colonoscopies*. Jama, 2006. **295**(20): p. 2366-73.
55. Grady, W.M. and J.M. Carethers, *Genomic and epigenetic instability in colorectal cancer pathogenesis*. Gastroenterology, 2008. **135**(4): p. 1079-99.
56. Stoler, D.L., et al., *The onset and extent of genomic instability in sporadic colorectal tumor progression*. Proc Natl Acad Sci U S A, 1999. **96**(26): p. 15121-6.
57. Hermsen, M., et al., *Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability*. Gastroenterology, 2002. **123**(4): p. 1109-19.
58. Fearon, E.R., et al., *Identification of a chromosome 18q gene that is altered in colorectal cancers*. Science, 1990. **247**(4938): p. 49-56.
59. Jen, J., et al., *Allelic loss of chromosome 18q and prognosis in colorectal cancer*. N Engl J Med, 1994. **331**(4): p. 213-21.
60. Lurje, G., W. Zhang, and H.J. Lenz, *Molecular prognostic markers in locally advanced colon cancer*. Clin Colorectal Cancer, 2007. **6**(10): p. 683-90.
61. Popat, S. and R.S. Houlston, *A systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC status and colorectal cancer prognosis*. Eur J Cancer, 2005. **41**(14): p. 2060-70.
62. Bronner, C.E., et al., *Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer*. Nature, 1994. **368**(6468): p. 258-61.
63. Shen, L., et al., *Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer*. Proc Natl Acad Sci U S A, 2007. **104**(47): p. 18654-9.
64. Samowitz, W.S., et al., *Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer*. Gastroenterology, 2005. **129**(3): p. 837-45.
65. Ward, R.L., et al., *Adverse prognostic effect of methylation in colorectal cancer is reversed by microsatellite instability*. J Clin Oncol, 2003. **21**(20): p. 3729-36.

66. Ogino, S., et al., *CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer*. Gut, 2009. **58**(1): p. 90-6.
67. Ogino, S. and A. Goel, *Molecular classification and correlates in colorectal cancer*. J Mol Diagn, 2008. **10**(1): p. 13-27.
68. Wang, Y., et al., *Gene expression profiles and molecular markers to predict recurrence of Dukes' B colon cancer*. J Clin Oncol, 2004. **22**(9): p. 1564-71.
69. Barrier, A., et al., *Colon cancer prognosis prediction by gene expression profiling*. Oncogene, 2005. **24**(40): p. 6155-64.
70. Barrier, A., et al., *Stage II colon cancer prognosis prediction by tumor gene expression profiling*. J Clin Oncol, 2006. **24**(29): p. 4685-91.
71. Eschrich, S., et al., *Molecular staging for survival prediction of colorectal cancer patients*. J Clin Oncol, 2005. **23**(15): p. 3526-35.
72. Agesen, T.H., et al., *ColoGuideEx: a robust gene classifier specific for stage II colorectal cancer prognosis*. Gut 2012. **61**(11): p.1560-7
73. Ciardiello, F. and G. Tortora, *EGFR antagonists in cancer treatment*. N Engl J Med, 2008. **358**(11): p. 1160-74.
74. Nicholson, R.I., J.M. Gee, and M.E. Harper, *EGFR and cancer prognosis*. Eur J Cancer, 2001. **37 Suppl 4**: p. S9-15.
75. Deschoolmeester, V., et al., *A review of the most promising biomarkers in colorectal cancer: one step closer to targeted therapy*. Oncologist 2010. **15**(7): p. 699-731.
76. Siena, S., et al., *Biomarkers predicting clinical outcome of epidermal growth factor receptor-targeted therapy in metastatic colorectal cancer*. J Natl Cancer Inst, 2009. **101**(19): p. 1308-24.
77. Roth, A.D., et al., *Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial*. J Clin Oncol 2010. **28**(3): p. 466-74.
78. French, A.J., et al., *Prognostic significance of defective mismatch repair and BRAF V600E in patients with colon cancer*. Clin Cancer Res, 2008. **14**(11): p. 3408-15.
79. Weisenberger, D.J., et al., *CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer*. Nat Genet, 2006. **38**(7): p. 787-93.
80. Richman, S.D., et al., *KRAS and BRAF mutations in advanced colorectal cancer are associated with poor prognosis but do not preclude benefit from oxaliplatin or irinotecan: results from the MRC FOCUS trial*. J Clin Oncol, 2009. **27**(35): p. 5931-7.
81. Kato, S., et al., *PIK3CA mutation is predictive of poor survival in patients with colorectal cancer*. Int J Cancer, 2007. **121**(8): p. 1771-8.
82. Vogelstein, B., D. Lane, and A.J. Levine, *Surfing the p53 network*. Nature, 2000. **408**(6810): p. 307-10.
83. Russo, A., et al., *The TP53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: influence of tumor site, type of mutation, and adjuvant treatment*. J Clin Oncol, 2005. **23**(30): p. 7518-28.
84. Bunz, F., et al., *Disruption of p53 in human cancer cells alters the responses to therapeutic agents*. J Clin Invest, 1999. **104**(3): p. 263-9.
85. Hemminki, A., et al., *Microsatellite instability is a favorable prognostic indicator in patients with colorectal cancer receiving chemotherapy*. Gastroenterology, 2000. **119**(4): p. 921-8.
86. Ribic, C.M., et al., *Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer*. N Engl J Med, 2003. **349**(3): p. 247-57.
87. Iacopetta, B., *TP53 mutation in colorectal cancer*. Hum Mutat, 2003. **21**(3): p. 271-6.

88. Lievre, A., et al., *KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer*. *Cancer Res*, 2006. **66**(8): p. 3992-5.
89. Ruzzo, A., et al., *Role of KRAS let-7 LCS6 SNP in metastatic colorectal cancer patients*. *Ann Oncol* 2011. **22**(1): p. 234-5.
90. Van Schaeybroeck, S., et al., *Implementing prognostic and predictive biomarkers in CRC clinical trials*. *Nat Rev Clin Oncol* 2011. **8**(4): p. 222-32.
91. Yantiss, R.K., et al., *Clinical, pathologic, and molecular features of early-onset colorectal carcinoma*. *Am J Surg Pathol*, 2009. **33**(4): p. 572-82.
92. Cejas, P., et al., *Analysis of the concordance in the EGFR pathway status between primary tumors and related metastases of colorectal cancer patients: Implications for cancer therapy*. *Curr Cancer Drug Targets* 2012 Feb;12(2):124-31.
93. Chee, C.E. and F.A. Sinicrope, *Targeted therapeutic agents for colorectal cancer*. *Gastroenterol Clin North Am* 2010. **39**(3): p. 601-13.
94. Gold, P. and S.O. Freedman, *Specific carcinoembryonic antigens of the human digestive system*. *J Exp Med*, 1965. **122**(3): p. 467-81.
95. Duffy, M.J., *Carcinoembryonic antigen as a marker for colorectal cancer: is it clinically useful?* *Clin Chem*, 2001. **47**(4): p. 624-30.
96. Beauchemin, N., et al., *Redefined nomenclature for members of the carcinoembryonic antigen family*. *Exp Cell Res*, 1999. **252**(2): p. 243-9.
97. Hammarstrom, S., *The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues*. *Semin Cancer Biol*, 1999. **9**(2): p. 67-81.
98. Fletcher, R.H., *Carcinoembryonic antigen*. *Ann Intern Med*, 1986. **104**(1): p. 66-73.
99. Wanebo, H.J., et al., *Preoperative carcinoembryonic antigen level as a prognostic indicator in colorectal cancer*. *N Engl J Med*, 1978. **299**(9): p. 448-51.
100. Ballesta, A.M., et al., *Carcinoembryonic antigen in staging and follow-up of patients with solid tumors*. *Tumour Biol*, 1995. **16**(1): p. 32-41.
101. Goldstein, M.J. and E.P. Mitchell, *Carcinoembryonic antigen in the staging and follow-up of patients with colorectal cancer*. *Cancer Invest*, 2005. **23**(4): p. 338-51.
102. Tan, E., et al., *Diagnostic precision of carcinoembryonic antigen in the detection of recurrence of colorectal cancer*. *Surg Oncol*, 2009. **18**(1): p. 15-24.
103. Wang, J.Y., et al., *Prognostic significance of pre- and postoperative serum carcinoembryonic antigen levels in patients with colorectal cancer*. *Eur Surg Res*, 2007. **39**(4): p. 245-50.
104. Takagawa, R., et al., *Preoperative serum carcinoembryonic antigen level as a predictive factor of recurrence after curative resection of colorectal cancer*. *Ann Surg Oncol*, 2008. **15**(12): p. 3433-9.
105. Bhatnagar, J., et al., *Comparison of carcinoembryonic antigen in tissue and serum with grade and stage of colon cancer*. *Anticancer Res*, 1999. **19**(3B): p. 2181-7.
106. Rieger, A. and B. Wahren, *CEA levels at recurrence and metastases; importance for detecting secondary disease*. *Scand J Gastroenterol*, 1975. **10**(8): p. 869-74.
107. Carpelan-Holmstrom, M., et al., *CEA, CA 242, CA 19-9, CA 72-4 and hCGbeta in the diagnosis of recurrent colorectal cancer*. *Tumour Biol*, 2004. **25**(5-6): p. 228-34.
108. Herszenyi, L., et al., *Tumor marker utility and prognostic relevance of cathepsin B, cathepsin L, urokinase-type plasminogen activator, plasminogen activator inhibitor type-1, CEA and CA 19-9 in colorectal cancer*. *BMC Cancer*, 2008. **8**: p. 194.
109. Bormer, O.P., *Major disagreements between immunoassays of carcinoembryonic antigen may be caused by nonspecific cross-reacting antigen 2 (NCA-2)*. *Clin Chem*, 1991. **37**(10 Pt 1): p. 1736-9.
110. Burtin, P., G. Chavanel, and H. Hirsch-Marie, *Characterization of a second normal antigen that cross-reacts with CEA*. *J Immunol*, 1973. **111**(6): p. 1926-8.



111. Hanada, H., et al., *Early detection of colorectal cancer metastasis and relapse by recognizing nonspecific cross-reacting antigen 2 in commercial carcinoembryonic antigen assays*. Clin Chem, 2009. **55**(9): p. 1747-8.
112. Siepen, D., et al., *Carcinoembryonic antigen (CEA) and two crossreacting antigens of 165 kD and 105 kD isolated from meconium exhibit identical amino terminal sequences*. Biochem Biophys Res Commun, 1987. **147**(1): p. 212-8.
113. Paulson, J.C., *Glycoproteins: what are the sugar chains for?* Trends in Biochemical Sciences, 1989. **14**(7): p. 272-276.
114. Yamashita, K., et al., *Carbohydrate structures of nonspecific cross-reacting antigen-2, a glycoprotein purified from meconium as an antigen cross-reacting with anticarcinoembryonic antigen antibody. Occurrence of complex-type sugar chains with the Gal beta 1----3GlcNAc beta 1----3Gal beta 1----4GlcNAc beta 1----outer chains*. J Biol Chem, 1989. **264**(30): p. 17873-81.
115. Hedin, A., et al., *A monoclonal antibody-enzyme immunoassay for serum carcinoembryonic antigen with increased specificity for carcinomas*. Proc Natl Acad Sci U S A, 1983. **80**(11): p. 3470-4.
116. Matsuoka, Y., *[CEA and related antigens as possible new tumor markers]*. Gan To Kagaku Ryoho, 1983. **10**(8): p. 1754-63.
117. Neumaier, M., U. Fenger, and C. Wagener, *Monoclonal antibodies for carcinoembryonic antigen (CEA) as a model system: identification of two novel CEA-related antigens in meconium and colorectal carcinoma tissue by Western blots and differential immunoaffinity chromatography*. J Immunol, 1985. **135**(5): p. 3604-9.
118. Grunert, F., et al., *Two CEA and three NCA species, although distinguishable by monoclonal antibodies, have nearly identical peptide patterns*. Int J Cancer, 1985. **36**(3): p. 357-62.
119. He, L. and G.J. Hannon, *MicroRNAs: small RNAs with a big role in gene regulation*. Nat Rev Genet, 2004. **5**(7): p. 522-31.
120. Hammond, S.M., *RNAi, microRNAs, and human disease*. Cancer Chemother Pharmacol, 2006. **58 Suppl 1**: p. s63-8.
121. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14*, 1993. **75**(5): p. 843-54.
122. Griffiths-Jones, S., et al., *miRBase: microRNA sequences, targets and gene nomenclature*. Nucleic Acids Res, 2006. **34**(Database issue): p. D140-4.
123. Kozomara, A. and S. Griffiths-Jones, *miRBase: integrating microRNA annotation and deep-sequencing data*. Nucleic Acids Res 2011. **39**(Database issue): p. D152-7.
124. Rodriguez, A., et al., *Identification of mammalian microRNA host genes and transcription units*. Genome Res, 2004. **14**(10A): p. 1902-10.
125. Kim, V.N. and J.W. Nam, *Genomics of microRNA*. Trends Genet, 2006. **22**(3): p. 165-73.
126. Yi, R., et al., *Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs*. Genes Dev, 2003. **17**(24): p. 3011-6.
127. Bartel, D.P., *MicroRNAs: genomics, biogenesis, mechanism, and function*. Cell, 2004. **116**(2): p. 281-97.
128. Chendrimada, T.P., et al., *TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing*. Nature, 2005. **436**(7051): p. 740-4.
129. Bartel, D.P., *MicroRNAs: target recognition and regulatory functions*. Cell, 2009. **136**(2): p. 215-33.
130. Schwarz, D.S., et al., *Asymmetry in the assembly of the RNAi enzyme complex*. Cell, 2003. **115**(2): p. 199-208.
131. Obernosterer, G., et al., *Post-transcriptional regulation of microRNA expression*. Rna, 2006. **12**(7): p. 1161-7.

132. Lee, E.J., et al., *Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors*. Rna, 2008. **14**(1): p. 35-42.
133. Orom, U.A., F.C. Nielsen, and A.H. Lund, *MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation*. Mol Cell, 2008. **30**(4): p. 460-71.
134. Tay, Y., et al., *MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation*. Nature, 2008. **455**(7216): p. 1124-8.
135. Pillai, R.S., S.N. Bhattacharyya, and W. Filipowicz, *Repression of protein synthesis by miRNAs: how many mechanisms?* Trends Cell Biol, 2007. **17**(3): p. 118-26.
136. Eulalio, A., et al., *Deadenylation is a widespread effect of miRNA regulation*. Rna, 2009. **15**(1): p. 21-32.
137. Eulalio, A., et al., *Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing*. Genes Dev, 2007. **21**(20): p. 2558-70.
138. Behm-Ansmant, I., et al., *mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes*. Genes Dev, 2006. **20**(14): p. 1885-98.
139. Huntzinger, E. and E. Izaurralde, *Gene silencing by microRNAs: contributions of translational repression and mRNA decay*. Nat Rev Genet 2011. **12**(2): p. 99-110.
140. Selbach, M., et al., *Widespread changes in protein synthesis induced by microRNAs*. Nature, 2008. **455**(7209): p. 58-63.
141. Baek, D., et al., *The impact of microRNAs on protein output*. Nature, 2008. **455**(7209): p. 64-71.
142. Calin, G.A., 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**(9): p. 2999-3004.
143. Calin, G.A. and C.M. Croce, *MicroRNA signatures in human cancers*. Nat Rev Cancer, 2006. **6**(11): p. 857-66.
144. Pasquinelli, A.E., et al., *Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA*. Nature, 2000. **408**(6808): p. 86-9.
145. Calin, G.A., et al., *Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia*. Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15524-9.
146. Lu, J., et al., *MicroRNA expression profiles classify human cancers*. Nature, 2005. **435**(7043): p. 834-8.
147. Volinia, S., et al., *A microRNA expression signature of human solid tumors defines cancer gene targets*. Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2257-61.
148. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs*. Genome Res, 2009. **19**(1): p. 92-105.
149. Zhang, T., et al., *The oncogenetic role of microRNA-31 as a potential biomarker in oesophageal squamous cell carcinoma*. Clin Sci (Lond) 2011. **121**(10): p. 437-47.
150. Albulescu, R., et al., *Tissular and soluble miRNAs for diagnostic and therapy improvement in digestive tract cancers*. Expert Rev Mol Diagn 2011. **11**(1): p. 101-20.
151. Nana-Sinkam, S.P., M. Fabbri, and C.M. Croce, *MicroRNAs in cancer: personalizing diagnosis and therapy*. Ann N Y Acad Sci 2010. **1210**: p. 25-33.
152. Song, B. and J. Ju, *Impact of miRNAs in gastrointestinal cancer diagnosis and prognosis*. Expert Rev Mol Med 2010. **12**: p. e33.
153. Heneghan, H.M., N. Miller, and M.J. Kerin, *MiRNAs as biomarkers and therapeutic targets in cancer*. Curr Opin Pharmacol 2010. **10**(5): p. 543-50.
154. Yanaihara, N., et al., *Unique microRNA molecular profiles in lung cancer diagnosis and prognosis*. Cancer Cell, 2006. **9**(3): p. 189-98.
155. Rosenfeld, N., et al., *MicroRNAs accurately identify cancer tissue origin*. Nat Biotechnol, 2008. **26**(4): p. 462-9.
156. Cummins, J.M., et al., *The colorectal microRNAome*. Proc Natl Acad Sci U S A, 2006. **103**(10): p. 3687-92.

157. Bandres, E., et al., *Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues*. Mol Cancer, 2006. **5**: p. 29.
158. Schetter, A.J., et al., *MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma*. Jama, 2008. **299**(4): p. 425-36.
159. Michael, M.Z., et al., *Reduced accumulation of specific microRNAs in colorectal neoplasia*. Mol Cancer Res, 2003. **1**(12): p. 882-91.
160. Wang, C.J., et al., *Clinicopathological significance of microRNA-31, -143 and -145 expression in colorectal cancer*. Dis Markers, 2009. **26**(1): p. 27-34.
161. Schepeler, T., et al., *Diagnostic and prognostic microRNAs in stage II colon cancer*. Cancer Res, 2008. **68**(15): p. 6416-24.
162. Monzo, M., et al., *Overlapping expression of microRNAs in human embryonic colon and colorectal cancer*. Cell Res, 2008. **18**(8): p. 823-33.
163. Guo, C., et al., *The noncoding RNA, miR-126, suppresses the growth of neoplastic cells by targeting phosphatidylinositol 3-kinase signaling and is frequently lost in colon cancers*. Genes Chromosomes Cancer, 2008. **47**(11): p. 939-46.
164. Mongroo, P.S. and A.K. Rustgi, *The role of the miR-200 family in epithelial-mesenchymal transition*. Cancer Biol Ther 2010. **10**(3): p. 219-22.
165. Slattery, M.L., et al., *MicroRNAs and colon and rectal cancer: differential expression by tumor location and subtype*. Genes Chromosomes Cancer 2011. **50**(3): p. 196-206.
166. Slaby, O., et al., *Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer*. Oncology, 2007. **72**(5-6): p. 397-402.
167. Diaz, R., et al., *Deregulated expression of miR-106a predicts survival in human colon cancer patients*. Genes Chromosomes Cancer, 2008. **47**(9): p. 794-802.
168. Motoyama, K., et al., *Over- and under-expressed microRNAs in human colorectal cancer*. Int J Oncol, 2009. **34**(4): p. 1069-75.
169. Xi, Y., et al., *Prognostic Values of microRNAs in Colorectal Cancer*. Biomark Insights, 2006. **2**: p. 113-121.
170. Lanza, G., et al., *mRNA/microRNA gene expression profile in microsatellite unstable colorectal cancer*. Mol Cancer, 2007. **6**: p. 54.
171. Nakajima, G., et al., *Non-coding MicroRNAs hsa-let-7g and hsa-miR-181b are Associated with Chemoresponse to S-1 in Colon Cancer*. Cancer Genomics Proteomics, 2006. **3**(5): p. 317-324.
172. Svoboda, M., et al., *Micro-RNAs miR125b and miR137 are frequently upregulated in response to capecitabine chemoradiotherapy of rectal cancer*. Int J Oncol, 2008. **33**(3): p. 541-7.
173. Ng, E.K., et al., *Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening*. Gut, 2009. **58**(10): p. 1375-81.
174. Yamamichi, N., et al., *Locked nucleic acid in situ hybridization analysis of miR-21 expression during colorectal cancer development*. Clin Cancer Res, 2009. **15**(12): p. 4009-16.
175. Huang, Z., et al., *Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer*. Int J Cancer 2010. **127**(1): p. 118-26.
176. Akcakaya, P., et al., *miR-185 and miR-133b deregulation is associated with overall survival and metastasis in colorectal cancer*. Int J Oncol 2011. **39**(2): p. 311-8.
177. Sarver, A.L., et al., *Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states*. BMC Cancer, 2009. **9**: p. 401.
178. Shibuya, H., et al., *Clinicopathological and prognostic value of microRNA-21 and microRNA-155 in colorectal cancer*. Oncology 2010. **79**(3-4): p. 313-20.
179. Manne, U., et al., *miRNAs as biomarkers for management of patients with colorectal cancer*. Biomark Med 2010. **4**(5): p. 761-70.

180. Mitchell, P.S., et al., *Circulating microRNAs as stable blood-based markers for cancer detection*. Proc Natl Acad Sci U S A, 2008. **105**(30): p. 10513-8.
181. Lawrie, C.H., et al., *Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma*. Br J Haematol, 2008. **141**(5): p. 672-5.
182. Chen, X., et al., *Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases*. Cell Res, 2008. **18**(10): p. 997-1006.
183. Wu, C.W., et al., *Detection of miR-92a and miR-21 in stool samples as potential screening biomarkers for colorectal cancer and polyps*. Gut 2012. **61**(5):p. 739-45.
184. Link, A., et al., *Fecal MicroRNAs as novel biomarkers for colon cancer screening*. Cancer Epidemiol Biomarkers Prev 2010. **19**(7): p. 1766-74.
185. Ahmed, F.E., et al., *Diagnostic microRNA markers for screening sporadic human colon cancer and active ulcerative colitis in stool and tissue*. Cancer Genomics Proteomics, 2009. **6**(5): p. 281-95.
186. Nagel, R., et al., *Regulation of the adenomatous polyposis coli gene by the miR-135 family in colorectal cancer*. Cancer Res, 2008. **68**(14): p. 5795-802.
187. Kalimutho, M., et al., *Differential expression of miR-144\* as a novel fecal-based diagnostic marker for colorectal cancer*. J Gastroenterol 2011 **46**(12):1391-402.
188. Kalimutho, M., et al., *Epigenetically silenced miR-34b/c as a novel faecal-based screening marker for colorectal cancer*. Br J Cancer 2011. **104**(11): p. 1770-8.
189. Koga, Y., et al., *Exosome can prevent RNase from degrading microRNA in feces*. J Gastrointest Oncol 2011. **2**(4): p. 215-22.
190. Hunter, M.P., et al., *Detection of microRNA expression in human peripheral blood microvesicles*. PLoS One, 2008. **3**(11): p. e3694.
191. Valadi, H., et al., *Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells*. Nat Cell Biol, 2007. **9**(6): p. 654-9.
192. Couzin, J., *Cell biology: The ins and outs of exosomes*. Science, 2005. **308**(5730): p. 1862-3.
193. Selcuklu, S.D., M.T. Donoghue, and C. Spillane, *miR-21 as a key regulator of oncogenic processes*. Biochem Soc Trans, 2009. **37**(Pt 4): p. 918-25.
194. Drebber, U., et al., *Altered levels of the onco-microRNA 21 and the tumor-suppressor microRNAs 143 and 145 in advanced rectal cancer indicate successful neoadjuvant chemoradiotherapy*. Int J Oncol 2011. **39**(2): p. 409-15.
195. Kulda, V., et al., *Relevance of miR-21 and miR-143 expression in tissue samples of colorectal carcinoma and its liver metastases*. Cancer Genet Cytogenet 2010. **200**(2): p. 154-60.
196. Asangani, I.A., et al., *MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdc4 and stimulates invasion, intravasation and metastasis in colorectal cancer*. Oncogene, 2008. **27**(15): p. 2128-36.
197. Allgayer, H., *Pdc4, a colon cancer prognostic that is regulated by a microRNA*. Crit Rev Oncol Hematol 2010. **73**(3): p. 185-91.
198. Fassan, M., et al., *PDCD4 nuclear loss inversely correlates with miR-21 levels in colon carcinogenesis*. Virchows Arch 2011. **458**(4): p. 413-9.
199. Ma, Y., et al., *miR-150 as a potential biomarker associated with prognosis and therapeutic outcome in colorectal cancer*. Gut 2012 **61**(10):1447-53.
200. Zhang, G., et al., *Clinical significance of miR-22 expression in patients with colorectal cancer*. Med Oncol 2012. [Epub ahead of print].
201. Yu, G., et al., *Prognostic values of the miR-17-92 cluster and its paralogs in colon cancer*. J Surg Oncol 2012 **106**(3):232-7.
202. Pu, X.X., et al., *Circulating miR-221 directly amplified from plasma is a potential diagnostic and prognostic marker of colorectal cancer and is correlated with p53 expression*. J Gastroenterol Hepatol 2010. **25**(10): p. 1674-80.

203. Cheng, H., et al., *Circulating plasma MiR-141 is a novel biomarker for metastatic colon cancer and predicts poor prognosis*. PLoS One 2011. **6**(3): p. e17745.
204. Lee, J.A., et al., *Prognostic Implications of MicroRNA-21 Overexpression in Invasive Ductal Carcinomas of the Breast*. J Breast Cancer 2011. **14**(4): p. 269-75.
205. Zaman, M.S., et al., *Up-regulation of microRNA-21 correlates with lower kidney cancer survival*. PLoS One 2012. **7**(2): p. e31060.
206. Wang, Z.X., et al., *Prognostic significance of serum miRNA-21 expression in human non-small cell lung cancer*. J Surg Oncol 2011. **104**(7): p. 847-51.
207. Balaguer, F., et al., *Colorectal cancers with microsatellite instability display unique miRNA profiles*. Clin Cancer Res 2011. **17**(19): p. 6239-49.
208. Earle, J.S., et al., *Association of microRNA expression with microsatellite instability status in colorectal adenocarcinoma 2010*. J Mol Diagn. **12**(4): p. 433-40.
209. Rossi, L., E. Bonmassar, and I. Faraoni, *Modification of miR gene expression pattern in human colon cancer cells following exposure to 5-fluorouracil in vitro*. Pharmacol Res, 2007. **56**(3): p. 248-53.
210. Ruzzo, A., et al., *High Let-7a MicroRNA Levels in KRAS-Mutated Colorectal Carcinomas May Rescue Anti-EGFR Therapy Effects in Patients with Chemotherapy-Refractory Metastatic Disease*. Oncologist 2012;17(6):823-9..
211. Akao, Y., Y. Nakagawa, and T. Naoe, *let-7 microRNA functions as a potential growth suppressor in human colon cancer cells*. Biol Pharm Bull, 2006. **29**(5): p. 903-6.
212. Ng, E.K., et al., *MicroRNA-143 targets DNA methyltransferases 3A in colorectal cancer*. Br J Cancer, 2009. **101**(4): p. 699-706.
213. Chen, X., et al., *Role of miR-143 targeting KRAS in colorectal tumorigenesis*. Oncogene, 2009. **28**(10): p. 1385-92.
214. Akao, Y., Y. Nakagawa, and T. Naoe, *MicroRNA-143 and -145 in colon cancer*. DNA Cell Biol, 2007. **26**(5): p. 311-20.
215. Borralho, P.M., et al., *miR-143 overexpression impairs growth of human colon carcinoma xenografts in mice with induction of apoptosis and inhibition of proliferation*. PLoS One 2011. **6**(8): p. e23787.
216. Song, B., et al., *miR-192 Regulates dihydrofolate reductase and cellular proliferation through the p53-microRNA circuit*. Clin Cancer Res, 2008. **14**(24): p. 8080-6.
217. Borralho, P.M., et al., *MicroRNA-143 reduces viability and increases sensitivity to 5-fluorouracil in HCT116 human colorectal cancer cells*. Febs J, 2009. **276**(22): p. 6689-700.
218. Song, B., et al., *Mechanism of chemoresistance mediated by miR-140 in human osteosarcoma and colon cancer cells*. Oncogene, 2009. **19**;28(46):4065-74.
219. Henry, J.C., A.C. Azevedo-Pouly, and T.D. Schmittgen, *MicroRNA replacement therapy for cancer*. Pharm Res 2011. **28**(12): p. 3030-42.
220. Krutzfeldt, J., et al., *Silencing of microRNAs in vivo with 'antagomirs'*. Nature, 2005. **438**(7068): p. 685-9.
221. Gondi, C.S. and J.S. Rao, *Concepts in in vivo siRNA delivery for cancer therapy*. J Cell Physiol, 2009. **220**(2): p. 285-91.
222. Flatmark, K., et al., *Nuclear localization of the metastasis-related protein S100A4 correlates with tumour stage in colorectal cancer*. J Pathol, 2003. **200**(5): p. 589-95.
223. Boye, K., et al., *EMMPRIN is associated with S100A4 and predicts patient outcome in colorectal cancer*. Br J Cancer 2012. **7**;107(4):667-74.
224. Flatmark, K., et al., *Disseminated tumour cells as a prognostic biomarker in colorectal cancer*. Br J Cancer 2011. **104**(9): p. 1434-9.
225. Boye, K., et al., *Nuclear S100A4 is a novel prognostic marker in colorectal cancer 2010*. Eur J Cancer. **46**(16): p. 2919-25.

226. Ingebrigtsen, V.A., et al., *B7-H3 expression in colorectal cancer: Nuclear localization strongly predicts poor outcome in colon cancer*. *Int J Cancer*. 2012 Dec 1;131(11):2528-36.
227. Liang, Y., et al., *Characterization of microRNA expression profiles in normal human tissues*. *BMC Genomics*, 2007. **8**: p. 166.
228. Paus, E., et al., *A single-chain-Fv-based immunofluorometric assay specific for the CEA variant NCA-2*. *J Immunol Methods*, 2003. **283**(1-2): p. 125-39.
229. Bjerner, J., et al., *Protein epitopes in carcinoembryonic antigen. Report of the ISOBM TD8 workshop*. *Tumour Biol*, 2002. **23**(4): p. 249-62.
230. Kuroki, M., et al., *Reaction profiles of seven enzyme immunoassay kits for carcinoembryonic antigen (CEA) analyzed with purified preparations of CEA and related normal antigens*. *Clin Biochem*, 1992. **25**(1): p. 29-35.
231. Kuroki, M., et al., *Antigenic reactivities of purified preparations of carcinoembryonic antigen (CEA) and related normal antigens using four different radioimmunoassay systems for CEA*. *J Immunol Methods*, 1983. **60**(1-2): p. 221-33.
232. Burtin, P., M.C. Sabine, and G. Chavanel, *A comparative study of the localization of CEA and NCA2 in cancerous and normal gastrointestinal tissues*. *Int J Cancer*, 1977. **19**(5): p. 634-41.
233. van Essen, H.F., et al., *Alcohol based tissue fixation as an alternative for formaldehyde: influence on immunohistochemistry*. *J Clin Pathol* 2010. **63**(12): p. 1090-4.
234. Jones, M.L., *How formalin affects the outcome of routine and special stains*. *Biotech Histochem*, 2007. **82**(3): p. 155-9.
235. Schee, K., et al., *Investigation of nonspecific cross-reacting antigen 2 as a prognostic biomarker in bone marrow plasma from colorectal cancer patients*. *Tumour Biol* 2012. **33**(1): p. 73-83.
236. van Rooij, E., *The art of microRNA research*. *Circ Res* 2011. **108**(2): p. 219-34.
237. Benes, V. and M. Castoldi, *Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available*. *Methods* 2010. **50**(4): p. 244-9.
238. geNORM. [cited; Available from: <http://medgen.ugent.be/~jvdesomp/genorm/>].
239. Motameny, S., et al., *Next Generation Sequencing of miRNAs â€œ Strategies, Resources and Methods 2010*. p. 70-84.
240. t Hoen, P.A., 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**(21): p. e141.
241. Willenbrock, H., et al., *Quantitative miRNA expression analysis: comparing microarrays with next-generation sequencing*. *Rna*, 2009. **15**(11): p. 2028-34.
242. Biasiolo, M., et al., *Impact of host genes and strand selection on miRNA and miRNA\* expression*. *PLoS One* 2011. **6**(8): p. e23854.
243. Hu, H.Y., et al., *Sequence features associated with microRNA strand selection in humans and flies*. *BMC Genomics*, 2009. **10**: p. 413.
244. Guo, L., T. Liang, and Z. Lu, *A comprehensive study of multiple mapping and feature selection for correction strategy in the analysis of small RNAs from SOLiD sequencing*. *Biosystems* 2011. **104**(2-3): p. 87-93.
245. Schee, K., O. Fodstad, and K. Flatmark, *MicroRNAs as biomarkers in colorectal cancer*. *Am J Pathol* 2010. **177**(4): p. 1592-9.
246. Gregersen, L.H., et al., *MicroRNA-145 targets YES and STAT1 in colon cancer cells 2010*. *PLoS One*. **5**(1): p. e8836.
247. Schepeler, T., et al., *Attenuation of the beta-catenin/TCF4 complex in colorectal cancer cells induces several growth-suppressive microRNAs that target cancer promoting genes*. *Oncogene* 2012 **31**(22):2750-60.
248. Sachdeva, M. and Y.Y. Mo, *MicroRNA-145 suppresses cell invasion and metastasis by directly targeting mucin 1*. *Cancer Res* 2010. **70**(1): p. 378-87.

249. Ibrahim, A.F., et al., *MicroRNA replacement therapy for miR-145 and miR-33a is efficacious in a model of colon carcinoma*. *Cancer Res* 2011. **71**(15): p. 5214-24.
250. Ma, Y., et al., *Candidate microRNA biomarkers in human colorectal cancer: Systematic review profiling studies and experimental validation*. *Int J Cancer* 2012 **130**(9):2077-87.
251. Vaz, C., et al., *Analysis of microRNA transcriptome by deep sequencing of small RNA libraries of peripheral blood*. *BMC Genomics* 2010. **11**: p. 288.
252. Chang, K.H., et al., *MicroRNA signature analysis in colorectal cancer: identification of expression profiles in stage II tumors associated with aggressive disease*. *Int J Colorectal Dis* 2011. **26**(11): p. 1415-22.
253. Braun, C.J., et al., *p53-Responsive micrornas 192 and 215 are capable of inducing cell cycle arrest*. *Cancer Res*, 2008. **68**(24): p. 10094-104.
254. Boni, V., et al., *miR-192/miR-215 influence 5-fluorouracil resistance through cell cycle-mediated mechanisms complementary to its post-transcriptional thymidilate synthase regulation*. *Mol Cancer Ther* 2010. **9**(8): p. 2265-75.
255. Valeri, N., C.M. Croce, and M. Fabbri, *Pathogenetic and clinical relevance of microRNAs in colorectal cancer*. *Cancer Genomics Proteomics*, 2009. **6**(4): p. 195-204.
256. Tang, J.T. and J.Y. Fang, *MicroRNA regulatory network in human colorectal cancer*. *Mini Rev Med Chem*, 2009. **9**(8): p. 921-6.
257. Slaby, O., et al., *MicroRNAs in colorectal cancer: translation of molecular biology into clinical application*. *Mol Cancer*, 2009. **8**: p. 102.
258. Li, M., et al., *MicroRNAs: control and loss of control in human physiology and disease*. *World J Surg*, 2009. **33**(4): p. 667-84.
259. Faber, C., T. Kirchner, and F. Hlubek, *The impact of microRNAs on colorectal cancer*. *Virchows Arch*, 2009. **454**(4): p. 359-67.
260. Aslam, M.I., et al., *MicroRNAs are novel biomarkers of colorectal cancer*. *Br J Surg*, 2009. **96**(7): p. 702-10.





## **PAPER I**

Kristina Schee, Kjersti Flatmark, Ruth Holm, Kjetil Boye, Elisabeth Paus.

**“Investigation of nonspecific cross-reacting antigen 2 as a prognostic biomarker in bone marrow plasma from colorectal cancer patients.”**

Tumour Biol. 2012 Feb;33(1):73-83. Epub 2011 Oct 18.



## **ERRATUM PAPER I**

Kristina Schee, Kjersti Flatmark, Ruth Holm, Kjetil Boye, Elisabeth Paus.

**“Investigation of nonspecific cross-reacting antigen 2 as a prognostic biomarker in bone marrow plasma from colorectal cancer patients.”**

Tumour Biol. 2012 Aug;33(4):1263-4



## **PAPER II**

Kristina Schee, Kjetil Boye, Torveig Weum Abrahamsen, Øystein Fodstad, Kjersti Flatmark.

**"Clinical relevance of microRNA miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145 in colorectal cancer."**

Under revision, BMC Cancer



**Clinical relevance of microRNA miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145 in colorectal cancer.**

Kristina Schee\*<sup>1</sup>, Kjetil Boye<sup>1,2</sup>, Torveig Weum Abrahamsen<sup>1</sup>, Øystein Fodstad<sup>1</sup>, Kjersti Flatmark<sup>1,3</sup>.

Departments of <sup>1</sup>Tumor Biology, <sup>2</sup>Oncology and <sup>3</sup>Gastrointestinal Surgery, Clinic for Cancer Surgery and Transplantation, The Norwegian Radium Hospital, Oslo University Hospital, N-0310 Oslo, Norway

\*Corresponding author

Email: Kristina Schee\* - [kristina.schee@rr-research.no](mailto:kristina.schee@rr-research.no)

This work was supported by grants from South-Eastern Norway Regional Health Authority and from the Norwegian Foundation for Health and Rehabilitation.

## **Abstract**

**Background:** MicroRNAs (miRNAs) regulate gene expression by binding to mRNA, and can function as oncogenes or tumor suppressors depending on the target. In this study, using RT-PCR, we examined the expression of six miRNAs (miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145) in tumors from 193 prospectively recruited patients with colorectal cancer, and associations with clinicopathological parameters and patient outcome were analyzed. The miRNAs were chosen based on previous studies for their biomarker potential and suggested biological relevance in colorectal cancer.

**Results:** MiR-101 was hardly expressed in the tumor samples, while for the other miRNAs, variable expression levels and expression ranges were observed, with miR-21 being most abundantly expressed relative to the reference (RNU44). In our study cohort, major clinical significance was demonstrated only for miR-31, as high expression was associated with advanced tumor stage and poor differentiation. No significant associations were found between expression of the investigated miRNAs and metastasis-free or overall survival.

**Conclusions:** Investigating the expression of six miRNAs previously identified as candidate biomarkers in colorectal cancer, few clinically relevant associations were detected in our patient cohort. Our results emphasize the importance of validating potential tumor markers in independent patient cohorts, and indicate that the role of miRNAs as colorectal cancer biomarkers is still undetermined.

**Keywords:** miRNA, colorectal cancer, prognostic biomarker



## **Background**

MicroRNAs (miRNAs) are a class of small, non-coding RNAs (19-22 nucleotides) that function as posttranscriptional gene regulators by binding to the 3'UTR of mRNA, and one miRNA may potentially down-regulate multiple mRNA targets. More than 1500 human miRNAs are currently annotated in the miRBase [1], and it has been predicted that as many as 30% of protein-encoding genes may be regulated by miRNAs [2]. The discovery that miRNAs may function as oncogenes or tumor suppressors depending on the target mRNA, has instigated intensive research to determine the role of these molecules in cancer. MiRNAs are chemically very stable, and can be detected by a range of high-throughput detection methods in tissue, serum and plasma as well as in urine and feces, and are for these reasons considered to have great potential as cancer biomarkers.

In colorectal cancer (CRC), treatment decisions are still based essentially on anatomical extent of disease at diagnosis, and the search for better biomarkers is warranted. Several miRNAs with potential biological and clinical relevance have been identified and are being explored as diagnostic, prognostic and predictive biomarkers [3-6]. Based on previous studies and our recent review of this topic, six candidate miRNAs, miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145 (Table 1), were chosen for analysis in a cohort of 193 prospectively recruited patients receiving curative surgery for CRC [7-13]. Expression of the miRNA was determined by RT-PCR and associations with clinicopathological parameters and outcome were analyzed.

## **Materials and Methods**

### *Patient cohort*

316 patients, recruited from five hospitals in the Oslo region between the year 1998 and 2000 [31], were prospectively included in the study at the time of primary surgery for assumed or verified colorectal cancer. The study was approved by the Regional Ethics Committee (Health Region II, Norway) and informed consent was obtained from the patients. At surgery, resected specimens were routinely processed for histopathological assessment and additional tumor tissue was sampled and snap-frozen in liquid nitrogen. A number of cases were excluded from statistical analysis for the following reasons: not invasive cancer (25), histology other than adenocarcinoma (5), distant metastasis at the time of surgery (34, tissue samples not available), preoperative chemoradiotherapy (2), inadequate surgical margins (7), unknown stage of disease (1), freshly frozen tissue samples not obtainable (46), and high Ct-values ( $>37$ ;  $n=3$ ). The study population thus consisted of 193 patients in TNM stage I-III (Table 2). Follow-up data was obtained from the participating hospitals and from the general practitioners (for the patients not attending scheduled controls). Metastasis was verified by radiological examination and survival data was obtained from the National Registry of Norway and updated by October 1<sup>st</sup> 2008 with the cause of death registered and classified as death from colorectal cancer, death of other cause or death of unknown cause.

### *MiRNA selection*

MiRNA selection was based on previous studies and our literature review [11], identifying miRNA with proposed clinical relevance in CRC, including published articles leading up to the year 2009. We wished to examine selected miRNAs in our CRC cohort and their relevance with clinicopathological data and outcome parameters (Table 2). The following six miRNAs were chosen for analysis; miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145 [7-13].

### *Sample preparation and RNA isolation*

Biopsies were sampled and snap frozen in liquid nitrogen and stored at -80 °C. The biopsies were sectioned using a cryostat microtome and hematoxylin-eosin stained slides were evaluated for tumor content by a pathologist (median tumor content in the samples was 50%, range 30 - 80%). The tumor tissue was sliced into 10 µm sections using a cryostat microtome, aliquoted into 1.5 ml Micro tubes (Sarstedt, Nümbrecht, Germany) and stored at -80 °C. RNA was isolated from the tumor tissue using TriReagent (Ambion Inc, TX) according to the manufacturer's protocol and the total RNA concentration was measured by Nanodrop (ND-1000).

### *RT-PCR*

Total RNA from 196 patients was used to reversely transcribe miRNAs using TaqMan MicroRNA assays (Applied Biosystems, Foster City, CA). Each reverse transcriptase reaction contained 10 ng of total RNA (5µl), 0.15 µl dNTP (100 mM total), 1.0 µl Multiscribe RT enzyme (50 U/µl), 1.5 µl 10X RT buffer, 0.19 µl RNase Inhibitor (20 U/µl) , 4.16 µl nuclease free water (Sigma-Aldrich, Ayshire, UK) and 3.0 µl 5X RT Primer. The 15 µl reaction volumes were incubated in 8-well PCR

strip tubes (Sarstedt) in a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems) as follows; 30 min at 16 °C, 30 min at 42 °C, 5 min at 85 °C. Real-time PCR was performed using Applied Biosystems 7500 real-time PCR system. The reversely transcribed miRNAs were diluted 1:20 before adding 1.3 µl to 10 µl 2X Universal PCR Master Mix (no AmpErase UNG), 7.7 µl water and 1.0 µl 20X MicroRNA Assay. A total volume of 20 µl per reactions was incubated in 96-well MicroAmp plates (Applied Biosystems) for 10 min 95 °C followed by 40 cycles of 15 sec. 95 °C and 60 sec. 60 °C. All samples were run in duplicates. RNU6B and RNU44 were tested as potential reference genes and performed equally well, and RNU44 was selected for further analysis [32]. Each miRNA was normalized against RNU44 and the relative expression was calculated using  $2^{-\Delta\Delta Ct}$  method.

### *Statistical analysis*

All statistical analyses were performed using SPSS version 18.0 (SPSS Inc., Chicago, MO) and P-values < 0.05 were considered to be statistically significant. Associations between miRNA expression and clinicopathological variables were explored using Mann-Whitney U and Kruskal-Wallis test as appropriate. Survival was estimated using the Kaplan-Meier method and compared using the log-rank test. Overall and metastasis-free survival was calculated from date of surgery until date of death or diagnosis of metastasis.

## **Results**

### *MiRNA expression in tumor samples*

The most abundantly expressed miRNA relative to the reference was miR-21, and it also exhibited the widest expression range among the examined candidates (median relative expression ratio was 7.7; range 0.4-61.0). In contrast, miR-101 was hardly detectable in any of the samples (0.02; 0-0.13), and miR-31 exhibited low expression but a wider expression range (0.04; 0-2.6). The remaining three miRNAs, miR-92a (1.9; 0.04-24.4), miR-106a (1.0; 0.1-18.1), and miR-145 (0.5; 0.04-29.8) exhibited intermediate expression levels and variability between samples (Figure 1).

#### *MiRNA expression and associations with clinicopathological parameters*

To explore the clinical significance of these findings, associations with clinicopathological variables were investigated. Somewhat surprisingly, few significant associations were detected between expression of miR-21, miR-92a, miR-101, miR-106a and miR-145 and clinicopathological variables, including age, gender, tumor stage, differentiation, localization and specific histomorphologic characteristics such as vascular invasion, perineural infiltration and lymphocyte infiltration (Table 2). MiR-92a and miR-106a were associated with differentiation, as higher median expression levels were found in intermediately differentiated tumors than in well and poorly differentiated tumors ( $p=0.003$  and  $p=0.01$ , respectively). Also, some associations were found between miR-31, miR-92a and miR106a expression and tumor localization, as miR-31 exhibited higher expression in colon tumors while miR-92a and miR106a had higher expression levels in rectal tumors ( $p=0.02$ ,  $p=0.05$  and  $p=0.05$ , respectively).

For miR-31, an association with tumor stage, and in particular with pT stage was found, as relative median expression of miR-31 increased with pT stage (0.015, 0.02, 0.05, and 0.14 for pT1, pT2, pT3, and pT4, respectively;  $p=0.004$ , Kruskal-Wallis test) (Figure 2 and Table 2). High miR-31 expression was also associated with poorly differentiated tumors, as relative mean expression was 0.2, 0.04 and 0.02 for poor, intermediate and well differentiated tumors, respectively ( $p=0.001$ , Kruskal-Wallis test), which is also in accordance with previous findings [12].

#### *MiRNA expression and associations with patient outcome*

To analyze associations with outcome, survival was estimated using the Kaplan-Meier method and compared using the log-rank test. As there are no generally recognized cut-off values for the miRNAs analyzed in this work, different values were explored to arrange data (including mean, median and tertiles). Regardless of the cut-off value used, we found no significant associations between expression of any of the analyzed miRNAs and metastasis-free (Figure 3) or overall survival. Similar results were obtained using univariate Cox regression analysis with miRNA expression levels as continuous variables (data not shown).

#### **Discussion**

Although miR-31 was expressed at relatively low levels compared with some of the other candidates, high expression was associated with advanced tumor stage at diagnosis, and

particularly with pT-stage, in accordance with previous results [9, 10]. There are multiple predicted targets for miR-31, but few have been fully validated at present. One proposed target of miR-31 that has been experimentally investigated is Special AT-rich Binding protein 2 (SATB2), which is involved in transcriptional regulation and chromatin remodeling [14]. In an immunohistochemical study performed in 146 colorectal tumors, low expression of SATB2 was associated with metastasis development and poor prognosis [15]. Another target that has been shown to be regulated by miR-31 is the T lymphoma Invasion And Metastasis gene 1 (TIAM1), which is a guanidine exchange factor for Rac GTPase and when over-expressed, it prevents TGF- $\beta$  and TNF- $\alpha$  dependent motility and invasion in CRC cell lines [16]. The postulated effects of miR-31 on SATB2 and TIAM1 are consistent with the associations between miR-31 expression and advanced tumor stage, observed by us and others, but clearly, the regulatory activity of miR-31 is still incompletely understood in CRC.

MiR-92a was included in the analyses because it has been proposed as an early-detection biomarker in plasma and stool [13, 17]. In general, one would expect an early-detection biomarker to be ubiquitously expressed in the tissue of interest, and although several tumors in our study had relatively high levels of miR-92a, low levels were found in a substantial proportion of the samples. Also, over-expression of miR-92a has been found in other cancer types, such as hepatocellular carcinoma and leukemia [18, 19], which suggest that further evaluation is necessary to determine its specificity and sensitivity as an early-detection biomarker. Although miR-92a was not primarily included in this study for its prognostic relevance, it was recently proposed as a key oncogenic component of the miR-17-92 cluster through targeting and down-

regulating the proapoptotic protein Bim in CRC, suggesting that the functional role of miR-92a in CRC should be further elucidated [20].

MiR-21 is one of the more extensively studied miRNAs in CRC and was included in our study because of its proposed association with advanced tumor stage and outcome in CRC [12, 21]. In the present work, miR-21 exhibited the highest relative expression and the widest expression range of the examined candidates, but no significant associations with clinicopathological data or outcome were found. Although some investigators have identified this miRNA as clinically relevant, other exploratory studies of miRNA expression in CRC have not been able to verify these findings [22-24]. It has been speculated that discrepancies might be explained by the composition of patient cohorts, particularly regarding tumor localization, as the association between miR-21 and survival has primarily been documented in colon cancer [25]. However, in our cohort no differences were found when comparing the clinical relevance of miR-21 expression in colon and rectum cancer. In most of the previous studies, miR-21 expression was reported relative to paired normal tissues, whereas only tumor tissue was available from our patients, which might influence interpretation of results. However, among the reports that did not identify miR-21 as relevant for outcome in CRC, both analysis of tumor tissue alone and paired tumor and normal samples were used, suggesting that this may not be the only explanation for the discrepancies.

When the primary objective is to identify cancer specific molecules, the inclusion of normal tissues is necessary, whereas, in the current project the aim was to evaluate previously identified potential biomarkers, which is a different setting. Importantly, normal tissue is often



not obtainable for analysis, and expression of molecular targets in normal tissues might vary considerably between patients, and not necessarily in concert with the corresponding tumor sample. Thus, it is probably both practicable and necessary to develop assays that are independent of normal tissue. Another related challenge concerns the definition of biologically relevant cut-off levels, which have not been determined for specific miRNA in different tissues. We explored multiple cut-off levels, but associations with clinicopathological parameters and outcome for all the candidates remained relatively similar.

MiR-101, miR-145 and miR-106a have previously been associated with cancer-relevant biological processes, such as growth, proliferation and inhibition of apoptosis, or with clinical outcome in CRC [7-9], but few associations with clinicopathological parameters or outcome were found in our cohort. MiR-101 was hardly detectable in tumor samples, which is in accordance with its proposed function as a tumor suppressor that is lost during tumorigenesis. Interesting recent findings in pancreatic cancer suggest miR-101 as a key regulator of stem cell protein markers; its loss favoring the stem cell phenotype and its re-expression constituting a possible therapeutic strategy [26]. Down-regulation of miR-145 was also identified as an early event in CRC carcinogenesis, which might explain why associations with clinical variables in invasive tumors were absent in our tumor panel. The biological relevance of miR-145 in CRC has, however, been repeatedly confirmed, and this miRNA is also being explored as a therapeutic target [27, 28]. MiR-106a was in a recent review identified as consistently up-regulated in CRC (relative to normal colon) which would be in agreement with our findings [29]. It has also been identified in stool samples in CRC patients, and has been suggested as an early detection

biomarker [30], but even if extensively studied in several cancer forms, its function and clinical relevance remain unclear.

## **Conclusions**

It has become evident over the last decade that miRNAs contribute to the pathogenesis of a broad variety of human disease, including cancer. Their relatively small number combined with large potential downstream regulatory effects and unique chemical stability make these molecules interesting biomarker candidates. Although the miRNAs analyzed in the present study were chosen on the basis of biomarker potential and biological relevance in CRC, major clinical significance could only be confirmed for miR-31 in our study cohort. It seems clear that the role of miRNAs as colorectal cancer biomarkers is still undetermined, emphasizing the need for further investigations in the exploratory setting and to validate potential biomarkers.

## **Competing interests**

The authors declare that they have no competing interests

## **Authors' contributions**

KS carried out the qRT-PCR and normalization, performed parts of the statistical analysis and drafted the manuscript. KB performed the main statistical analysis. TWA carried out the tissue preparation and RNA isolation. ØF contributed with critical revisions of the manuscript. KF

participated in the study design and coordination and helped draft the manuscript. All authors have read and approved the final manuscript.

### **Acknowledgements**

This work was supported by the South-Eastern Norway Regional Health Authority, by the “Legacy in support of cancer research” and by the The Norwegian Radium Hospital Research Foundation.

## REFERENCES

1. Griffiths-Jones S: **The microRNA Registry**. vol. 32. pp. D109-D111; 2004:D109-D111.
2. Lewis BP, Burge CB, Bartel DP: **Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets**. *Cell* 2005, **120**:15-20.
3. Ahmed FE, Jeffries CD, Vos PW, Flake G, Nuovo GJ, Sinar DR, Naziri W, Marcuard SP: **Diagnostic microRNA markers for screening sporadic human colon cancer and active ulcerative colitis in stool and tissue**. *Cancer Genomics Proteomics* 2009, **6**:281-295.
4. Bandres E, Agirre X, Bitarte N, Ramirez N, Zarate R, Roman-Gomez J, Prosper F, Garcia-Foncillas J: **Epigenetic regulation of microRNA expression in colorectal cancer**. *Int J Cancer* 2009, **125**:2737-2743.
5. Shibuya H, Iinuma H, Shimada R, Horiuchi A, Watanabe T: **Clinicopathological and prognostic value of microRNA-21 and microRNA-155 in colorectal cancer**. *Oncology* 2010, **79**:313-320.
6. Sarver AL, French AJ, Borralho PM, Thayanithy V, Oberg AL, Silverstein KA, Morlan BW, Riska SM, Boardman LA, Cunningham JM, et al: **Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states**. *BMC Cancer* 2009, **9**:401.
7. Diaz R, Silva J, Garcia JM, Lorenzo Y, Garcia V, Pena C, Rodriguez R, Munoz C, Garcia F, Bonilla F, Dominguez G: **Deregulated expression of miR-106a predicts survival in human colon cancer patients**. *Genes Chromosomes Cancer* 2008, **47**:794-802.
8. Strillacci A, Griffoni C, Sansone P, Paterini P, Piazzini G, Lazzarini G, Spisni E, Pantaleo MA, Biasco G, Tomasi V: **MiR-101 downregulation is involved in cyclooxygenase-2 overexpression in human colon cancer cells**. *Exp Cell Res* 2009, **315**:1439-1447.
9. Wang CJ, Zhou ZG, Wang L, Yang L, Zhou B, Gu J, Chen HY, Sun XF: **Clinicopathological significance of microRNA-31, -143 and -145 expression in colorectal cancer**. *Dis Markers* 2009, **26**:27-34.

10. Bandres E, Cubedo E, Agirre X, Malumbres R, Zarate R, Ramirez N, Abajo A, Navarro A, Moreno I, Monzo M, Garcia-Foncillas J: **Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues.** *Mol Cancer* 2006, **5**:29.
11. Schee K, Fodstad O, Flatmark K: **MicroRNAs as biomarkers in colorectal cancer.** *Am J Pathol* 2010, **177**:1592-1599.
12. Slaby O, Svoboda M, Fabian P, Smerdova T, Knoflickova D, Bednarikova M, Nenutil R, Vyzula R: **Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer.** *Oncology* 2007, **72**:397-402.
13. Ng EK, Chong WW, Jin H, Lam EK, Shin VY, Yu J, Poon TC, Ng SS, Sung JJ: **Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening.** *Gut* 2009, **58**:1375-1381.
14. Aprelikova O, Yu X, Palla J, Wei BR, John S, Yi M, Stephens R, Simpson RM, Risinger JI, Jazaeri A, Niederhuber J: **The role of miR-31 and its target gene SATB2 in cancer-associated fibroblasts.** *Cell Cycle* 2010, **9**:4387-4398.
15. Wang S, Zhou J, Wang XY, Hao JM, Chen JZ, Zhang XM, Jin H, Liu L, Zhang YF, Liu J, et al: **Down-regulated expression of SATB2 is associated with metastasis and poor prognosis in colorectal cancer.** *J Pathol* 2009, **219**:114-122.
16. Cottonham CL, Kaneko S, Xu L: **miR-21 and miR-31 converge on TIAM1 to regulate migration and invasion of colon carcinoma cells.** *J Biol Chem* 2010, **285**:35293-35302.
17. Wu CW, Ng SS, Dong YJ, Ng SC, Leung WW, Lee CW, Wong YN, Chan FK, Yu J, Sung JJ: **Detection of miR-92a and miR-21 in stool samples as potential screening biomarkers for colorectal cancer and polyps.** *Gut* 2012, **61**: 739-45.
18. Ohyashiki JH, Umezu T, Kobayashi C, Hamamura RS, Tanaka M, Kuroda M, Ohyashiki K: **Impact on cell to plasma ratio of miR-92a in patients with acute leukemia: in vivo assessment of cell to plasma ratio of miR-92a.** *BMC Res Notes* 2010, **3**:347.
19. Shigoka M, Tsuchida A, Matsudo T, Nagakawa Y, Saito H, Suzuki Y, Aoki T, Murakami Y, Toyoda H, Kumada T, et al: **Deregulation of miR-92a**

**expression is implicated in hepatocellular carcinoma development.**  
*Pathol Int* 2010, **60**:351-357.

20. Tsuchida A, Ohno S, Wu W, Borjigin N, Fujita K, Aoki T, Ueda S, Takanashi M, Kuroda M: **miR-92 is a key oncogenic component of the miR-17-92 cluster in colon cancer.** *Cancer Sci* 2011, **102**:2264-2271.
21. 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* 2008, **299**:425-436.
22. Akcakaya P, Ekelund S, Kolosenko I, Caramuta S, Ozata DM, Xie H, Lindfors U, Olivecrona H, Lui WO: **miR-185 and miR-133b deregulation is associated with overall survival and metastasis in colorectal cancer.** *Int J Oncol* 2011, **39**:311-318.
23. Motoyama K, Inoue H, Takatsuno Y, Tanaka F, Mimori K, Uetake H, Sugihara K, Mori M: **Over- and under-expressed microRNAs in human colorectal cancer.** *Int J Oncol* 2009, **34**:1069-1075.
24. Xi Y, Formentini A, Chien M, Weir DB, Russo JJ, Ju J, Kornmann M, Ju J: **Prognostic Values of microRNAs in Colorectal Cancer.** *Biomark Insights* 2006, **2**:113-121.
25. Nielsen BS, Jorgensen S, Fog JU, Sokilde R, Christensen IJ, Hansen U, Brunner N, Baker A, Moller S, Nielsen HJ: **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.
26. Bao B, Ali S, Banerjee S, Wang Z, Logna F, Azmi AS, Kong D, Ahmad A, Li Y, Padhye S, Sarkar FH: **Curcumin Analogue CDF Inhibits Pancreatic Tumor Growth by Switching on Suppressor microRNAs and Attenuating EZH2 Expression.** *Cancer Res* 2012, **72**:335-345.
27. Akao Y, Nakagawa Y, Hirata I, Iio A, Itoh T, Kojima K, Nakashima R, Kitade Y, Naoe T: **Role of anti-oncomirs miR-143 and -145 in human colorectal tumors.** *Cancer Gene Ther* 2010, **17**:398-408.
28. Xu Q, Liu LZ, Qian X, Chen Q, Jiang Y, Li D, Lai L, Jiang BH: **MiR-145 directly targets p70S6K1 in cancer cells to inhibit tumor growth and angiogenesis.** *Nucleic Acids Res* 2012, **40**:761-74

29. Ma Y, Zhang P, Yang J, Liu Z, Yang Z, Qin H: **Candidate microRNA biomarkers in human colorectal cancer: Systematic review profiling studies and experimental validation.** *Int J Cancer* 2012, **130**:2077-87
30. Link A, Balaguer F, Shen Y, Nagasaka T, Lozano JJ, Boland CR, Goel A: **Fecal MicroRNAs as novel biomarkers for colon cancer screening.** *Cancer Epidemiol Biomarkers Prev* 2010, **19**:1766-1774.
31. Flatmark K, Pedersen KB, Nesland JM, Rasmussen H, Aamodt G, Mikalsen SO, Bjornland K, Fodstad O, Maelandsmo GM: **Nuclear localization of the metastasis-related protein S100A4 correlates with tumour stage in colorectal cancer.** *J Pathol* 2003, **200**:589-595.
32. Andersen CL, Jensen JL, Orntoft TF: **Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets.** *Cancer Res* 2004, **64**:5245-5250.

## Figure legends

### Figure 1 - MiRNA expression in tumor samples.

Boxplot showing the relative expression distribution of miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145. RT-PCR was performed and Ct values for each miRNA was normalized against RNU44 and the relative expression was calculated using  $2^{-dCt}$  method. Circles represent outliers while stars represent extreme outliers.

### Figure 2 - MiR-31 expression according to pT-stage at diagnosis.

Boxplot showing qRT-PCR relative quantities (using the  $2^{-dCt}$  method) of miR-31 according to pT-stage, indicating that expression of miR-31 increased with increasing pT stage ( $p=0.004$ , Kruskal-Wallis test). Circles represent outliers while stars represent extreme outliers.

### Figure 3 - MiRNA expression and metastasis-free survival

Kaplan-Meier survival plots of metastasis-free survival for the six selected miRNAs. The 193 patients were divided into low and high expression of the respective miRNA based on the median value (low expression  $n=97$  and high expression  $n=96$ ).



## Tables

**Table 1 - Mature sequence, miRBase accession number and proposed clinical relevance for the six chosen miRNAs.**

miRNA name	Mature microRNA sequence	miRBase Accession number	Proposed clinical relevance	Comment	Reference
hsa-miR-21	UAGCUUAUCAGACUGAUGUUGA	MIMAT0000076	Overall survival	High expression associated with poor OS	[21]
hsa-miR-31	AGGCAAGAUGCUGGCAUAGCU	MIMAT0000089	Tumor stage/ differentiation	High expression associated with advanced tumor stage and poorly differentiated tumors	[10]
hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	MIMAT0000092	Plasma marker	Elevated levels as a possible diagnostic marker	[13]
hsa-miR-101	UACAGUACUGUGAUACUGAA	MIMAT0000099	Increased invasiveness	Decreased expression associated with invasiveness	[8]
hsa-miR-106a	AAAAGUGCUUACAGUGCAGGUAG	MIMAT0000103	Disease free and overall survival	Down-regulation associated with poor disease free and overall survival.	[7]
hsa-miR-145	GUCCAGUUUCCAGGAUCCCU	MIMAT0000437	Tumor size	Low expression associated with large tumor size	[9]

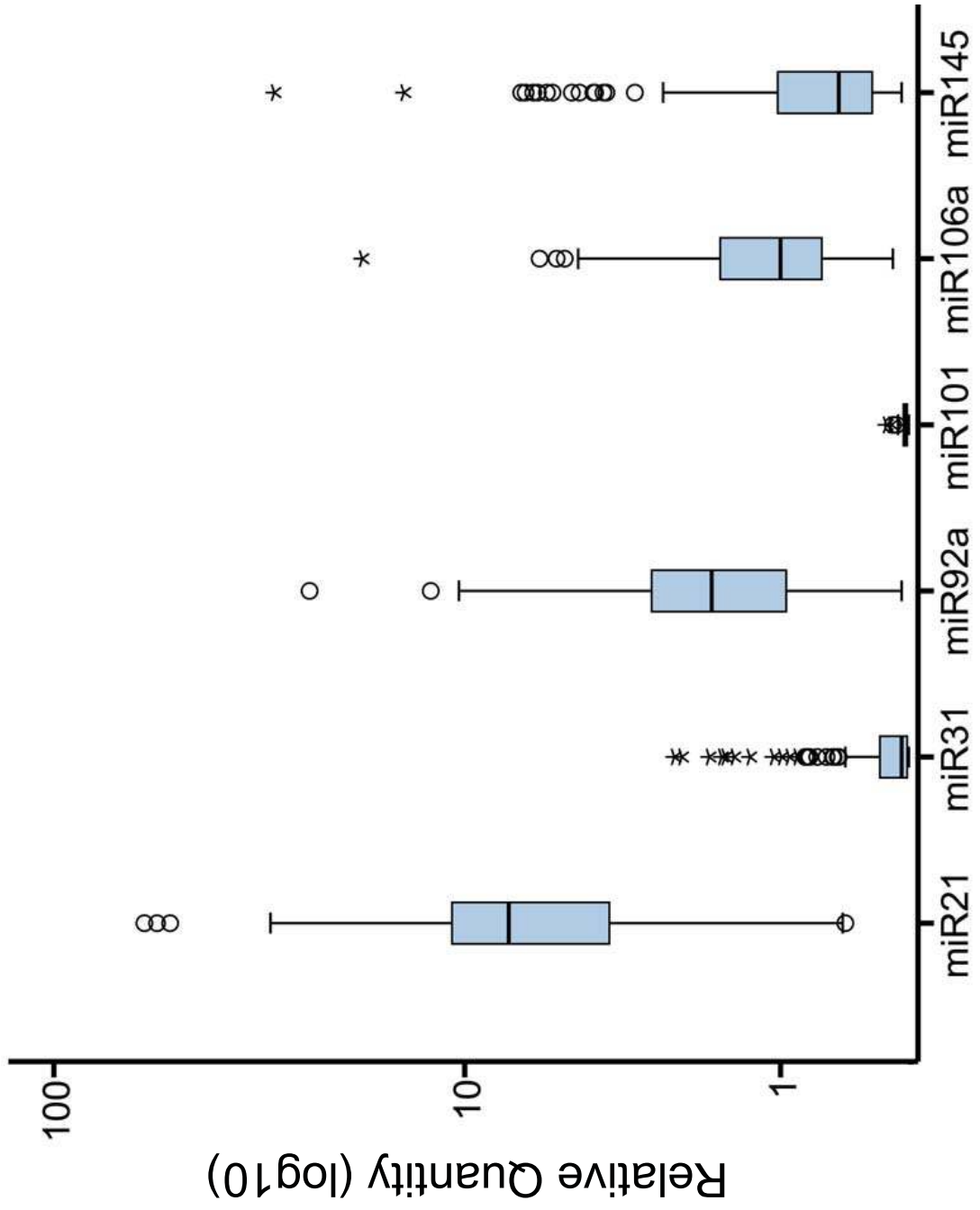
**Table 2 - Median expression levels of the six selected miRNAs and associations with clinicopathological data.**

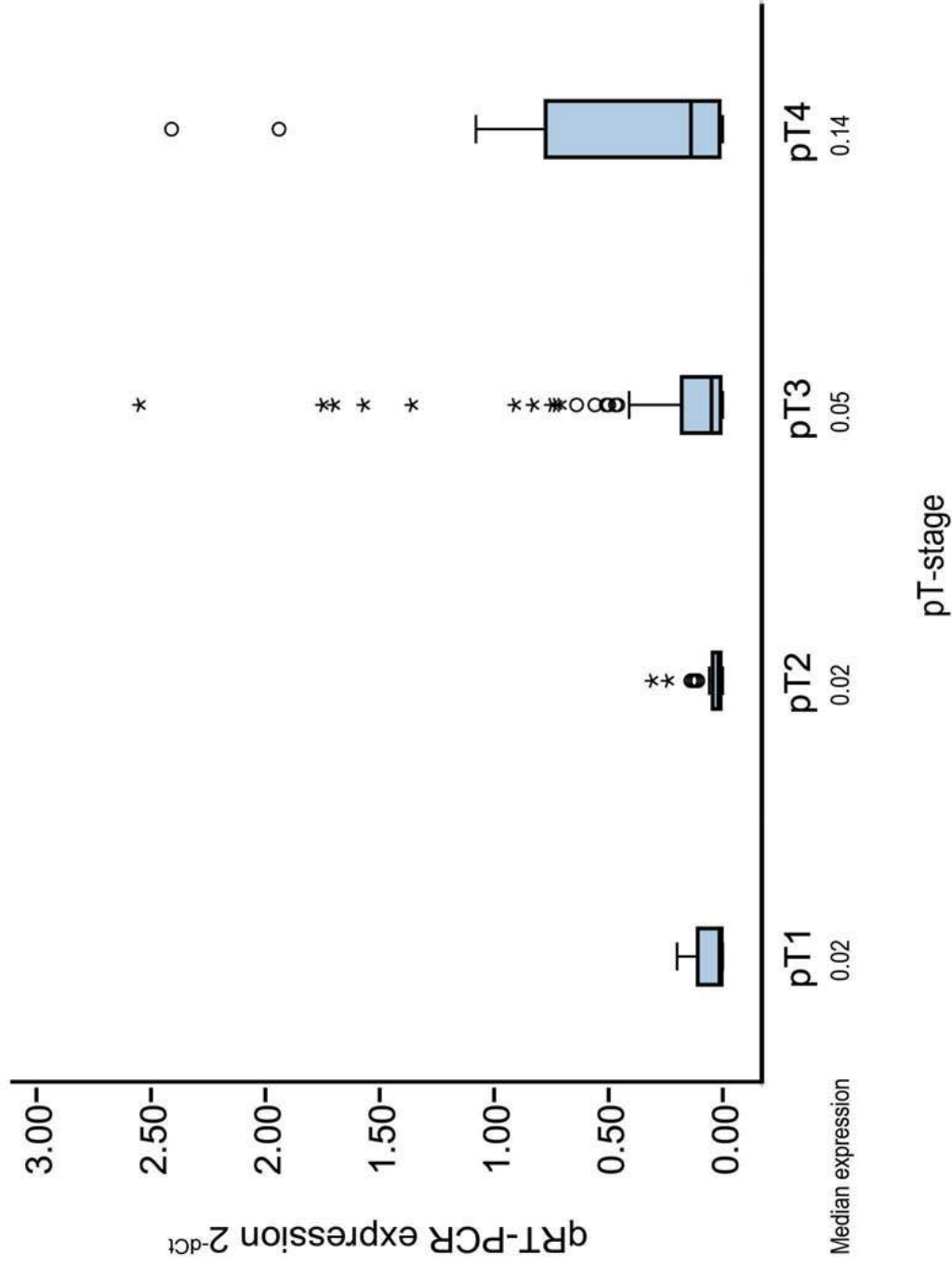
	Number (percent)	miR-21	miR-31	miR-92a	miR-101	miR-106a	miR-145
<b>Gender</b>							
Female	81 (42)	7.78	0.07	1.64	0.02	1.16	0.43
Male	112 (58)	7.53	0.03	2.12	0.02	0.93	0.48
<i>p-value</i>		0.45	0.14	0.52	0.78	0.43	0.99
<b>TNM</b>							
I	35 (18)	5.22	0.02	2.41	0.02	1.24	0.34
II	97 (50)	7.67	0.06	2.09	0.02	1.10	0.48
III	61 (32)	7.78	0.07	1.59	0.02	0.85	0.51
<i>p-value</i>		0.23	0.02	0.80	0.86	0.54	0.30
<b>pT</b>							
1	4 (2)	7.90	0.02	2.51	0.02	1.16	0.36
2	36 (19)	5.15	0.02	2.45	0.02	1.26	0.36
3	133 (69)	7.67	0.05	1.74	0.02	0.88	0.46
4	20 (10)	8.50	0.14	2.58	0.02	1.33	0.58
<i>p-value</i>		0.37	0.004	0.61	0.76	0.52	0.70
<b>pN</b>							
0	132 (68)	7.53	0.04	2.12	0.02	1.15	0.44
1	39 (20)	7.67	0.05	1.73	0.02	0.89	0.46
2	22 (11)	8.47	0.09	1.39	0.02	0.82	0.59
<i>p-value</i>		0.82	0.31	0.60	0.95	0.54	0.63
<b>Differentiation</b>							
Well	6 (3)	3.58	0.02	1.09	0.01	0.38	0.24
Intermediate	167 (87)	7.67	0.04	2.14	0.02	1.16	0.45
Poor	20 (10)	6.83	0.20	0.95	0.02	0.70	0.69
<i>p-value</i>		0.28	0.001	0.003	0.33	0.01	0.12
<b>Tumor localization</b>							
Colon	129 (67)	7.52	0.07	1.62	0.02	0.87	0.43
Rectum	64 (33)	7.77	0.02	2.58	0.02	1.27	0.67
<i>p-value</i>		0.50	0.02	0.05	0.32	0.05	0.08
<b>Lymphocyte infiltration</b>							
High	26 (14)	7.93	0.09	1.96	0.02	0.88	0.63
Intermediate	125 (65)	7.78	0.03	2.14	0.02	1.01	0.51
Low	40 (21)	6.91	0.08	1.46	0.02	1.04	0.33
<i>p-value</i>		0.47	0.19	0.49	0.82	0.37	0.14
<b>Vascular invasion</b>							
Present	38 (20)	7.73	0.07	1.54	0.02	1.17	0.59
Absent	155 (80)	7.36	0.04	2.01	0.02	0.98	0.43
<i>p-value</i>		0.30	0.23	0.43	0.27	0.94	0.05
<b>Perineural invasion</b>							
Present	16 (8)	8.60	0.16	1.68	0.02	1.05	0.35
Absent	177 (92)	7.54	0.04	1.95	0.02	0.98	0.48
<i>p-value</i>		0.42	0.43	0.60	0.29	0.83	0.49
<b>Perinodal growth*</b>							

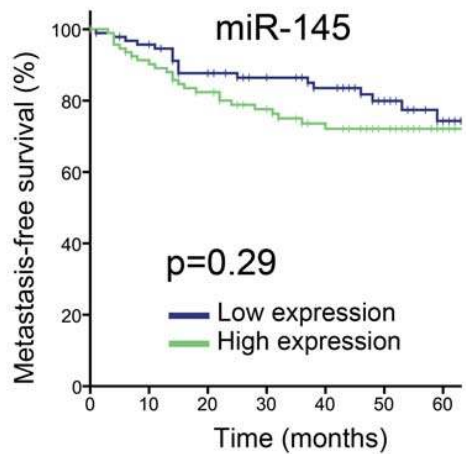
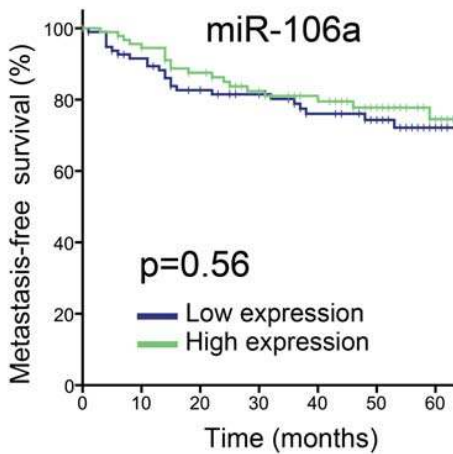
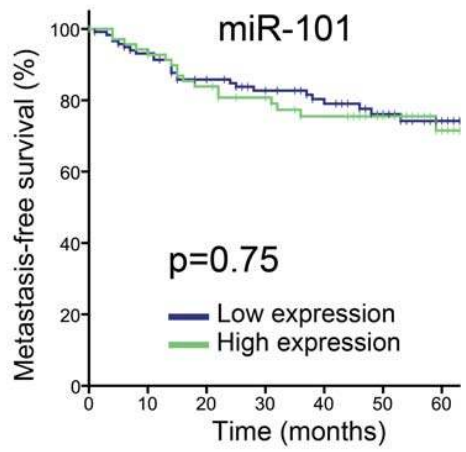
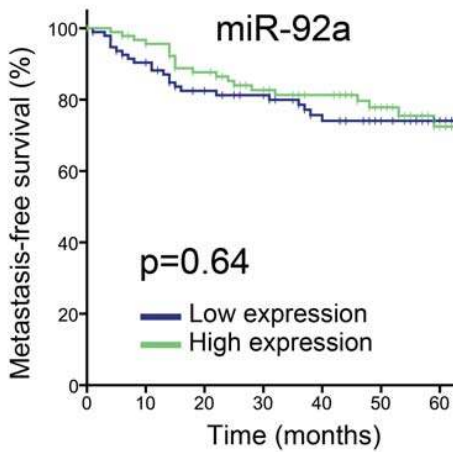
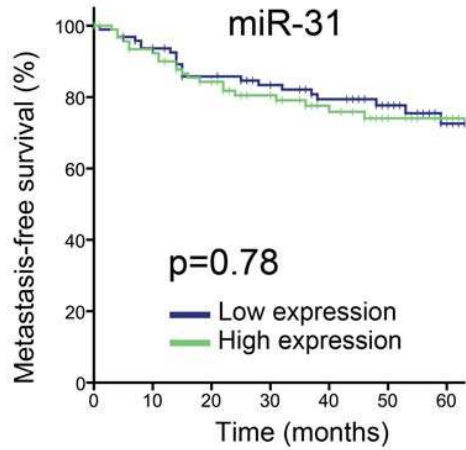
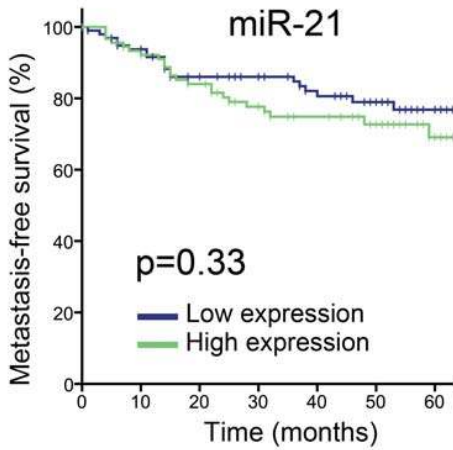
Present	38 (62)	8.09	0.11	1.46	0.02	0.81	0.52
Absent	23 (38)	7.36	0.03	2.01	0.02	1.10	0.36
<i>p-value</i>		<i>0.77</i>	<i>0.36</i>	<i>0.30</i>	<i>1.00</i>	<i>0.18</i>	<i>0.49</i>

\*Associations between miRNA expression and clinicopathological variables were explored using Mann-Whitney or Kruskal-Wallis test as appropriate; p-values are given in italic. Total number of patients included in the analyses was 193, median age was 73 years.

\*\* Perinodal growth was only assessed in the lymph node positive patients







## **PAPER III**

Kristina Schee, Susanne Lorenz, Merete Molton Worren, Clara-Cecilie Günther, Marit Holden,  
Eivind Hovig, Øystein Fodstad, Leonardo Meza-Zepeda, Kjersti Flatmark.

**"Deep sequencing the miRNA transcriptome in colorectal cancer. "**

(Manuscript)





## Deep sequencing the microRNA transcriptome in colorectal cancer

Schee K<sup>1</sup>, Lorenz S<sup>1,2</sup>, Molton Warren M<sup>3</sup>, Günther C-C<sup>4</sup>, Holden M<sup>4</sup>, Hovig E<sup>1,3,5</sup>, Fodstad Ø<sup>1</sup>,  
Meza-Zepeda LA<sup>1,2</sup>, Flatmark K\*<sup>1,6</sup>.

Department of <sup>1</sup>Tumor Biology, and <sup>6</sup> Gastrointestinal Surgery, Clinic for Cancer, Surgery and Transplantation, The Norwegian Radium Hospital, Oslo University Hospital, N-0310 Oslo, Norway. <sup>2</sup> Genomics Core Facility, Institute for Molecular Biosciences, University of Oslo, Oslo, Norway. <sup>3</sup> Institute for informatics, University of Oslo, Oslo, Norway. <sup>4</sup> Norwegian Computing Center, Oslo, Norway, <sup>5</sup> Oslo University Hospital, Section for Medical Informatics, The Norwegian Radium Hospital, P.O. Box 4950, Nydalen, N-0424 Oslo, Norway,

\*Corresponding author

Email: Kjersti Flatmark\*- [kjersti.flatmark@rr-research.no](mailto:kjersti.flatmark@rr-research.no)

This work was supported by grants from South-Eastern Norway Regional Health Authority and from the Norwegian Cancer Society.

## **Abstract**

Colorectal cancer (CRC) is one of the leading causes of cancer related deaths and the search for prognostic biomarkers which might improve treatment decisions is warranted. MicroRNAs are short non-coding RNA molecules involved in regulating gene expression and have been proposed as possible biomarkers in CRC. In order to characterize the miRNA transcriptome, 90 CRC tumors with long-term follow-up were deep sequenced. 523 mature miRNAs were expressed in our cohort, and they exhibited largely uniform expression patterns across tumor samples. Few associations were found between clinical parameters and miRNA expression, among them, increased expression of miR-10b-5p was associated with tumors located in the colon relative to rectum and poorly differentiated tumors, while decreased expression of miR-192-5p was associated with increased pT stage and poorly differentiated tumors, and no prognostic biomarker candidates emerged. Examination of the five most abundantly expressed miRNAs (miR-10a-5p, miR-21-5p, miR-22-3p, miR-143-3p and miR-192-5p) revealed that their collective expression represented 54% of the detected miRNA sequences. Pathway analysis of the top five highly expressed miRNAs uncovered targets within the CRC pathway including Wnt, MAPK and p53 thus suggesting that the top five highly expressed miRNAs are cancer relevant.

## Introduction

MicroRNA (miRNA) are evolutionary conserved small (20-22 nt long), non-coding RNAs that regulate gene expression by binding to the 3'UTR of mRNA, thereby inhibiting translation [1]. They can bind with partial complementarity to mRNA to potentially downregulate several mRNAs. This makes the downstream studies somewhat challenging with multiple potential targets for each miRNA. Today there are approximately 1500 miRNAs annotated in the microRNA database (miRBase) [2] and it is estimated that up to 60% of protein coding genes may be regulated by miRNAs [3]. MiRNAs are essential for normal mammalian development and are involved in fine-tuning many biological processes, such as cell proliferation, differentiation, apoptosis and metabolism, and their involvement in cancer has sparked increased interest in miRNA biology [4-6]. They have been proposed as possible biomarkers because of their regulatory functions, chemical stability and the possibility of measuring miRNA in serum, plasma, stool and tissue samples [7-10].

Colorectal cancer (CRC) is one of the most common cancer forms in Western countries and a leading cause of cancer related deaths. It is a heterogeneous disease characterized by accumulation of genetic and epigenetic events, and influenced by lifestyle [11, 12]. Treatment decisions are still essentially based on the anatomical extent of disease at diagnosis, and the search for better biomarkers is warranted. MiRNAs have been examined for their potential role as prognostic, diagnostic and therapeutic biomarkers in CRC using hybridization based array technologies and quantitative RT-PCR (qRT-PCR) [13-16]. Using expression microarrays,

expression of a large amount of pre-selected miRNAs may be detected, and the miRNA detection is based on signal intensities and is not directly quantifiable. Relative expression can be calculated using qRT-PCR, but when expanding to multiple parallel analyses, the number of miRNAs possible to analyze and RNA quantity may represent limitations. Deep sequencing has emerged as an attractive approach for global miRNA analysis, advantages including pooling of samples for high-throughput purposes, a wide detectable expression range, the ability to analyze expression of all annotated miRNAs and the possibility of detecting novel miRNAs.

In this work, deep sequencing was used to determine miRNA expression in 90 CRC tumor samples, and associations between expression levels and clinicopathological data and outcome were analyzed.

## **MATERIALS AND METHODS**

### ***Patient cohort and sample preparation***

Between the year 1998-2000, 316 patients were recruited from five hospitals in the Oslo region [17], and prospectively included in the study at the time of primary surgery for assumed or verified colorectal cancer. The study was approved by the Regional Ethics Committee (Health Region II, Norway) and informed consent was obtained from the patients. Resected specimens were routinely processed for histopathological assessment at the time of surgery and additional tumor tissue was sampled and snap-frozen in liquid nitrogen. 120 cases were not included in the study for the following reasons: not invasive cancer (25), histology other than adenocarcinoma

(5), distant metastasis at the time of surgery (34), preoperative chemoradiotherapy (2), inadequate surgical margins (7) and unknown stage of disease (1). In addition, frozen tissue samples were not obtainable in 46 cases. From the 196 samples in TNM stage I-III, 90 tumor samples were randomly selected for deep sequencing. After sequencing two samples from the cohort were deemed (degraded/unsuitable) and were removed from further studies leaving a sample cohort of 88 patients (Table 1). Follow-up data was obtained from the participating hospitals and from the general practitioners (for the patients not attending scheduled controls). Metastasis was verified by radiological examination and survival data was obtained from the National Registry of Norway and updated by October 1st 2008.

#### ***RNA isolation and deep sequencing***

RNA was isolated from tumor tissue using TriReagent (Ambion Inc, TX) according to the manufacturer's protocol and the total RNA concentration was measured by Nanodrop (ND-1000). The quality was assessed on an Agilent 2100 Bioanalyzer. Small RNA sequencing libraries were created following the Illumina®TruSeq™ Small RNA Sample Preparation protocol. In brief, 3' and 5' RNA adapter, specifically modified to target the ends of small RNA molecules, were ligated to 1 µg of high quality total RNA. Reverse transcription was performed to generate cDNA libraries and PCR was used to amplify and add unique index sequences to each library. Small RNA libraries were pooled and 32 bases were sequenced for each cDNA molecule using an Illumina® Genome Analyzer IIx. Indexes were sequenced in order to identify the source of each read. Two independent sequencing runs were performed to produce a sufficient amount of reads for each sample. The data for both runs (run 1 and run 2) were combined for downstream

expression analysis, as well as analyzed separately to determine the technical reproducibility of the experiments.

### ***Sequencing data analysis***

Real-time analysis, base calling and filtering of low quality reads were done by Illumina's software packages (SCS2.9/RTA1.9 and Off-line Basecaller v1.9). Novoalign (V2.08.01 Novocraft 2010; [www.novocraft.com](http://www.novocraft.com)) was used to cut remaining adapter sequence and map the reads to the reference human genome (hg19). All reads mapping to 10 or more genomic regions were excluded from further analysis. The mapped reads were annotated using known databases. The miRBase data base release 18 (November 2011) was used to identify miRNAs, the NCBI build "Homo\_sapiens.NCBI.36.58" to identify other small RNA species and mRNA.

To calculate the read count for miRNAs, the reads that mapped uniquely within a mature miRNA sequence with a maximum of one mismatch were considered hits. The reads mapping to more than one mature miRNA sequence were assigned according to the frequency of uniquely mapped reads found for these miRNAs. That means when two miRNAs shared a given number of multiple mapped reads, we identified the ratio of unique reads between these two miRNAs. This ratio was applied to divide the number of multiple mapped reads and assign them. If multiple hits were found to be perfectly mapped to one genomic region and mapped with mismatch to another one, only the perfect matches were considered.

For normalization of read counts, four different approaches were tested. We calculated the normalization factor for all samples by dividing the total number of reads, the number of reads aligned to the genome, allowing multiple hits or that map uniquely, or the number of reads

mapped to annotated mature miRNAs with 1 million. The normalized expression values for each miRNA were generated by dividing the read count of the miRNA with the according normalization factor. The normalization methods gave very similar results when comparing the mean change calculated by the difference in percent for each miRNA readcount per miRNA (data not shown).

### ***Quantitative Real Time-PCR***

Total RNA from 88 patients was used to reversely transcribe miRNAs using TaqMan microRNA assays (Applied Biosystems, Foster City, CA) as described in the protocol. All samples were run in duplicates. Ct values for miRNAs were normalized against RNU44 and the relative expression was calculated using  $2^{-\Delta\Delta Ct}$  method [18].

### ***Statistical analysis***

Normalized sequence data was log2-transformed and filtered by excluding miRNAs for which all individuals had read counts below 20 (before normalization). Overall and metastasis-free survival was calculated from date of surgery until date of death or diagnosis of metastasis. To identify miRNAs associated with overall and metastasis-free survival univariate Cox proportional hazard regression was applied to each miRNA, testing for associations with metastasis-free or overall survival. Then, for all the miRNAs simultaneously, the LASSO method in the Cox proportional hazards model [19], as implemented previously [20], was used to discover a set of miRNAs associated with the endpoints. A p-value less than 0.01 in the univariate Cox analysis or

a coefficient different from 0 in the LASSO analysis were necessary to consider a miRNA as significantly associated with an outcome parameter. To account for multiple testing, adjusted p-values were calculated by controlling the false discovery rate (FDR), using the Benjamini-Hochberg procedure [21].

For individual miRNAs, survival was additionally estimated using the Kaplan-Meier method and compared using the log-rank test. Data was arranged into high and low values based on the median gene expression, and associations with overall and metastasis-free survival were examined. Associations between miRNA expression values and clinicopathological variables, including TNM, pT, pN, lymphocyte infiltration, vascular invasion, neural invasion, differentiation, tumor localization and perinodal infiltration were tested using two-tailed Fisher's exact test or linear-by-linear association chi-square test. These analyses were performed using SPSS version 18.0 (SPSS Inc., Chicago, MO) and p-values < 0.05 were considered to be statistically significant.

### ***Hierarchical clustering and SAM analysis***

Hierarchical clustering was performed to visualize expression patterns of all miRNAs and significance analysis of microarrays (SAM) [22] was used to identify miRNAs associated with clinicopathological parameters, using J-Express (2012 version) [23]. The normalized expression values were log<sub>2</sub> transformed and unsupervised two-way hierarchical clustering was performed using Euclidean distance and weighted average linkage (WPGMA) to cluster miRNAs and



samples simultaneously. The input for SAM was also normalized and log<sub>2</sub> transformed expression values and clinicopathological parameters were used as response variables. Ten thousand repeat permutations of the data were used to determine if the expression of any miRNAs were significantly related with one of the parameters. The following clinicopathological parameters were analyzed: TNM, pT, pN, lymphocyte infiltration, vascular invasion, neural invasion, differentiation, tumor localization and perinodal infiltration. The false discovery rate expressed as q-values of 0.05 was used to evaluate statistical significance.

## **RESULTS**

### ***Small RNA sequencing and annotation***

The length of the detected sequences varied between 13 and 29 nucleotides after removal of the adapter sequence. The main portion of reads, 97.9%, were between 19 and 23 bases (Supplementary; size distribution plot). In average, 2.6 million reads mapping to the human genome were obtained per tumor sample. We identified the frequencies of reads falling into different classes of small RNA or other genomic regions and calculated the median frequencies comparing all 88 tumor samples (Figure 1). The frequency of reads mapping to miRNAs ranged from 37 to 77% in the libraries and gave a median of 61% (Figure 1B). For intronic/intergenic regions a median read frequency of 33% was found (Figure 1D), for premature miRNAs and snoRNAs the median read frequencies were 4% and 2%, respectively (Figure 1C). In addition a small fraction of reads mapped to snRNA, miscRNA, tRNA, rRNA, and mRNA together comprising a frequency of ~0.05% (Figure 1E). MiRNAs with less than 10 reads across all patient samples were considered not expressed. In total, 523 miRNAs were expressed in the data set.

### ***Normalization and technical replicates***

In general, to compare expression values between samples it is necessary to normalize against the read count by calculating a normalization factor. 48 samples were deep sequenced twice, denoted run 1 and run 2, and these data sets were used to compare the expression from the separate runs, and also to evaluate the impact of the different normalization methods.

Assuming a good reproducibility of the data in run1 and run2, normalized read counts should be very similar for each miRNA. Figure 2 shows the distribution of normalized and non-normalized miRNA read count ("Readcounts") for run1 compared to run2. All tested normalization methods generated quite similar normalized read counts for the majority of miRNAs when comparing the two runs. The results also demonstrated a very good correlation between the technical replicates. The normalization using the read counts mapped to annotated mature miRNAs was used for further evaluation.

### ***The five most highly expressed miRNAs***

The five most abundantly expressed miRNAs in this cohort were miR-10a-5p, miR-21-5p, miR-22-3p, miR-143-3p and miR-192-5p (Figure 3). The read counts for these miRNAs accounted for 53.6% of the total number of miRNA sequences detected in the patient samples, while the top 20 miRNAs accounted for 80.3% of the reads. The remaining 503 miRNAs represented only 19.3% of the reads. Of the top five miRNA, miR-192-5p had the highest median expression

(156,413 reads) while miR-22-3p had the lowest (35,284 reads). The sum of all reads in the top five varied with 10 mill reads from the highest, miR-192-5p with 13,338,186 reads, to the lowest, miR-22-3p with 3,052,128 reads.

Many miRNA genes are located in close proximity to other miRNA genes in gene clusters, and two of the top five most highly expressed miRNAs are part of such clusters, and expression levels of miRNAs belonging to these clusters are depicted in Figure 4. No co-expression was apparent for the miRNAs belonging to these gene clusters, which is in concordance with previous results [24].

#### ***Pathway analysis for the five most highly expressed miRNAs***

To investigate the biological influence of the most highly expressed miRNAs, target genes were identified using TarBase6.0 [25]. This database contains target genes experimentally verified in addition to sequence-based target predictions. In total, 1490 target genes were identified as potentially regulated by miR-10a-5p, miR-21-5p, miR-22-3p, miR-143-3p and miR-192-5p. The gene identities were uploaded into the web-based DAVID functional annotation tool for pathway analysis using the KEGG database [26, 27]. The pathway analysis showed that a high number of the identified genes are involved in cancer related pathways, specifically the “Colorectal Cancer Pathway”, including the Wnt, MAPK and ErbB signaling networks. The top ten pathways with the most significant gene-enrichment are shown in Table 2. Focusing on the CRC pathway we found that genes regulated by the top five most highly expressed miRNAs

included oncogenes (K-Ras), tumor suppressors (APC and TGF $\beta$ RII), and DNA repair genes (hMSH6) (Figure 5).

### ***Hierarchical clustering***

The miRNA expression patterns observed with hierarchical clustering are shown in Figure 6 with miRNAs are on the vertical axis and patient samples on the horizontal axis. Most of the miRNAs exhibited very similar expression levels among patient samples. In areas of the plot, some miRNAs appeared to be differentially expressed, but these were almost exclusively located among the miRNA that had very low expression.

### ***Correlation between qRT-PCR and deep sequencing data***

Expression values for six miRNA miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145, was previously determined using qRT-PCR (Schee et al, under revision BMC cancer). The miRNA expression measured by qRT-PCR was compared to the deep sequencing data using linear regression analysis of normalized Ct values (qRT-PCR) and log<sub>2</sub>-transformed deep sequencing data. The R<sup>2</sup> values for the 6 miRNAs tested were 0.06, 0.38, 0.10, 0.001, 0.03, and 0.28 for miR-21, miR-31, miR-92a, miR-101, miR-106a, and miR-145, respectively. The sum of total expression levels was also calculated for these miRNA for each method, and the relative levels are shown in Figure 7. The relative expression between individual miRNAs was reasonably consistent between methods for four of the miRNAs (miR-21, miR-31, miR-92a and miR-106a), whereas

there were clear discrepancies for miR-101 and miR-145. MiR-101 was hardly detectable with qRT-PCR, but exhibited detectable expression values with deep sequencing, while miR-145 was detected by qRT-PCR and hardly detected using deep sequencing.

#### ***Associations between miRNA and clinicopathological parameters***

SAM analysis of the expression data and the clinicopathological parameters revealed that high expression of miR-10b-5p was associated with tumors located in the colon ( $q < 0.001$ ) compared with tumors located in the rectum (Figure 8). The expression of this miRNA showed a twofold increase ( $\log_2$  fold change = 2.1) for tumors in the colon with median normalized read counts of 34193 and 15395 in colon and rectum, respectively. MiR-10b-5p was also associated with poorly differentiated tumors ( $p = 0.007$ ), but this value did not retain significance when correcting for multiple testing. Analyzing associations between the top five most highly expressed miRNA and clinicopathological parameter using linear-by-linear chi-square test showed that low expression of miR-192-5p was associated with high pT stage ( $p = 0.007$ ) and poorly differentiated tumors ( $p = 0.034$ ) (Table 3).

#### ***Associations between miRNA expression and outcome***

In the LASSO and univariate Cox analysis, with metastasis development as endpoint, 5 miRNAs (miR-339-5p, miR-7-1-3p, miR-365b-3p, miR-454-3p, miR-194-3p and miR-15b-3p) emerged, although none of these remained significantly associated with metastasis development after adjusting for multiple testing. When overall survival was used as the endpoint in the LASSO and

univariate Cox analysis, only one miRNA (miR-101-5p) was identified. The miRNAs obtained from the LASSO and Cox analysis with endpoints for both metastasis and overall-free survival were in addition analyzed with Kaplan-Meier plots and none were found to have a p-value <0.05.

## **DISCUSSION**

In the present work, deep sequencing was used to quantify miRNA expression in CRC tumor samples. This approach may contribute potential advantages in global miRNA expression analysis, but also entails new challenges regarding data analysis, as the amount of data collected after deep sequencing contains millions of reads which need to be mapped to genome and normalized [28]. In the 88 CRC patients successfully analyzed, 523 mature miRNAs were detected. Other small RNA sequences were also detected, but the low detection frequencies of other RNA classes and genomic regions showed that selection for miRNAs had been successful, and in accordance with previous results [24, 29]. In addition, the excellent agreement observed between technical replicates suggested adequate reproducibility the deep sequencing results.

The five miRNAs most abundantly expressed in the examined CRC cohort were miR-10a-5p, miR-21-5p, miR-22-3p, miR-143-3p and miR-192-5p, and all of these have previously have been shown to be dysregulated in CRC [30-36]. These miRNAs were also among the most highly expressed miRNAs in a previous study performed with deep sequencing of CRC [37]. Interestingly, the top five most highly expressed miRNAs accounted for as much as 54% of the

total number of miRNA sequences detected. This is in concordance with a previous deep sequencing study performed on peripheral blood samples, in which the let-7 family accounted for 77% of the total miRNA read counts [24]. The relative importance of high versus low miRNA expression is difficult to interpret, since the absence or abundant presence of miRNAs may represent equally important biological regulatory signals. However, overrepresentation of a small number of miRNAs may imply that these miRNA play important roles as negative regulators of downstream targets and the biological pathways affected by these targets. The predicted targets of the top 5 most highly expressed miRNAs were associated with cancer-relevant pathways, including the CRC pathway. Among the predicted targets in the CRC pathway were oncogenes, tumor suppressors and DNA repair genes which are involved in several important signaling pathways including Wnt, MAPK, cell cycle, TGF- $\beta$ , and p53. Predicted targets were also involved with downregulating DNA repair genes affecting microsatellites and are thereby involved in microsatellite instability pathway. These results suggest that the identified top five most highly expressed miRNA are cancer relevant, and probably relevant in CRC but further investigation is necessary to validate the targets is needed and to assess downstream effects.

When comparing miRNA expression levels between samples, the most apparent finding was the low variability, clearly illustrated in Figure 6, showing the outcome of hierarchical clustering. Several specific miRNAs have previously been reported to be differentially expressed in CRC. In most cases these results were obtained using expression microarrays or RT-PCR, and often, normalization against normal tissue expression was performed [16, 38, 39]. One of the

supposed advantages of deep sequencing compared to microarray analysis and qRT-PCR is improved specificity and sensitivity, suggesting that this method would return more correct measurements of each miRNA than the other methods. When comparing our previous results, measuring expression of six miRNA using qRT-PCR with deep sequencing data, correlation on the individual sample level was poor for all the miRNAs examined. Deep-sequencing is often validated by qRT-PCR, but comprehensive comparisons between the two approaches have not been performed. In a deep sequencing study on 9 bladder cancer samples, selected miRNAs from deep sequencing were validated by qRT-PCR, and reported to correlate well when analyzed by fold expression differences in a bar plot [40]. In another study performed on 10 neuroblastoma samples, correlation coefficients between 0.1 and 1 were found when comparing the sum of miRNA expression from RT-PCR and deep sequencing [41]. Thus, it seems unclear whether qRT-PCR can be used for validation purposes, since qRT-PCR and deep sequencing data are generated with different methods and appear on different scales with variable expression ranges, making it difficult to compare the two datasets on a patient-to-patient basis.

One of the aims of this study was to investigate associations between miRNA expression and clinicopathological parameters and outcome. Given the low variability observed between samples, it is not surprising that very few such associations were detected. Using the Cox regression and LASSO analysis, only a small number of miRNA were found to be associated with metastasis development and overall survival, and none of these remained significant after correction for multiple testing. Increased expression of miR-10b was associated with tumors



located in the colon relative to the rectum and with poorly differentiated tumors. MiR-10b has previously been reported to be downregulated in CRC and increased expression was associated with increased pT-stage [42]. We did, however, not observe association with pT stage. Among the five most highly expressed miRNAs, low expression of miR-192 was associated with high pT stage and poor tumor differentiation. MiR-192 has previously been reported to be downregulated in CRC compared to normal mucosa [43, 44] and increased expression has been associated with inhibition of cellular proliferation in colon cancer cell lines containing wild-type p53 [31]. The significant associations with tumor differentiation and pT stage at diagnosis may thus represent biologically relevant findings, although no associations were found with metastasis-free or overall survival.

Differential expression analysis of miRNAs using deep sequencing has previously been performed in a small cohort of 8 CRC samples [37]. In this study, 37 miRNAs were dysregulated relative to corresponding normal tissue (19 downregulated and 18 upregulated), and 16 miRNA had not previously been reported to be associated with CRC. Previous studies of miRNA expression in CRC have reported quite variable results, both with respect to expression levels and associations with clinicopathological parameters and outcome [39]. Whether these discrepancies result from variations in the methods used, differences between the clinical cohorts analyzed or biological variance are unclear. A contributing explanation to the observed discrepancies may be that many investigators have normalized tumor miRNA expression values against normal samples, using either paired samples or a mixed “normal cohort”, in order to compare samples. The colorectum is a heterogeneously composed organ, consisting of multiple

cell types which all may be affected by genetic predispositions and external influences. MiRNA expression in different cell types in the normal colorectum has not been characterized and in addition, the miRNA transcriptome from different “normal” samples have not been extensively compared. The few studies which show miRNA expression in normal colon tissues reveal that the expression is highly variable and in many cases not consistent between individuals [45]. Exploration of miRNA expression in normal tissue and establishing a baseline is needed to establish the relevance of using this normalization strategy.

In the present work, deep sequencing was performed to characterize the miRNA transcriptome of CRC, using tumor samples from a large patient cohort with long-term follow-up. Deep sequencing was technically successful, and a total of 523 mature miRNA were expressed in the samples. Most of the miRNAs exhibited relatively uniform expression between tumor samples, and only few associations were found between expression of specific miRNA and clinical parameters. Specifically, no miRNA emerged as a prognostic biomarker candidate, which in our opinion is in agreement with the highly variable results obtained in similar studies in CRC. The five most highly expressed miRNAs, representing 54% of the detected miRNA sequences, have been predicted to regulate targets involved in cancer pathways, and may represent interesting candidates for future studies of the role of miRNAs in CRC development and progression.

## FIGURE LEGENDS

**Figure 1. Frequencies of different RNA classes and genomic regions present in the small RNA libraries.** **A** Number of reads that mapped to human genome (hg19) for each sample. **B** and **C** Frequencies of reads mapping into annotated mature miRNAs (**B**) and premature miRNAs (**C**) for the different samples using the microRNA database (miRBase release 18). The fractions of reads that did not map to any annotated sequence are illustrated in **D** (intronic/intergenic regions). An overview about the median frequencies of all different RNA classes found in the data set is given in **E**.

**Figure 2. Comparison of the normalization methods and the technical replicates.** We tested four different methods to calculate the normalization factor for each sample using either the total number of reads (green), the number of reads mapped to the genome (blue), mapped to annotated miRNAs (light blue) or number of reads only mapped uniquely (1 place) in the genome (purple). Un-normalized read counts are shown in black.

**Figure 3. Boxplot of the top five most highly expressed miRNAs.**

The five most abundantly expressed miRNAs in our CRC cohort, total number of reads in log<sub>2</sub> scale. The circles represent outliers and the stars represent extreme outliers.

**Figure 4. Two of the top five highly expressed miRNAs shown with miRNAs from their respective gene clusters.** **A.** MiR-192-5p is more highly expressed than miR-1994-5p which is

found in the same gene cluster. **B.** MiR-143-39 is found in the same gene cluster as miR-145-5p, but is more highly expressed.

**Figure 5. The colorectal cancer pathway.** Pathway analysis of the top five highly expressed miRNAs and which targets they inhibit. Illustration is taken from KEGG.

**Figure 6. Heat maps of the expression levels of 88 colorectal cancer samples.** The normalized expression values were log<sub>2</sub> transformed and analyzed by hierarchical clustering using weighted average linkage (WPGMA). The global map contains all expressed miRNAs shown vertically and all the patient samples horizontally.

**Figure 7. Comparison between deep sequencing data and qRT-PCR.** Sum of the total expression of each miRNA from deep sequencing and PCR (n= 88 CRC tumors) for the following miRNAs; miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145.

**Figure 8. Boxplot of miR-10b expression (log<sub>2</sub>) between colon and rectum.** Increased expression of miR-10b was associated with tumor localized in the colon relative to the rectum (p < 0.01).

## TABLES

**Table 1.** Clinical and histopathological characteristics of the investigated patient cohort (n=88).

<b>Sex</b>	Male	36
	Female	52
<b>TNM</b>	1	10
	2	51
	3	27
<b>pT</b>	1	2
	2	10
	3	70
	4	6
<b>pN</b>	0	61
	1	16
	2	11
<b>Tumor localization</b>	Colon	63
	Rectum	26
<b>Differentiation</b>	Poor	10
	Intermediate	76
	Well	2
<b>Perinodal infiltration</b>	No	70
	Yes	18
<b>Vascular invasion</b>	No	69
	Yes	19
<b>Neural infiltration</b>	No	82
	yes	6

**Table 2. Pathway analysis of the predicted targets of the top ten highly expressed miRNAs.**

The number of proteins present in their representative KEGG pathway with respective p-values.

<b>Term</b>	<b>Number of proteins</b>	<b>p-value</b>
Pathways in cancer	55	$3.3 \times 10^{-7}$
Cell cycle	28	$2.2 \times 10^{-6}$
Colorectal cancer	21	$1.0 \times 10^{-5}$
Pancreatic cancer	19	$1.4 \times 10^{-5}$
Prostate cancer	20	$8.4 \times 10^{-5}$
Bladder cancer	12	$4.3 \times 10^{-4}$
Wnt signaling pathway	26	$5.1 \times 10^{-4}$
Chronic myeloid leukemia	16	$9.7 \times 10^{-4}$
TGF-beta signaling pathway	17	$1.7 \times 10^{-3}$
ErbB signaling pathway	17	$1.7 \times 10^{-3}$

**Table 3. Associations between miRNA expression and clinicopathological parameters.**

<b>Parameter</b>		<b>Total nr</b>	<b>miR-192-5p median expression (range)</b>	<b>miR-10b-5p median expression (range)</b>
<b>Differentiation</b>	Poor	10	81392 (7188, 219191)	47179 (35309, 65812)
	Intermediate	76	158384 (7120, 325277)	20627 (3998, 203299)
	Well	2	206945 (174026, 239864)	25680 (19396, 31964)
<b>pT</b>				-
	1	2	256044 (211076, 301012)	-
	2	10	188341 (59803, 307865)	-
	3	70	153743 (7120, 325277)	-
<b>Tumor localization</b>	4	6	17813 (7187, 162816)	-
	Colon	63	-	34193 (6518, 203299)
	Rectum	25	-	15396 (3998, 43504)

## REFERENCES

1. Kusenda, B., et al., *MicroRNA biogenesis, functionality and cancer relevance*. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub, 2006. **150**(2): p. 205-15.
2. Griffiths-Jones, S., et al., *miRBase: microRNA sequences, targets and gene nomenclature*. Nucleic Acids Res, 2006. **34**(Database issue): p. D140-4.
3. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs*. Genome Res, 2009. **19**(1): p. 92-105.
4. Tsuchiya, S., Y. Okuno, and G. Tsujimoto, *MicroRNA: biogenetic and functional mechanisms and involvements in cell differentiation and cancer*. J Pharmacol Sci, 2006. **101**(4): p. 267-70.
5. Tang, J.T. and J.Y. Fang, *MicroRNA regulatory network in human colorectal cancer*. Mini Rev Med Chem, 2009. **9**(8): p. 921-6.
6. Sassen, S., E.A. Miska, and C. Caldas, *MicroRNA: implications for cancer*. Virchows Arch, 2008. **452**(1): p. 1-10.
7. Chen, X., et al., *Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases*. Cell Res, 2008. **18**(10): p. 997-1006.
8. Huang, Z., et al., *Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer*. Int J Cancer, 2010. **127**(1): p. 118-26.
9. Link, A., et al., *Fecal MicroRNAs as novel biomarkers for colon cancer screening*. Cancer Epidemiol Biomarkers Prev, 2010. **19**(7): p. 1766-74.
10. Ju, J., *miRNAs as biomarkers in colorectal cancer diagnosis and prognosis*. Bioanalysis, 2010. **2**(5): p. 901-6.
11. Giovannucci, E., *Diet, body weight, and colorectal cancer: a summary of the epidemiologic evidence*. J Womens Health (Larchmt), 2003. **12**(2): p. 173-82.
12. Migliore, L., et al., *Genetics, cytogenetics, and epigenetics of colorectal cancer*. J Biomed Biotechnol. **2011**: p. 792362.
13. Lin, M., et al., *MicroRNA expression profiles in human colorectal cancers with liver metastases*. Oncol Rep, 2011. **25**(3): p. 739-47.
14. Mori, Y., et al., *Novel candidate colorectal cancer biomarkers identified by methylation microarray-based scanning*. Endocr Relat Cancer. 2011. **18**(4): p. 465-78.
15. Mosakhani, N., et al., *MicroRNA profiling differentiates colorectal cancer according to KRAS status*. Genes Chromosomes Cancer, 2012. **51**(1): p. 1-9.
16. Bandres, E., et al., *Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues*. Mol Cancer, 2006. **5**: p. 29.
17. Flatmark, K., et al., *Disseminated tumour cells as a prognostic biomarker in colorectal cancer*. Br J Cancer, 2011. **104**(9): p. 1434-9.
18. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) Method*. Methods, 2001. **25**(4): p. 402-8.
19. Tibshirani, R., *The lasso method for variable selection in the Cox model*. Stat Med, 1997. **16**(4): p. 385-95.
20. Bovelstad, H.M., et al., *Predicting survival from microarray data--a comparative study*. Bioinformatics, 2007. **23**(16): p. 2080-7.
21. Benjamini, Y. and R. Heller, *Screening for partial conjunction hypotheses*. Biometrics, 2008. **64**(4): p. 1215-22.
22. Wu, Y., et al., *Identification of differentially expressed genes using multi-resolution wavelet transformation analysis combined with SAM*. Gene. 2012. **2**: p.302-8.
23. Dysvik, B. and I. Jonassen, *J-Express: exploring gene expression data using Java*. Bioinformatics, 2001. **17**(4): p. 369-70.

24. Vaz, C., et al., *Analysis of microRNA transcriptome by deep sequencing of small RNA libraries of peripheral blood*. BMC Genomics, 2010. **11**: p. 288.
25. Vergoulis, T., et al., *TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support*. Nucleic Acids Res. 2012. **40**(Database issue): p. D222-9.
26. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. Nat Protoc, 2009. **4**(1): p. 44-57.
27. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists*. Nucleic Acids Res, 2009. **37**(1): p. 1-13.
28. Creighton, C.J., J.G. Reid, and P.H. Gunaratne, *Expression profiling of microRNAs by deep sequencing*. Brief Bioinform, 2009. **10**(5): p. 490-7.
29. Morin, R.D., et al., *Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells*. Genome Res, 2008. **18**(4): p. 610-21.
30. Han, L., et al., *DNA methylation regulates MicroRNA expression*. Cancer Biol Ther, 2007. **6**(8): p. 1284-8.
31. Song, B., et al., *miR-192 Regulates dihydrofolate reductase and cellular proliferation through the p53-microRNA circuit*. Clin Cancer Res, 2008. **14**(24): p. 8080-6.
32. Akao, Y., et al., *Role of anti-oncomirs miR-143 and -145 in human colorectal tumors*. Cancer Gene Ther, 2010. **17**(6): p. 398-408.
33. Chen, X., et al., *Role of miR-143 targeting KRAS in colorectal tumorigenesis*. Oncogene, 2009. **28**(10): p. 1385-92.
34. Gao, W., et al., *Deregulated expression of miR-21, miR-143 and miR-181a in non small cell lung cancer is related to clinicopathologic characteristics or patient prognosis*. Biomed Pharmacother, 2010. **64**(6): p. 399-408.
35. Kulda, V., et al., *Relevance of miR-21 and miR-143 expression in tissue samples of colorectal carcinoma and its liver metastases*. Cancer Genet Cytogenet, 2010. **200**(2): p. 154-60.
36. Ma, Y., et al., *Candidate microRNA biomarkers in human colorectal cancer: systematic review profiling studies and experimental validation*. Int J Cancer, 2012. **130**(9): p. 2077-87.
37. Hamfjord, J., et al., *Differential expression of miRNAs in colorectal cancer: comparison of paired tumor tissue and adjacent normal mucosa using high-throughput sequencing*. PLoS One, 2012. **7**(4): p. e34150.
38. Schepeler, T., et al., *Diagnostic and prognostic microRNAs in stage II colon cancer*. Cancer Res, 2008. **68**(15): p. 6416-24.
39. Xi, Y., et al., *Prognostic Values of microRNAs in Colorectal Cancer*. Biomark Insights, 2006. **2**: p. 113-121.
40. Han, Y., et al., *MicroRNA expression signatures of bladder cancer revealed by deep sequencing*. PLoS One, 2011. **6**(3): p. e18286.
41. Schulte, J.H., et al., *Deep sequencing reveals differential expression of microRNAs in favorable versus unfavorable neuroblastoma*. Nucleic Acids Res, 2010. **38**(17): p. 5919-28.
42. Chang, K.H., et al., *MicroRNA signature analysis in colorectal cancer: identification of expression profiles in stage II tumors associated with aggressive disease*. Int J Colorectal Dis. 2011. **26**(11): p. 1415-22.
43. Earle, J.S., et al., *Association of microRNA expression with microsatellite instability status in colorectal adenocarcinoma*. J Mol Diagn. 2010. **12**(4): p. 433-40.
44. Karaayvaz, M., et al., *Prognostic significance of miR-215 in colon cancer*. Clin Colorectal Cancer. 2011. **10**(4): p. 340-7.
45. Ng, E.K., et al., *MicroRNA-143 targets DNA methyltransferases 3A in colorectal cancer*. Br J Cancer, 2009. **101**(4): p. 699-706.



Figure 1

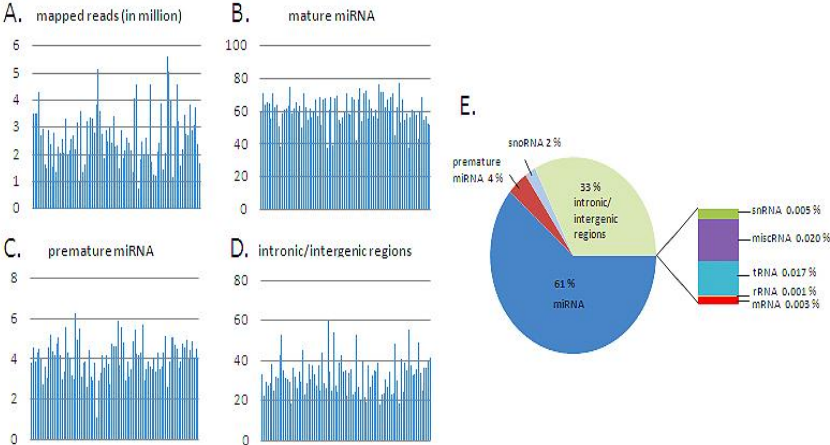


Figure 2

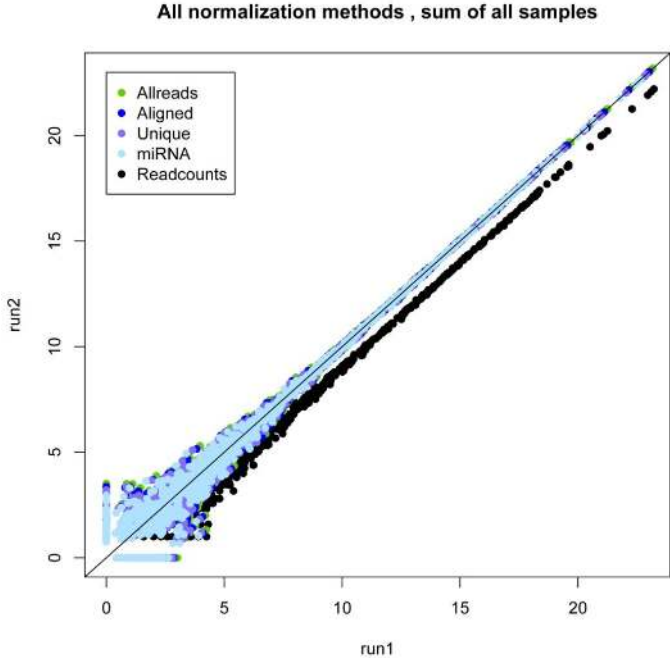


Figure 3



Figure 5

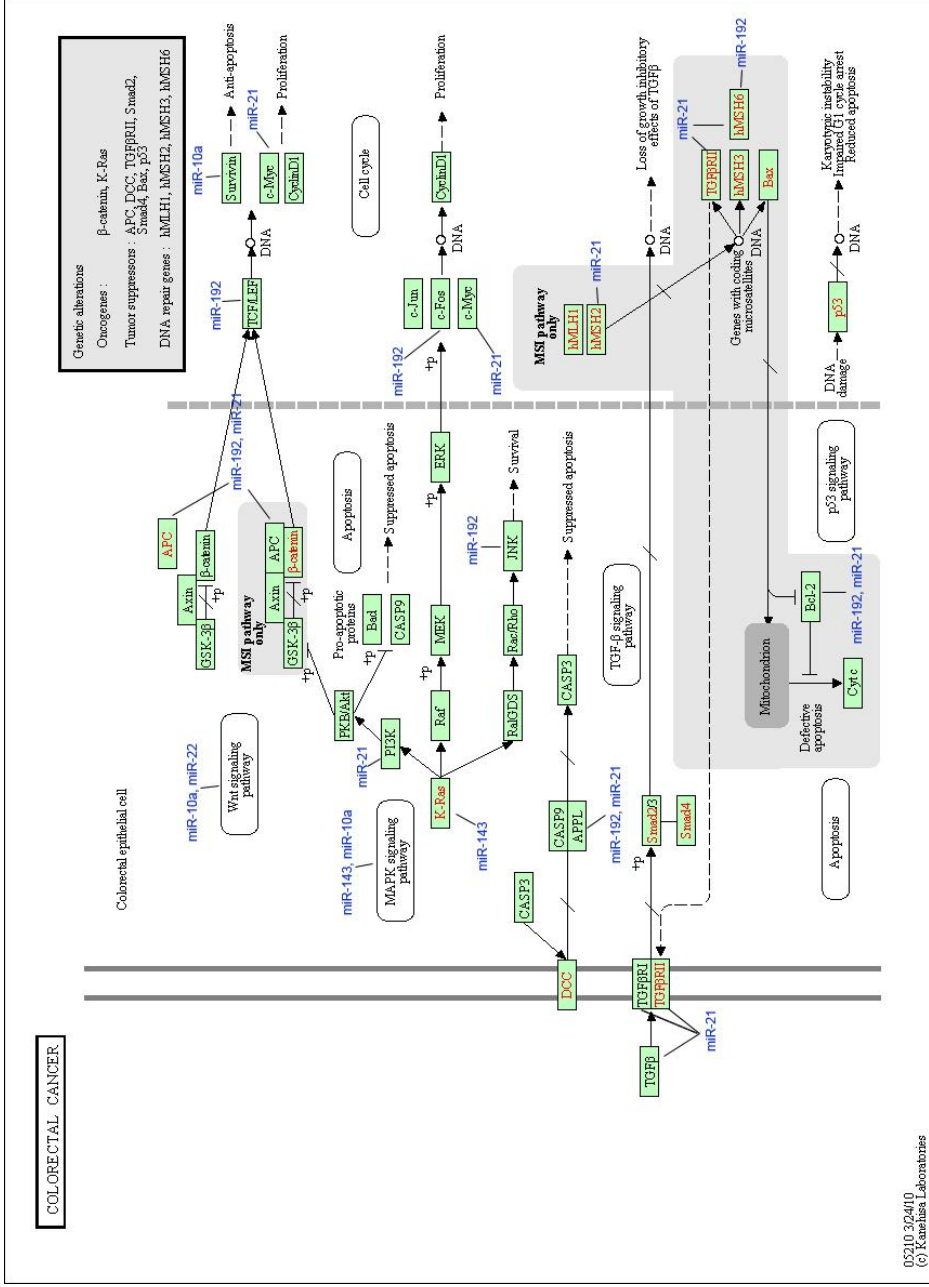


Figure 6

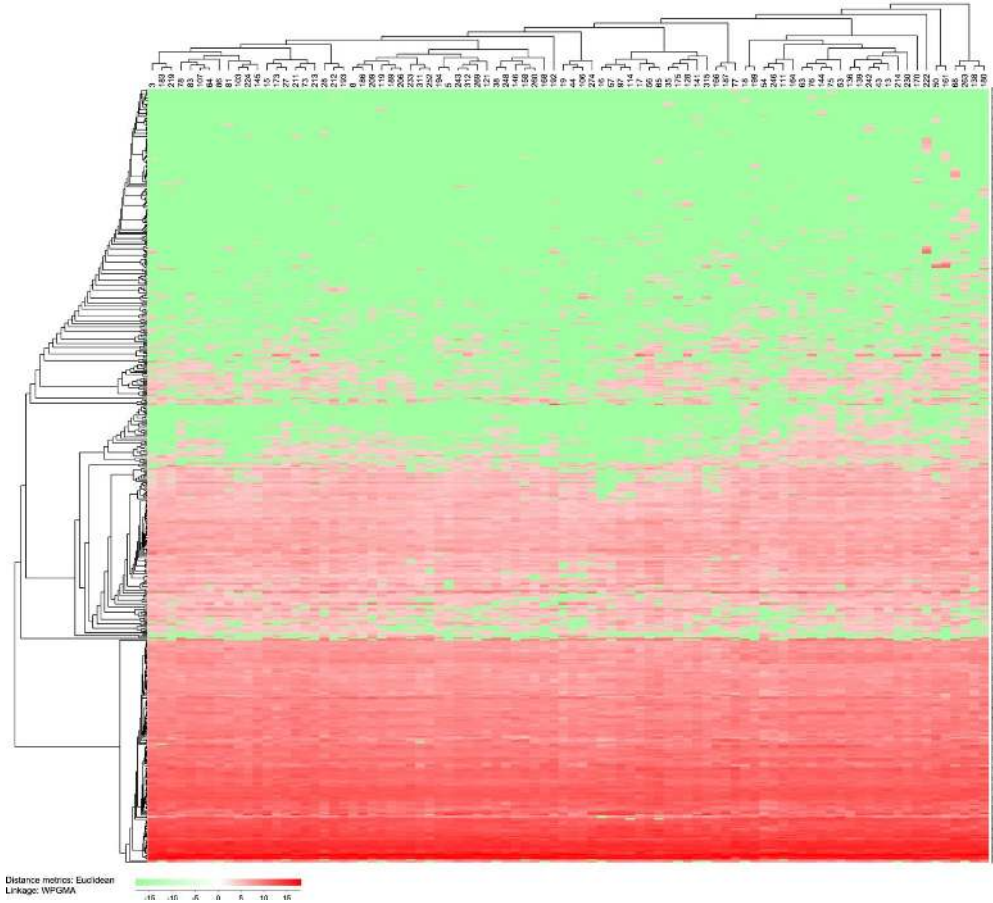


Figure 7

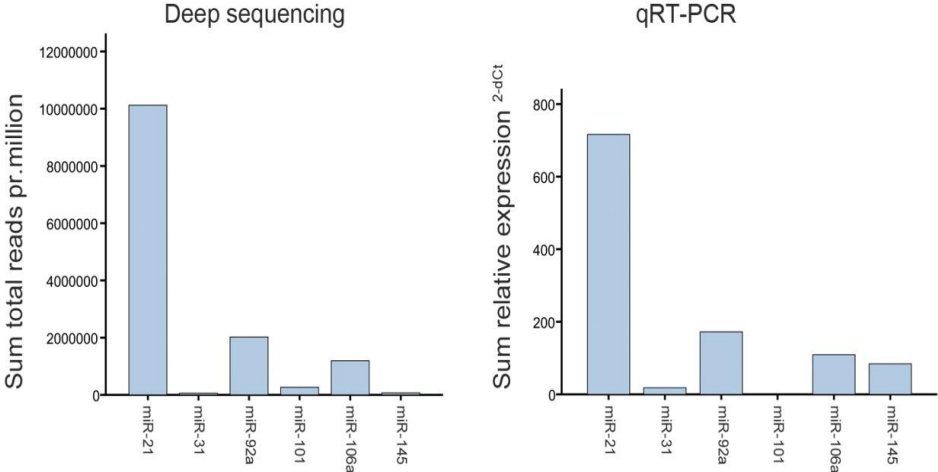
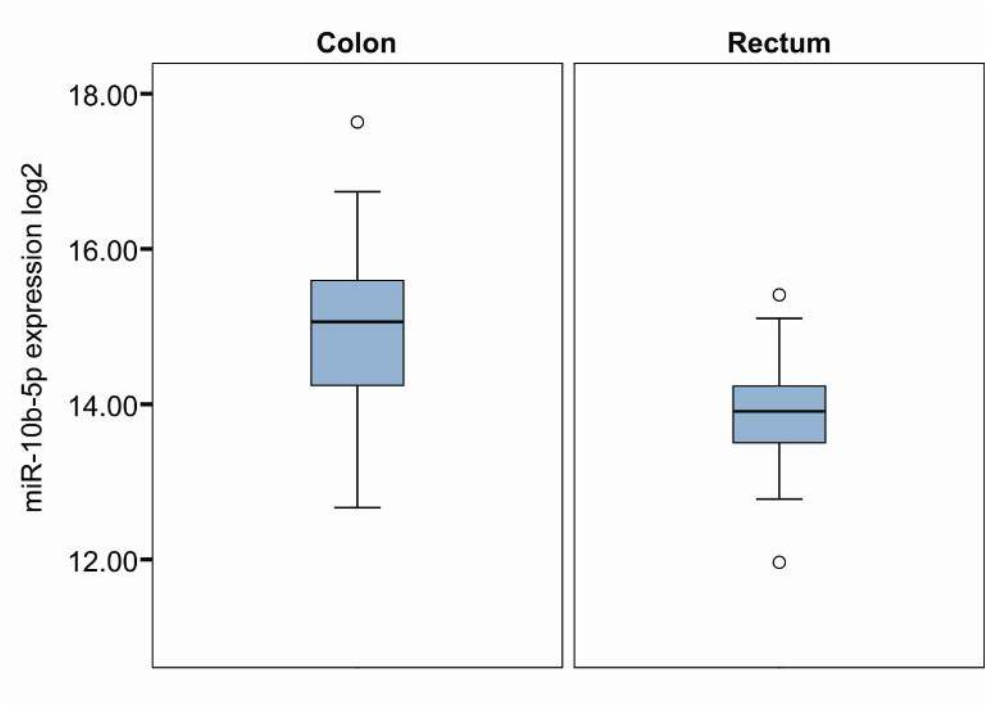


Figure 8



## **Erratum**

### **Reference list dissertation page 74**

**The year has been added for the following references on pages 74-85:**

1, 9, 16, 22, 23, 40, 43, 48, 49, 75, 77, 89, 90, 93, 123, 139, 149, 150, 151, 152, 153, 164, 165, 175, 176, 178, 179, 184, 188, 189, 194, 195, 197, 198, 202, 203, 204, 205, 206, 207, 208, 215, 219, 224, 225, 233, 235, 236, 237, 239, 242, 244, 245, 246, 248, 249, 251, 252, 254.

**Year, issue and page numbers have been added to the following references on pages 74-85:**

13, 45, 51, 72, 92, 183, 187, 199, 200, 201, 210, 218, 223, 226, 247, 250.

### **Reference list paper II:**

**The year has been added for the following references for paper II:**

5, 11, 14, 16, 17, 18, 20, 22, 25, 26, 27, 28, 29, 30.

### **Reference list paper III:**

**The year has been added for the following references for paper III:**

8, 9, 10, 12, 13, 14, 15, 17, 22, 24, 25, 32, 34, 35, 36, 37, 40, 41, 42, 43, 44.