UiT The Arctic University of Norway

Faculty of Health Sciences
Department of Clinical Medicine

Assessing new prognostic biomarkers in resected colon cancer patients

A retrospective study evaluating the prognostic impact of expression of new biomarkers in colon cancer patients utilizing *in situ* hybridization and immunohistochemistry on tissue microarrays.

Hallgeir Selven

A dissertation for the degree of Philosophiae Doctor - March 2023.

“I have expressed the opinion that we are on the eve of a discovery for the arrest, or, perhaps, the cure, of cancer.”

- Sir Samuel William Langston Parker, British surgeon, “The modern treatment of cancerous diseases by caustics or enucleation”, 1856.

Table of Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

The present work was carried out at the Translational Research Group, Department of Clinical Medicine, Faculty of Health Sciences, UiT The Arctic University of Norway. It was financed by the university. Data collection started already in 2012. Several years of standstill followed, until I was accepted as a full-time PhD-student in January 2018.

I would like to thank my initial main supervisor, Roy Bremnes, for giving me the opportunity to enter the world of scientific research. Unfortunately, health issues hampered your possibilities to contribute in the way we planned. A special thanks goes to Thomas Kilvær, who I have been privileged to call my friend and colleague for a long time. You stepped up as my new main supervisor, and without your continuous help and encouragement, I never would have reached the finish line. I am afraid I have given you a substantial amount of grey hair over this period, but I am truly baffled by your knowledge, work-capacity, availability, and positivity. Thanks to my co-supervisors, Lill-Tove Rasmussen Busund and Sigve Andersen for invaluable feedback and support.

Thanks to all members of the Translational Research Group, who have contributed to this work. Unfortunately, the covid-pandemic severely limited our social contact, but I've learnt a lot from every single one of you. (I'm sorry I couldn't keep up with your running standards, but I gave it some tries – at the cost of several repetitive strain injuries…)

I want to thank my former head of department, Tone Nordøy, for employing me at the Department of Oncology in 2011. I am also very grateful that your successor, Magnar Johansen, continued to show faith in me as an oncologist. Magnar even initiated the resumption of this project after the beforementioned standstill, and generously provided me with leave from my clinical work to finish this project. Thanks to all my colleagues at the Department of Oncology. You are truly wonderful people, and I am really looking forward to working with you again.

Finally, I want to thank my family and friends. To my mother and father, I probably have not said this enough, but I am extremely grateful for all you have done for me through the years. You have given me nothing but unconditional love and support, and for that you have my sincere gratitude.

To my lovely children; Hedvig, Åse and Sverre: Thank you for being who you are, constantly exploring the world (and the limits of my patience). Dear Sigrid, my fiancée, making life perfect for both the kids and me. Your calmness and loving nature keep our family together, and I love you with every microRNA in my body.

In loving memory of my grandmother, Ragnhild Elise Myrhaug (September 1, 1932 – December 26, 1995), who succumbed to metastatic colon cancer.

LIST OF PAPERS

PAPER I

Selven H, Busund LR, Andersen S, Bremnes RM, Kilvær TK.

High expression of microRNA-126 relates to favorable prognosis for colon cancer patients.

Sci Rep. 2021 May 5;11(1):9592. doi: 10.1038/s41598-021-87985-3.

PAPER II

Selven H, Andersen S, Pedersen MI, Lombardi APG, Busund LR, Kilvær TK.

High expression of miR-17-5p and miR-20a-5p predicts favorable disease-specific survival in stage I-III colon cancer.

Sci Rep. 2022 Apr 30;12(1):7080. doi: 10.1038/s41598-022-11090-2.

PAPER III

Selven H, Busund LR, Pedersen MI, Lombardi APG, Andersen S, Kilvær TK.

High expression of IRS-1, RUNX3 and SMAD4 are positive prognostic factors in stage I-III colon cancer.

Cancers 2023, 15(5), 1448; doi: 10.3390/cancers15051448.

LIST OF ABBREVIATIONS

AJCC	American Joint Committee on Cancer
AKT	Protein kinase B
Bcl-2	B-cell lymphoma 2
B-CLL	B-cell chronic lymphoblastic leukemia
BRAF	A protein kinase encoded for by the proto-oncogene B-Raf
c-MYC	Cellular myelocytomatosis oncogene
CPP	Cancer patient pathway
CRC	Colorectal cancer
CSP	Cold snare polypectomy
DFS	Disease-free survival
DICER1	Dicer1 Ribonuclease III
DNA	Deoxyribonucleic acid
Dpc4	Deleted in pancreas cancer 4
DROSHA	Drosha ribonuclease III
EMR	Endoscopic mucosal resection
EMT	Epithelial-mesenchymal transition
ESD	Endoscopic submucosal resection
FAP	Familial adenomatous polyposis
FFPE	Formalin-fixed, paraffin-embedded
FGF	Fibroblast growth factor
HCC	Hepatocellular carcinoma
HNPCC	Hereditary non-polyposis colorectal cancer
HSP	Hot snare polypectomy
IHC	Immunohistochemistry
IRS	Insulin receptor substrate
ISH	In-situ hybridization
KRAS	Kirsten rat sarcoma virus
LNA	Locked nucleic acid
MAPK	Mitogen-activated protein kinase
miR	MicroRNA
MIR17HG	miR-17~92 cluster host gene
miRISC	MicroRNA-induced silencing complex
miRNA	MicroRNA
mRNA	Messenger RNA
NSCLC	Non-small cell lung cancer
OS	Overall survival
PI3K	Phosphoinositide 3-kinase
Pre-miRNA	Precursor microRNA
Pri-miRNA	Primary microRNA
PTEN	Phosphatase and tensin homolog
qRT-PCR	Real-time reverse transcription polymerase chain reaction
RAS	From Rat sarcoma virus
RNA	Ribonucleic acid
RUNX	Runt-related transcription factor
SMAD4	From Small Mothers Against Decapentaplegic 4
TGF-β	Transforming growth factor beta
TMA	Tissue microarray
TNM	Tumor-node-metastasis
TRAIL	TNF-related apoptosis inducing ligand
UICC	Union for International Cancer Control
UNN	University Hospital of North Norway
VEGF	Vascular endothelial growth factor

1 INTRODUCTION

Cancer is a common and serious disease. It is estimated that approximately 40% of the Norwegian population will be diagnosed with cancer by the age of 80.[1] Worldwide, there is an estimated incidence of more than 18 million cancer cases annually.[2] Historically, awareness of cancer as a disease can be traced centuries and probably millennia back in time. Regarding cancer treatment, surgical resection of tumors was performed in ancient times, and the knowledge that it is possible to cure some solid tumors with surgery was known from the 19th century. The introduction of chloroform anesthesia in the mid-19th century facilitated better surgical interventions for solid tumors. Mortality was still high though, mainly due to septicemia. The introduction of antibiotics, following Fleming's discovery of Penicillin in 1928, improved survival significantly. Cancer surgery was further improved in the mid-20th century, aiming to prevent cancer cell dissemination by isolation of the cancer by first dividing its feeding and draining vessels. Advanced planning of cancer surgery became easier in the 1970s, as radiologic imaging improved. The late 20th century and the beginning of the 21st century introduced the concepts of laparoscopic cancer surgery and more refined surgical techniques.[3] Likewise, with the discovery of X-rays and ionic radiation in the late 19th century, initial attempts at radiotherapy were performed at that time. Chemotherapy was introduced in the late 1940s and significantly improved the prognosis for cancer patients. The knowledge of the pathogenesis of cancer has increased immensely over the last decades and the treatment has improved. With this acquired knowledge, new treatments arise. More recent supplements to anti-cancer treatments are targeted therapies (i.e., monoclonal antibodies, tyrosine kinase inhibitors) and modern immunotherapy (immune checkpoint inhibitors). Despite increasing knowledge and improved therapy, cancer is still a deadly disease. Further research to increase the understanding of the pathogenesis, helping to tailor treatment and improve the prognosis for cancer patients is very much needed. Contrary to Sir William Langston Parker's optimistic quote 167 years ago, today we are more realistic, acknowledging that cancer is a disease we must cope with for a long, long time to come.

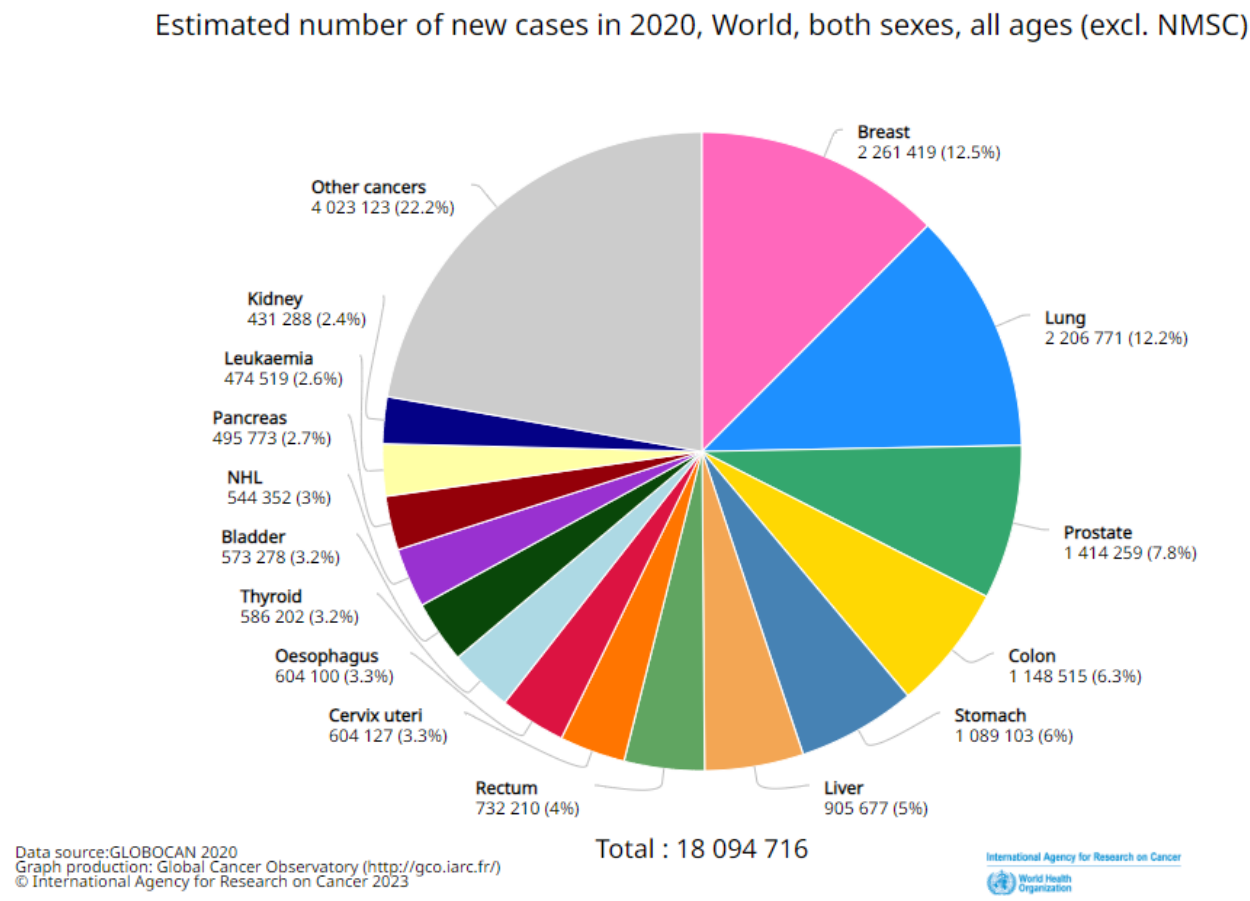

***Figure 1**: Estimated global cancer incidence 2020. Data source: GLOBOCAN 2020.*[2]

1.1 Colon cancer

1.1.1 Incidence and mortality

Colon cancer is a common malignancy. Globally, it is the 4th most common malignancy in both sexes combined, with an estimated number of 1.150.000 new cases in 2020. This represents approximately 6% of all new cancer cases diagnosed this year (

Figure 1).[2] It is the 4th and 5th most common cancer in females and males, respectively. Colon cancer is also one of the leading causes of cancer-related mortality. It is estimated that approximately 550.000 patients will succumb to colon cancer in 2020, making it the 5th most common cause of cancer related deaths.[4] Norway is one of the countries in the world with the highest incidence of colon cancer (**Figure 2**, **Figure 3**), and the reasons are unknown. It is estimated that approximately 3% of the Norwegian population will be diagnosed with colon cancer by the age of 75.[5] The median age at diagnosis in Norway is 73, making colon cancer a

rare disease in the young population (<40 years). However, lately there has been an alarming increase in colon cancer incidence among young adults, both in Northern America and Europe.[6,7]

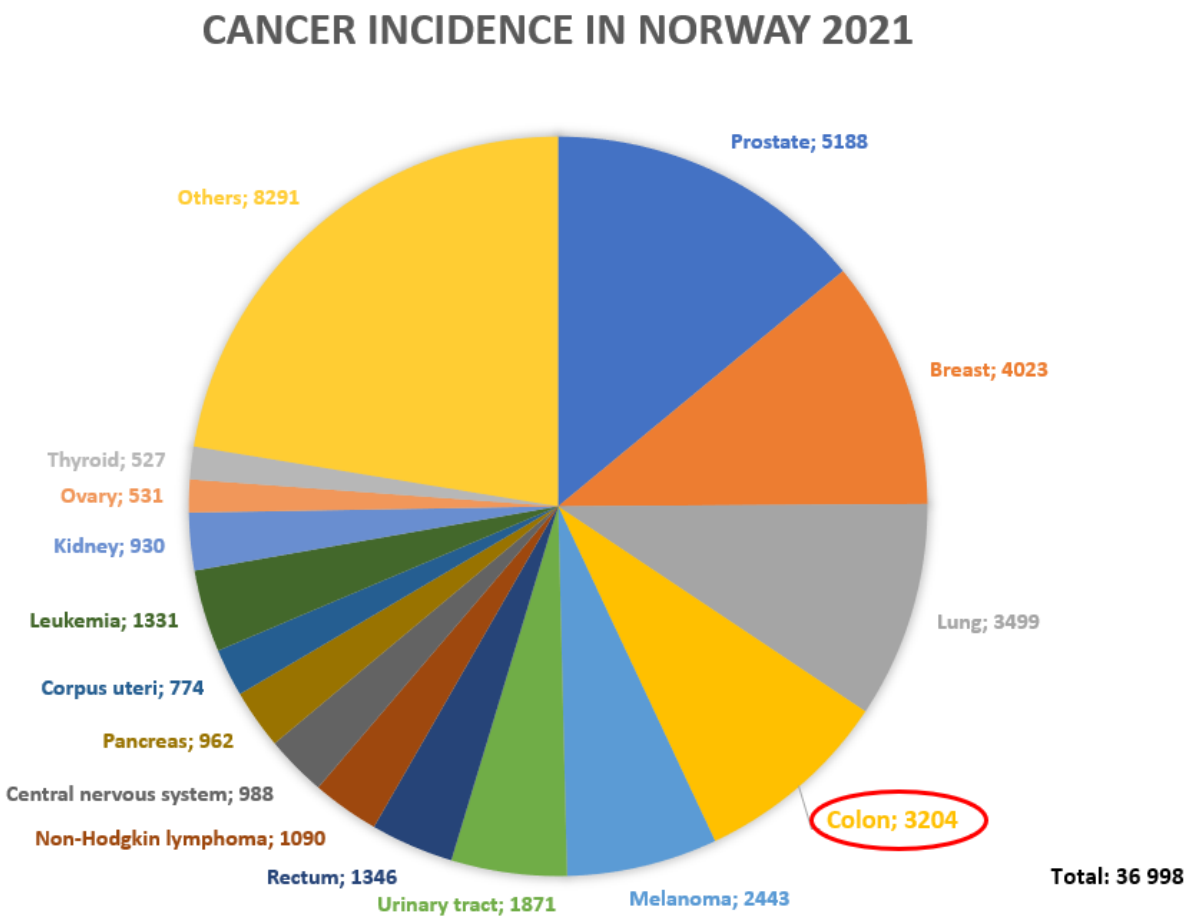

***Figure 2**: Cancer incidence in Norway 2021. Data source: www.kreftregisteret.no; The Cancer Registry of Norway.*[1]

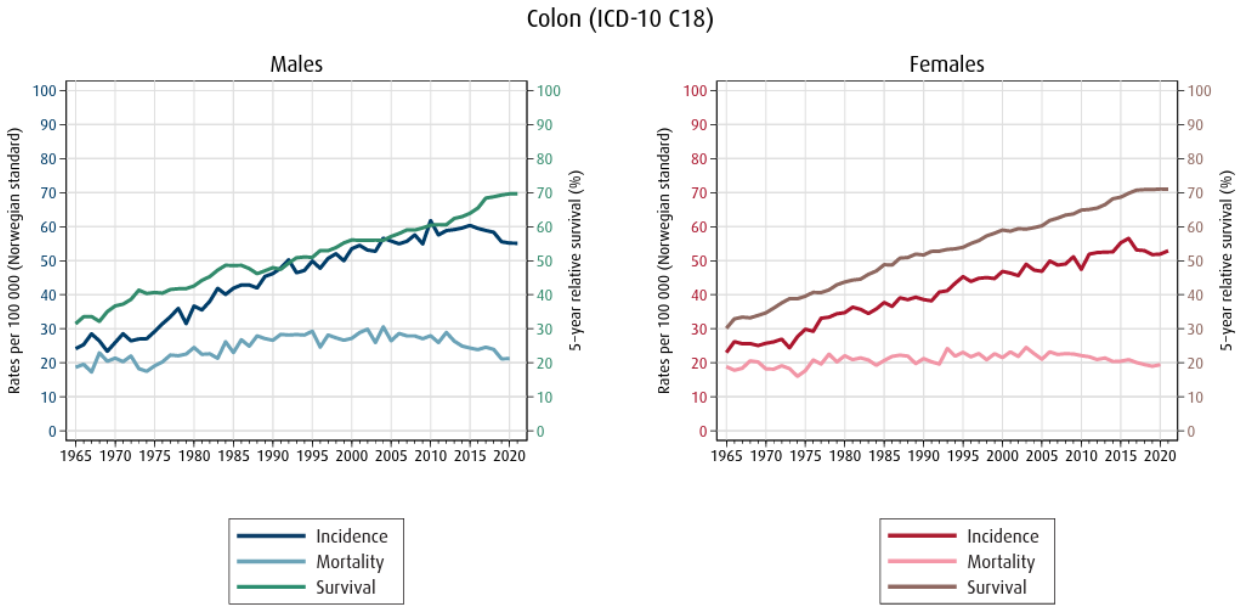

***Figure 3**: Trends in incidence and mortality rates and 5-year relative survival proportions. (Adapted from www.kreftregisteret.no; the Cancer Registry of Norway.[1])*

1.1.2 Risk factors

Several dietary and lifestyle factors influence the risk of contracting colon cancer. Physical inactivity, overweight and central obesity are consistent risk factors.[8] Consumption of red meat, processed meats and refined carbohydrates also seem to contribute to risk. High alcohol consumption and smoking early in life increase the risk of colon cancer. The abovementioned risk-factors are all modifiable, and it's proposed that the majority of colon cancers can be preventable by changes in diet and lifestyle.[9] It is also an explanation for demographic differences in colon cancer incidence, as the Western lifestyle is associated with higher incidence.[10] Non-modifiable risk factors are age, ethnicity, comorbidities (e.g inflammatory bowel disease) or a family history of colorectal cancer.[11] Approximately 2-5% of all colon cancer arise due to hereditary syndromes.[12]

1.1.3 Pathogenesis

Most colon cancer cases arise in a preexisting, benign colon tumor (adenoma). An adenoma typically appears due to sporadic mutations in the adenomatous polyposis coli-pathway or DNA mismatch repair, and histologically contains dysplasia. Adenomas may

increase in size and progress from low-grade to high-grade dysplasia, and eventually to carcinoma *in-situ* and even to invasive carcinoma. This is a process that normally takes several years (**Figure 4**).[13] Most malignant tumors of the colon are adenocarcinomas (~95%).

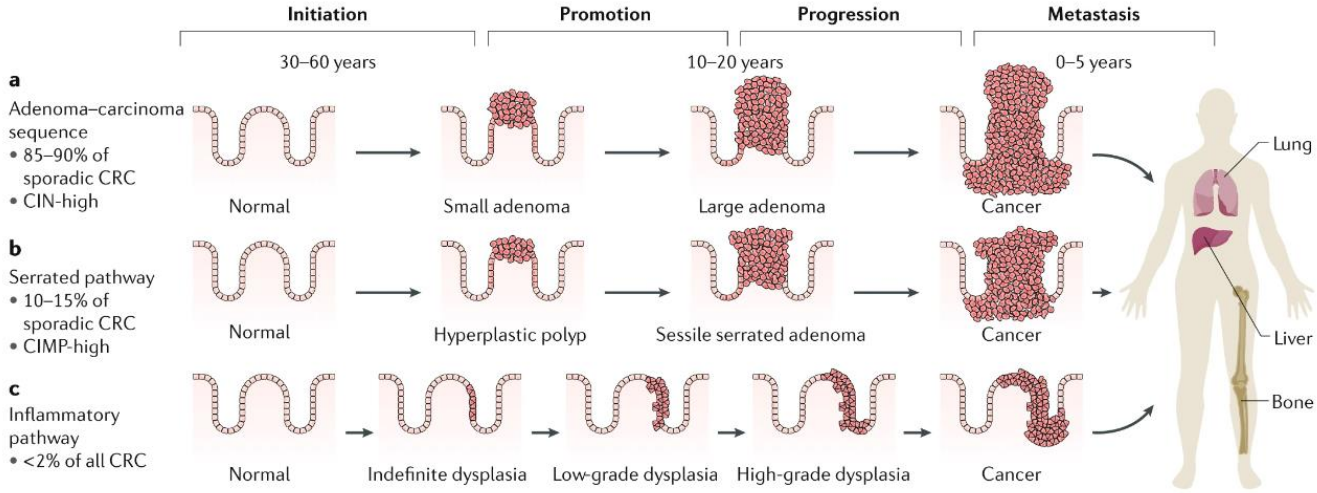

***Figure 4**: Pathways of colorectal carcinogenesis, illustrating the four stages: initiation, promotion, progression and metastasis. Each of the timespans differs substantially, and the exact timing is difficult to estimate. Illustrated pathways in the figure are the adenoma-carcinoma sequence (**a**), serrated pathway (**b**), and the inflammatory pathway (**c**). (Adapted from Keum and Giovannucci with permission.*[14]*)*

People with hereditary syndromes develop cancer at an earlier age (mentioned in section 1.1.2). Familial adenomatous polyposis (FAP) has a prevalence of 1 in 10 000 individuals. The classic feature is the development of hundreds to thousands of colonic adenomas from early adolescence. If untreated, colorectal cancer is inevitable. The average age of CRC onset for untreated individuals is 39 years (7% before the age of 21, and 95% by the age of 50).[12] In Lynch syndrome (also called hereditary nonpolyposis colorectal cancer [HNPCC]), affected individuals are predisposed for various cancers, among them colon cancer. Lynch syndrome accounts for 2-4% of all CRCs. Affected individuals develop more frequent adenomas than the general population, but polyposis is rare. Approximately 50-80% of individuals with Lynch syndrome develop CRC in their lifetime.[15]

1.1.4 Clinical presentation and screening

Colon cancer presents clinically in a very heterogeneous manner. Early-stage colon cancer usually presents with no/mild symptoms. These patients are mainly diagnosed due to screening. For the more advanced stages, more pronounced symptoms usually occur, and consequently, most colon cancer patients are diagnosed after the onset of symptoms. A typical symptom is blood in the stool. It can present as occult blood in the stool if there's a small bleeding. If there is a profuse bleeding (usually in distal cancers), it presents as hematochezia – the passage of fresh blood per anus. In more proximal cancers, the blood is digested throughout its passage through the colon, and presents as melena – the passage of black, tarry stools. Other symptoms are changes in bowel habits persisting over several weeks, or abdominal pain.[16] Iron-deficiency anemia of unknown reason may also be a sign of colon cancer, due to a longstanding, small bleeding. Approximately 15-25% of colon cancer patients presents with a need for emergency surgery, due to bowel obstruction, perforation, or a major bleeding.[17,18]

The diagnosis can sometimes be delayed, as other medical conditions present with similar symptoms. The most important differential diagnoses being inflammatory bowel diseases (ulcerative colitis and Crohn's disease), inflammatory bowel syndrome and hemorrhoids. Norway, as well as several other countries, has introduced a cancer patient pathway (CPP), often referred to as a "fast track" in cancer diagnosis.[19] From the age of 40, if you experience an intestinal bleeding of unknown reason, the presence of a tumor or polyp at lower endoscopy, or experience changes in bowel habits for longer than 4 weeks, you qualify for a shortened time to diagnostic procedures in the cancer patient pathway.[20] The Norwegian colorectal cancer screening program was initiated in May 2022.[21] In this program, asymptomatic adults turning 55 years of age in the current year are invited to participate. Initially, screening will be performed testing for occult blood in the stool (FOBT). It is estimated that 6% of the population will be diagnosed with blood in the stool by this test, and these individuals will be referred to a colonoscopy. Among individuals with blood in the stool, 1 out of 20 is estimated to have colorectal cancer. On the contrary, 1 out of 4 patients with colorectal cancer do not have blood in their stool, and consequently won't be identified by this method. With time, primary colonoscopy is planned to be part of the screening program, increasing the chance of diagnosing colorectal cancer patients without blood in their stool at an early stage. In a systematic review of screening for colorectal cancer, the absolute

effect was much more favorable for older adults (≥ 60 years), with numbers needed to screen of 343 and 492 for flexible sigmoidoscopy and fecal occult blood test, respectively.[22]

1.1.5 Staging and prognosis

TNM staging

Colon cancer prognostication is based on the TNM-system and histopathological criteria according to the American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC).[23] The tumor extent (T), lymph node involvement (N), and presence of metastasis (M) divides colon cancer into stages (**Table 1**). In stage I, the cancer has grown through the muscularis mucosa into the submucosa (T1), or into the muscularis propria (T2). There is no involvement of local lymph nodes nor signs of distant metastasis. In stage II, the cancer has grown into the outermost layers of the colon (T3). It may also have grown through the wall of the colon without (T4a) or with (T4b) growing into nearby tissues or organs. Still there is no involvement of local lymph nodes nor distant metastasis. In stage III, you can have every T-stage, but there is involvement of local lymph nodes (N+). It has not spread to distant sites. In stage IV, you can have any T- or N-stage, but the cancer has spread to distant sites (M+).

***Table 1**: TNM classification for colon cancer, 8th edition. Adapted from "TNM Classification of Malignant Tumors, 8th edition".*[23]

Stage	Tumor	Node	Metastasis
Stage 0	Tis	N0	M0
Stage I	T1, T2	N0	M0
Stage II	T3, T4	N0	M0
Stage IIA	T3	N0	M0
Stage IIB	T4a	N0	M0
Stage IIC	T4b	N0	M0
Stage III	Any T	N1	M0
Stage IIIA	T1, T2 T1	N1 N2a	M0 M0
Stage IIIB	T1, T2 T2, T3 T3, T4a	N2b N2a N1	M0 M0 M0
Stage IIIC	T3, T4a T4a T4b	N2b N2a N1, N2	M0 M0 M0
Stage IV	Any T	Any N	M1
Stage IVA	Any T	Any N	M1a
Stage IVB	Any T	Any N	M1b
Stage IVC	Any T	Any N	M1c

The prognosis of colon cancer relies heavily on the stage of the disease upon diagnosis. If diagnosed at an early stage of the disease, it comes with a favorable prognosis. In the more advanced stages, the prognosis is less favorable. In Norway, the 5-year relative survival is >95%, ~80% and ~20% for localized, locoregional and metastatic disease, respectively.[5] Among histopathologic indicators of poor prognosis in stage II and III colon cancer, pT4 is the most robust, and pT4 *vs* pT3 hazard ratios of DFS and OS (in stage II) are similar to pN1 *vs* pN0 disease.[24]

Other prognosticators

Histological grade of the tumor is an important factor, as patients with poorly differentiated tumors have an impaired prognosis compared to patients with well differentiated tumors.[25] The presence of lymphovascular invasion is an independent predictor

of poor prognosis, as it is associated with lymph node involvement and metastasis.[26] Preoperatively elevated levels of carcinoembryonic antigen (CEA) are also associated with impaired prognosis.[27] Additionally, a high degree of tumor budding is associated with poor survival.[28] These prognosticators don't indicate a need for adjuvant chemotherapy on their own, but if several of these characteristics are present, they might do.

1.1.6 Treatment

Surgical treatment

To cure colon cancer, surgery is required. For early-stage cancers, endoscopic resection may be an option. This is performed by cold or hot snare polypectomy (CSP or HSP), endoscopic mucosal resection (EMR), or endoscopic submucosal dissection (ESD).[29] For the more advanced stages, resection of a segment of the colon is needed. The extent of the surgery depends on tumor location, the presence of synchronous cancers, if there are multiple polyps in the colon, genetic aberrations, and comorbidity among others.[30] The most common technique is the complete mesocolic excision (CME), including the primary tumor, all corresponding lymphatic and blood vessels, and draining lymph nodes.[31] In T4-tumors, surgery is done by en bloc resection of the primary tumor along with all the mesentery containing the primary blood supply, lymphatics of the involved colonic segment, and any involved adjacent organs.

Regarding resection margins in colon cancer, a resection margin of 10 cm both proximally and distally has been the standard of care. However, it is been proven that a resection margin of 5 cm is sufficient in colon cancer.[30] The circumferential resection margin (CRM) should be >1mm to be considered negative in colon cancer.[32] A CRM ≤1 mm is considered positive.

Adjuvant chemotherapy

Adjuvant chemotherapy is administered to eradicate potential microscopic disease after surgery, thereby minimizing the risk of recurrent disease. The effect of adjuvant 5-FU-based chemotherapy in colon cancer was reported in studies performed in the late 1980s and early 1990s.[33,34] In 2004, the MOSAIC-trial demonstrated benefit of adding oxaliplatin to the

5-FU/Leucovorin-regimen, being standard treatment at the time.[35] The effect is striking in stage III patients, with updated data showing a 10-year overall survival of 67% *vs* 59% (FOLFOX4 *vs* LV5FU2).[36] Stage III patients not receiving chemotherapy have reported 5-year DFS of approximately 50%.[37] Hence, these patients should be offered adjuvant chemotherapy unless there are contraindications (high age, comorbidity etc.). For stage II patients the effect is less clear, and only patients with a high risk of recurrence are offered adjuvant chemotherapy (high risk being perforated tumor, T4-tumors, or low number of examined lymph nodes). Lately, the regimens for adjuvant chemotherapy have changed, making the treatment more individualized. Patients with a low risk of recurrence (T1-T3, N1) are candidates for three months of adjuvant therapy, reducing the risk of side-effects from the treatment and lowering health expenditures. High risk patients (T4, N2), should still be offered six months of adjuvant chemotherapy.[38]

Treatment of metastatic disease

Approximately 20% of patients with colorectal cancer present with synchronous metastases.[39] Additionally, 30-40% of patients treated with curative intent will be diagnosed with recurrent disease.[40] The most common metastatic sites are liver, lung and peritoneum affecting 70%, 32% and 21% of patients with metastatic disease, respectively.[41] Less common metastatic sites include the central nervous system, bones, distant lymph nodes and ovaries, among others. As blood is drained from the colon through the portal system to the liver, it is not surprising that most metastases arise in the liver. In case of solitary metastases to the liver, it may even be looked upon as a continuation of the colon and treated in a curative manner. Unfortunately, approximately 70% of patients who undergo curative hepatectomy for colorectal liver metastases develop a recurrence.[42] Still, if the recurrence is resectable (liver or lung), undergoing another resection is associated with improved survival compared to an unresected recurrence.

Metastatic colon cancer can be treated in different ways. Oligometastatic disease, most frequently seen in the liver or lungs, can be treated by metastasectomy.[43] A population-based analysis published in 2020, including 16 372 patients, showed better overall survival for patients who underwent surgical resection of liver metastases compared to controls (median OS 38 months *vs* 13 months, $p<0.001$). Similar results were seen for lung metastases, favoring patients who underwent metastasectomy (median OS 45 months *vs* 19 months,

p<0.001). For lung metastases, recent studies have questioned whether metastasectomy contributes to increased survival or not.[44] For patients not eligible for surgery, radiofrequency ablation or stereotactic radiotherapy are alternatives to surgery.

For polymetastatic, non-operable colon cancer, systemic treatment usually is the best option for eligible patients. Chemotherapy has been the first-line treatment for this group of patients, usually in the form of fluorouracil-based chemotherapy combined with oxaliplatin or irinotecan.[45] In the case of RAS/BRAF wild-type tumors, monoclonal antibodies directed against the epidermal growth factor receptor (cetuximab or panitumumab) can be given together with chemotherapy for synergistic effect.[46] If there is a mutation of RAS or BRAF, antiangiogenic therapy can be administered together with chemotherapy, mainly bevacizumab, a monoclonal antibody targeting the vascular endothelial growth factor (VEGF).[47] In clinical trials, median overall survival for previously untreated advanced or metastatic KRAS wild-type tumors treated with first-line chemotherapy and cetuximab or bevacizumab was ~30 months.[48] Response rates are reported at 55-60%. On progression after first line treatment, patients who were given an irinotecan-based regimen usually switch to an oxaliplatin-based regimen in the second line and *vice versa*. Among the 10% of patients with metastatic disease who harbor a BRAF V600E-mutation, the prognosis is impaired. These patients should be considered for more intense first line treatment, usually FOLFOXIRI (fluorouracil, oxaliplatin, irinotecan and calcium folinate). As second line treatment, these patients can benefit from BRAF-inhibitors (i.e., encorafenib) in combination with an EGFR-inhibitor (cetuximab).[49] In Norway, the multikinase inhibitor regorafenib and antimetabolite trifluridin/tipiracil /TAS-102) are approved treatment alternatives in third- or fourth line settings.[50,51] Monotherapy with these agents should not be offered routinely as the overall survival benefit is low (~1.5 months) and the toxicity profile is unfavorable. In rare cases, it can be offered to patients with a satisfactory performance status. The combination of trifluridin/tipiracil/TAS-102 and bevacizumab is showing promising results in clinical trials.[52]

Recent studies have shown that microsatellite instable/deficient mismatch repair tumors are likely to respond to immunotherapy (immune checkpoint inhibitors). This comprises approximately 15% of patients with colorectal cancer.[53] The NICHE-study has evaluated the effect of neoadjuvant immunotherapy in dMMR colon cancer, and results are very promising.[54,55] In the metastatic setting, patients with dMMR colorectal cancer are likely to respond to immunotherapy as palliative treatment, and this comprises approximately 5% of

patients with metastatic disease.[56] A flowchart for selection of systemic therapy in metastatic colorectal patients in Norway is shown in **Figure 5**.

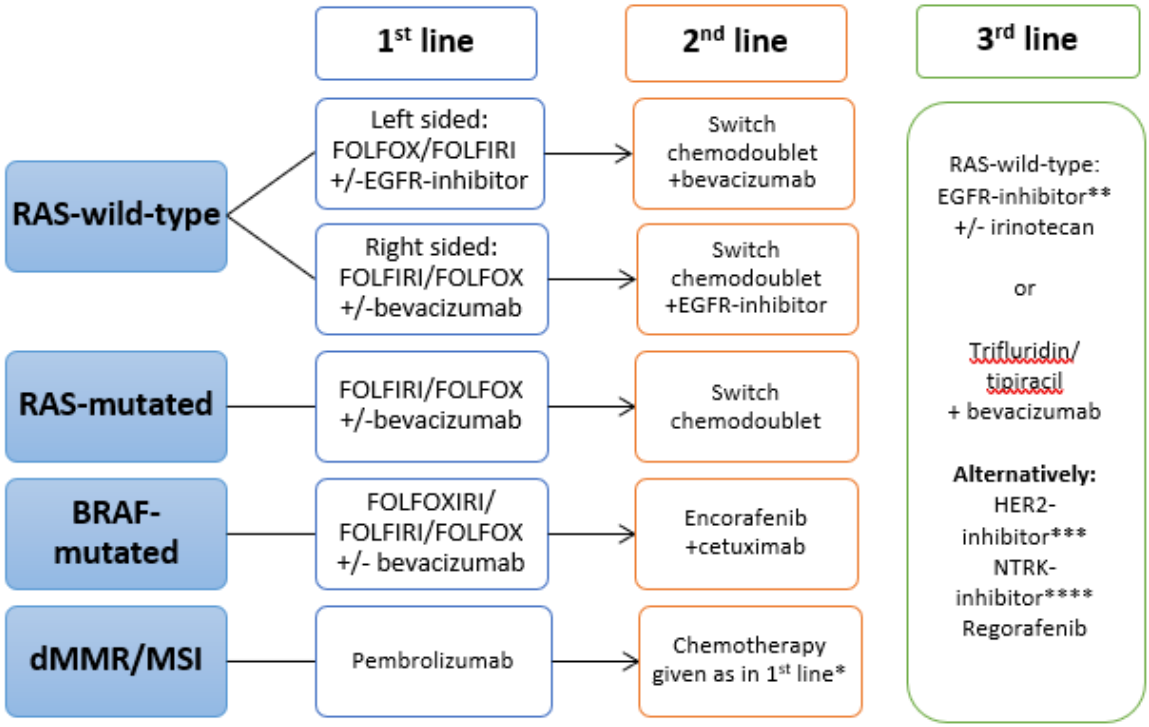

***Figure 5**: Flowchart for selection of systemic therapy for metastatic colorectal cancer in Norway. Figure adapted and translated with permission from The Norwegian Gastrointestinal Cancer Group (www.ngicg.no).*

1.1.7 Rectal cancer *vs* colon cancer

Colon cancers and rectal cancers are often referred to together as colorectal cancers, and a large amount of research is carried out on colorectal cancers combined. Rectal cancers are somehow different from colon cancers as they show less effect of adjuvant chemotherapy. Thus, despite lymph node involvement, rectal cancer patients haven't been given adjuvant chemotherapy routinely. But to think of a distal colon cancer as a completely different entity than a proximal rectal cancer is highly debatable. Embryologically, the proximal colon arise from the midgut (cecum, ascending colon, proximal 2/3 of transverse colon), whereas the distal colon arise from the hindgut (distal 1/3 of transverse colon, descending colon, sigmoid colon, rectum, superior anus).[57] Approximately 10% of colorectal cancers origin from the rectosigmoid, and the proper treatment for these tumors are often a subject for debate at multidisciplinary team meetings (involving radiologists, colorectal cancer surgeons, oncologists and pathologists).[58] Treatment of rectosigmoid/rectal cancers is a landscape that is continuously evolving and may be about to change in the years to come. Of particular interest is the finding that mismatch repair-deficient, locally advanced rectal cancers respond tremendously to immunotherapy up-front.[59]

1.2 MicroRNAs

MicroRNAs (miRNAs, miRs) are small, non-coding RNAs approximately 22 (21-23) nucleotides long. They bind imperfectly to the 3' untranslated region of target messenger RNAs (mRNAs).[60] miRNAs regulate gene expression post-transcriptionally by inhibiting translation and inducing degradation of mRNAs.

The first miRNA was reported in the nematode Caenorhabiditis elegans in 1993.[61] The regulatory function of miRNAs was increasingly understood in the early 2000s, and miRNAs' role in cancer was highlighted in 2002 when deletions of miR-15 and miR-16 were discovered in the majority of B-cell chronic lymphocytic leukemias (B-CLL).[62] Since then, miRNAs' role in cancer development has been extensively investigated. Dysregulation of miRNAs is common in cancer, and miRNAs may function either as oncogenes or tumor suppressors. The true function of the miRNA is depending on cellular context among others, and the same miRNA may have an oncogenic role in one cancer form, but may have a tumor suppressor role in another cancer form.[63] To complicate things further, it can even have a tumor suppressor effect at a relatively low expression, but an oncogenic effect as the tumor expression changes to relatively high levels.

1.2.1 miR biogenesis

In the nucleus, miRNAs are transcribed by the RNA polymerase II enzyme to produce a primary-miRNA (pri-miRNA), several hundred to a thousand nucleotides long (**Figure 6**). Pri-miRNAs can produce single miRNAs or contain clusters of miRNAs processed from a common primary transcript. The microprocessor complex, including the RNase (III) enzyme DROSHA and cofactor DiGeorge syndrome critical region 8 (DCGR8), cleaves the long pri-miRNA, producing a 60-70 nucleotide precursor miRNA (pre-miRNA).[64] Pre-miRNAs are exported to the cytoplasm by exportin 5. In the cytoplasm, a ribonuclease enzyme DICER1, further processes the pre-miRNA to produce the mature ~22 nucleotide miRNA-duplex.[65] One strand of mature miRNA (the guide strand) is integrated in the miRNA-induced silencing complex (miRISC), directing the miRISC to target mRNAs mediating gene suppression by mRNA degradation and translational repression (**Figure 7**).[66] The two strands of the miRNA-duplex, the 5p and the 3p strands, arise from the 5´ and the 3´ end of the duplex, respectively. They can both function as the guide strand or the passenger strand, depending on cell type,

developmental stage and miRNA thermodynamic stability among others.[67] miRNAs are believed to regulate the expression of approximately one third of all the protein-coding genes.[68]

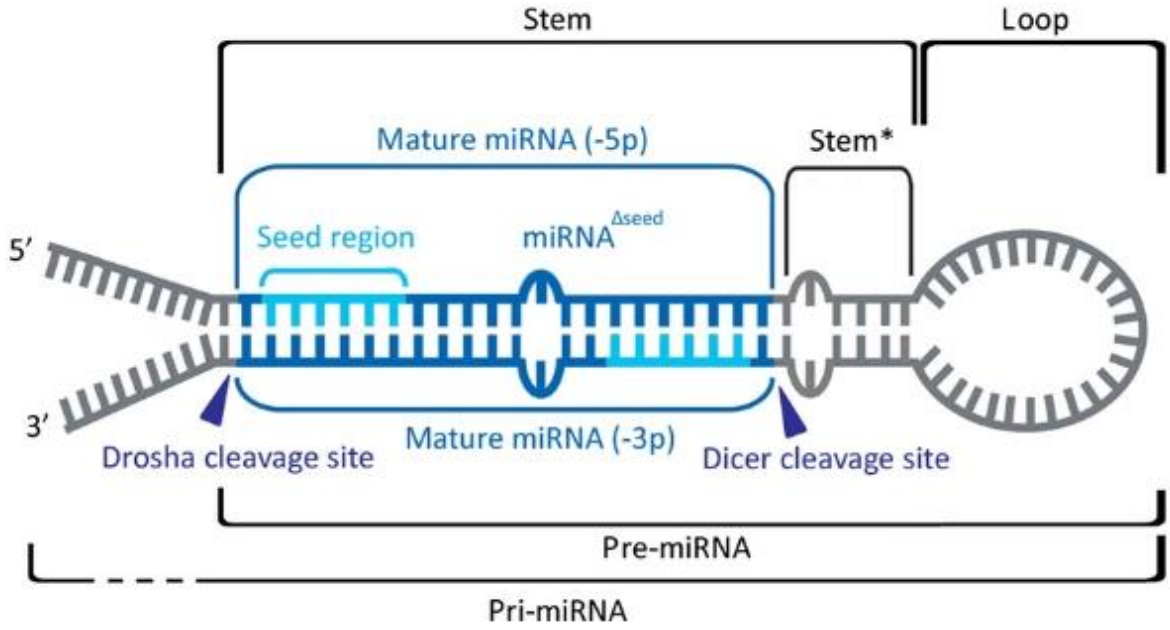

***Figure 6**: Secondary structure of miR. (Figure adapted from Jevsinek Skok et al.[69])*

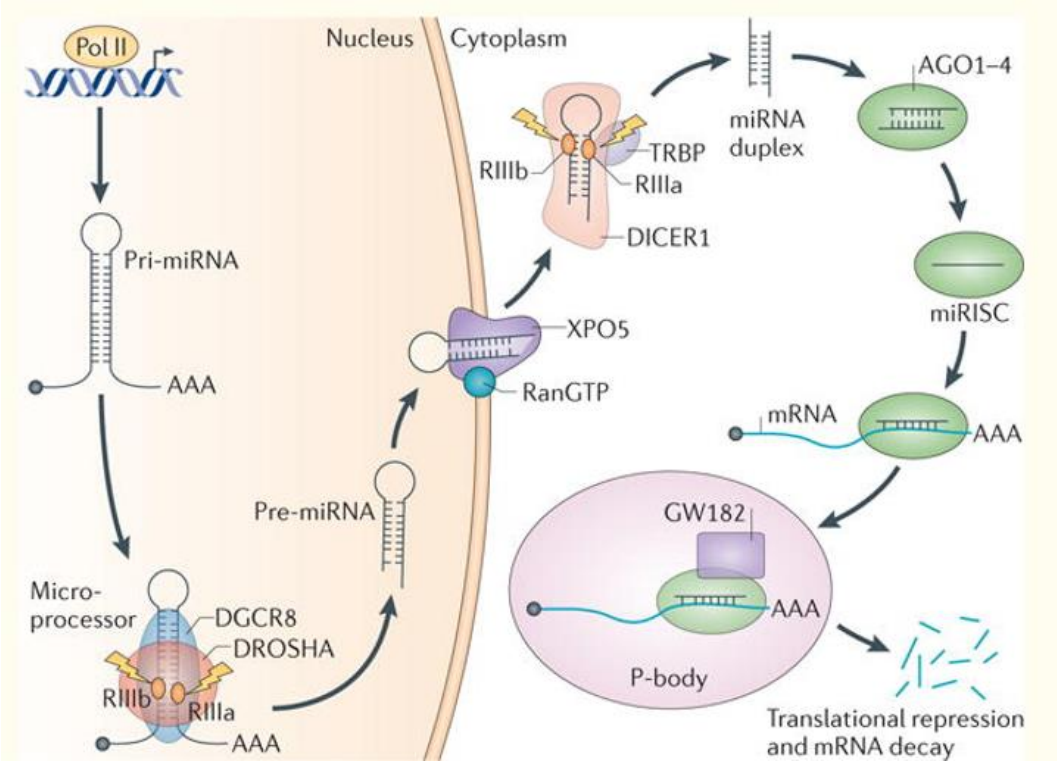

***Figure 7**: Overview of miR biogenesis pathway. miR genes are transcribed as Pri-miRNAs by RNA polymerase in the nucleus. They are cleaved by the microprocessor complex (including DROSHA), producing pre-miRNAs. Pre-miRNAs are transported to the cytoplasm, and further processed by DICER1, an enzyme producing the mature miRNAs. One strand of the mature miRNAs is located into the miRISC, targeting binding to mRNAs, mediating gene suppression and translational repression. (Adapted from Lin and Gregory et al with permission.[70])*

How miRs target mRNAs are explained by two proposed models, the standard model, and the expanded model (**Figure 8**). In the standard model, miR and target mRNA form exact base-pairs, absent of any irregularities in the seed region. The seed is the sequence spanning positions 2-7 inclusive from the 5' end of a miR, which is identical for the different miRs within the same miR-family. The expanded model allows wobble base-pairing (between U and G) or creating bulges on either the miR or mRNA side. The miRs of the miR-17-92 cluster contain at least two G/U bases in their seed region, and potentially rely more on the expanded model.[71]

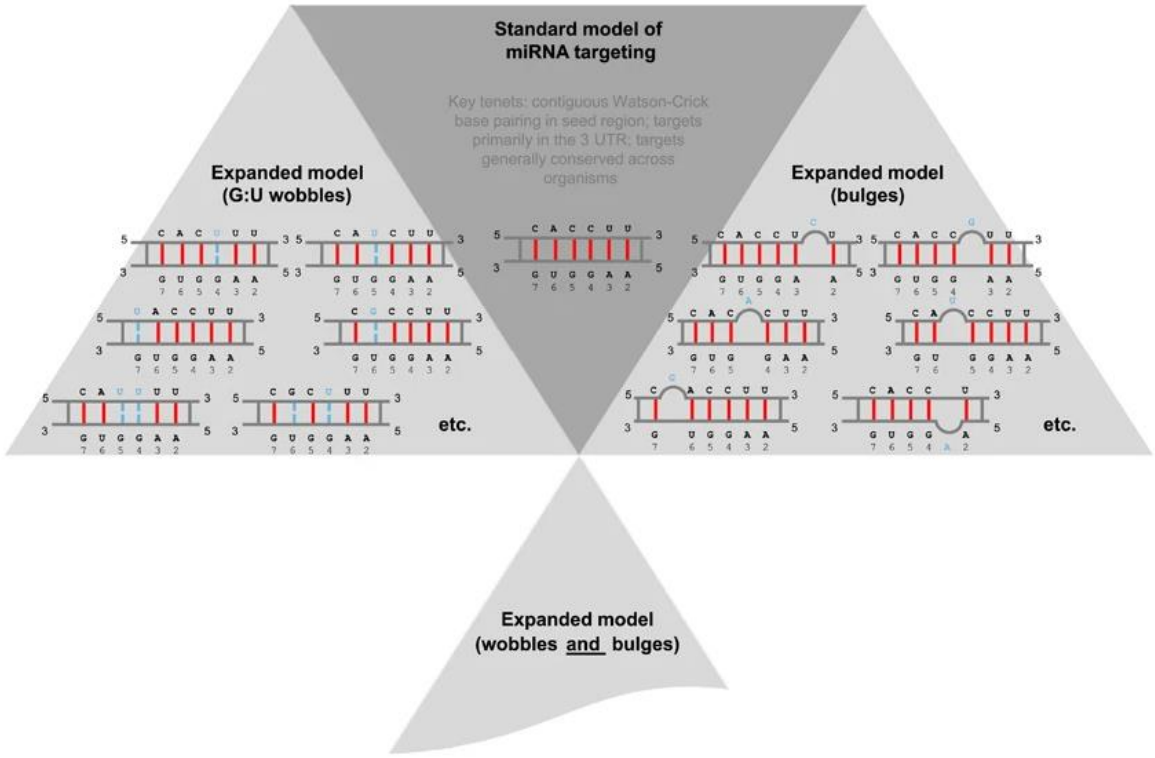

***Figure 8**: Standard vs Expanded models of miR targeting and the corresponding targetome. The standard model is based on exact base pairing, whereas the expanded model allows wobble base-pairing or creating bulges on of the sides. (Figure adapted from Mogilyansky et al.[72])*

1.2.2 miR-analyses

The expression of miRs can be assessed in different samples and with different techniques (qRT-PRC *vs* ISH to be discussed later). In cancer research, resected, cancerous tissue is commonly used. This is a highly invasive technique, and there is an ongoing search for precise, less invasive methods to assess miR expression. For several cancer types, blood (serum/plasma) has proven to give reliable results, detecting circulating miRs.[73] Circulating miRs remain stable due to their incorporation in apoptotic bodies, microvesicles or exosomes, thereby protecting them from degradation. The detection of circulating miRs must be considered minimally invasive, as it only requires a blood sample. For colorectal cancer, another non-invasive method is also available. Expression of miRs can be assessed from stool-samples, and has been shown to correlate with TNM-stages.[74] Whether these minimally-invasive/non-invasive techniques can be used for screening, diagnostic, prognostic or predictive purposes, remains to be elucidated.

Numerous studies have focused on the significance of miRs in colorectal cancer, assessing both over-expressed and under-expressed miRs.[75,76] Of specific interest is the proposed oncogenic miR-17~92 cluster, believed to play an important role in colorectal cancer progression when over-expressed.[77] Another miR of interest in colorectal cancer, miR-126, has been proposed to have a tumor suppressor effect.[78] We wanted to assess the prognostic impact of miR-17-5p and miR-20a-5p, two important members of the miR-17~92 cluster, and miR-126 in our cohort.

1.2.3 miR-17~92 cluster

The miR-17~92 cluster is located at chromosomal locus 13q31.3 in the non-protein-coding gene MIR17HG (the miR-17~92 cluster host gene). It comprises six tandem stem-loop hairpin structures that yield six mature miRs (miR-17, miR-18a, miR-19a, miR-19b, miR-20a and miR-92a).[79] There are two miR-17~92 cluster paralogs in mammals: miR-106b~25 located on chromosome 7 and miR-106a~363 located on the X-chromosome, comprising 6 and 3 mature miRs, respectively. These 15 mature miRs are grouped into four families, namely the miR-17 (miR-17, miR-20a, miR-106a, miR-20b, miR-106b, miR-93), miR-18 (miR-18a, miR-18b), miR-19 (miR-19a, miR-19b-1, miR-19b-2) and miR-92 (miR-92a-1, miR-92a-2, miR-363, miR-25) families. The miR-17~92-cluster was discovered in 2005 and was the first cluster shown to have an oncogenic potential. In B-cell lymphomas, high expression miR-17~92 acted with c-MYC expression, thereby enhancing tumorigenesis.[80] Hence, it was given the name “Oncomir-1”.

Expression of these miRs promotes cell proliferation, suppress apoptosis of malignant cells, and induces tumor angiogenesis. Physiologically, the miR-17~92 cluster has essential roles in normal development of the heart, lungs, and immune system, and is highly expressed in embryonic cells. In miR-17~92-deficient mice, severely hypoplastic lungs and ventricular septal defects were observed. An increase in apoptosis specific to the B-cell compartment was also observed during fetal development.[81]

1.2.4 miR-17

In breast and prostate cancer, miR-17 can act as a tumor suppressor by targeting AIB1 (amplified in breast cancer 1) and PCAF (p300/CBP-associated factor), respectively.[82,83] On the contrary, miR-17 has been shown to be significantly over-expressed in triple-negative breast cancers (being the most aggressive breast cancer)[84], and miR-17 has been shown to induce prostate tumor growth and invasion by regulating TIMP3 (TIMP metallopeptidase inhibitor 3).[85] In cervical cancer, miR-17 function as a tumor suppressor, targeting TP53INP1 (tumor protein p53-induced nuclear protein 1).[86] miR-17 inhibits melanoma growth by stimulating CD8+ T-cells mediating a host immune response, due to its regulation of the transcription factor STAT3 (signal transducer and activator of transcription 3).[87] In gastric cancer, high circulating levels of miR-17-5p (and miR-20a) was associated with poor tumor differentiation status and tumor progression. High expression levels of miR17-5p/20a were significantly correlated with poor overall survival.[88]

In a review on gastrointestinal cancers, most studies reported overexpression of miR-17, and that this over-expression was associated with impaired survival.[89] Noteworthy, the authors writing the review found that half of the publications on the prognostic impact of miR-17 in gastrointestinal cancers were non-significant. A Danish group found that in colon cancer, upregulation of miR-17-expression takes place early in the normal-adenoma-adenocarcinoma sequence.[90] There was significantly increased expression in low grade adenomas *vs* normal epithelium, as well as in high grade adenomas *vs* low grade adenomas. They did not, however, find an increase in expression from high grade adenomas to adenocarcinomas.

Potential targets are numerous in gastrointestinal cancer. In HCC, miR-17-5p-induced phosphorylation of HSP27 enhance migration of HCCs through the p38 MAPK pathway.[91] miR-17-5p downregulates TGFBR2 expression in gastric cancer, promoting cell growth and migration.[92] In pancreatic cancer, over-expression of miR-17-5p leads to inactivation of the tumor suppressor PTEN (phosphatase and tensin homolog), enhancing cell survival.[93] PTEN is also a potential target in colorectal cancer, where inhibition of PTEN leads to activation of the AKT/PI3K-pathway. PTEN-inhibition has been hypothesized to be the mode of action, in which miR-17-5p is involved in cytotoxic drug resistance among colorectal cancer patients.[94] miR-17-5p inhibits the transcription factor E2F1, another known tumor suppressor.[90]

1.2.5 miR-20a

miR-20a is considered an oncomiR participating in cell proliferation and cancer progression and is dysregulated in various cancer types. As a member of the miR-17 family, targets and mechanisms of action of miR-20 are closely related to those of miR-17-5p. The oncogene MYC induces the miR-17 family, dysregulating cell cycle progression, apoptosis, and tumor invasion via interactions with PTEN, E2F genes and the TGF-β pathway.[95,96] Although proposed mechanisms of action are many.

In non-small cell lung cancer (NSCLC), high plasma levels of miR-20a correlated with shorter DFS and/or OS. Over-expression of miR-20a also promoted radio-resistance in NSCLC patients.[97] In gastric cancer, similar to miR-17-5p, elevated plasma levels of miR-20a correlated with poor overall survival and was associated with tumor differentiation status and tumor progression.[88] In hepatocellular carcinoma (HCC), miR-20a was shown to be a tumor suppressor, and low expression correlated with significantly lower recurrence-free survival and overall survival. In HCC cell lines, the Mcl-1 (myeloid cell leukemia sequence 1) protein was identified as a direct target of miR-20a, an antiapoptotic member of the Bcl-2 family, protecting cells from apoptosis and tumor carcinogenesis.[98] In oral squamous cell carcinoma (OSCC), miR-17 and miR-20a were negatively correlated with TNM stage and lymphatic metastasis. In OSCC cell lines, miR-17/20a were shown to inhibit OSCC cell migration through a potential pathway mediated by ITGβ8 (integrin β8).[99]

In colorectal cancer cell lines, miR-20a is shown to be up-regulated. Over-expression of miR-20a-5p resulted in loss of drosophila mothers against decapentaplegic protein4 (SMAD4), switching the effect of TGF-β from a tumor suppressor to a tumor promoter.[100] This promoted the invasion and metastasis of colorectal cancer cells and induced epithelial-mesenchymal-transition (EMT). It was also found to participate in cell proliferation, cell apoptosis and invasion by targeting the BID- and TRAIL-pathways, among others.[101]

1.2.6 miR-126

miR-126 is located within the 7th intron of epidermal growth factor like domain 7 (EGFL7) on chromosome 9.[102] The host gene itself, EGFL7, is one of the major targets of miR-126. Transcription of EGFL7 and miR-126 occurs simultaneously. Mature miR-126 binds to EGFL7 to prevent its translation. The result is a decrease in EGFL7 protein levels,

mediated by a negative feedback mechanism.[103] It is the most highly enriched miR in endothelial cells and mediates developmental angiogenesis *in vivo*. Intracellular inhibitors of angiogenic signaling (Vascular endothelial growth factor (VEGF)-pathway) are repressed by miR-126, thereby promoting blood-vessel formation by enhancing the pro-angiogenic actions of VEGF and fibroblast growth factors (FGF). Knockdown of miR-126 in zebrafish resulted in loss of vascular integrity and hemorrhage during embryonic development, proving that miR-126 is essential for the maintenance of vascular integrity in vivo.

Dysregulation of miR-126 is seen in various cancers. In most cancers, miR-126 functions as a tumor suppressor. In esophageal cancer, reduced expression of miR-126 was linked with cancer dedifferentiation and lymph node metastasis.[104] In oral squamous cell carcinoma, low expression of miR-126 was associated with cancer progression, lymph node metastasis and poor prognosis.[105] The metalloproteinase disintegrin and metalloproteinase domain containing protein 9 (ADAM9) is highly expressed in pancreatic cancer. ADAM9 is a target of miR-126, and low expression of miR-126 resulted in high expression of ADAM9, causing invasive growth of pancreatic cancer cells related to epithelial-mesenchymal-transition (EMT).[106] In small cell lung cancer, miR-126 over-expression induced delayed G1 phase of the cancer cell cycle, whereas in breast cancer the G1/G0 to S phase was inhibited.[107,108] In gastric cancer, results are more conflicting. miR-126 over-expression has been shown to reduce gastric cancer cell proliferation by inducing cell cycle arrest in G0/G1 phase, reducing migration and invasion in vitro, and reducing carcinogenesis and metastasis in vivo.[109] Another study in gastric cancer showed that miR-126 inhibits sex-determining region Y-box 2 (SOX2), contributing to gastric carcinogenesis.[110] In patient-derived prostate cancer xenograft rat models and cell lines, it was shown that miR-126 promotes metastasis in prostate cancer.[111] In ovarian cancer patients, elevated levels of miR-126 were found in the sera of these patients compared to healthy controls.[112] In colorectal cancer, miR-126 normally function as a tumor suppressor.

1.3 Protein biomarkers

Assessing the expression of miRs is useful and could supply prognostic value. To further understanding of these miRs, we wanted to assess some of the proteins believed to be regulated by them. Following a literature search, we decided to concentrate on RUNX3, SMAD4, IRS-1 and IRS-2, believed to be correlated to the miRs.

1.3.1 RUNX3

The Runt domain transcription factors (RUNX), RUNX1, RUNX2 and RUNX 3, are integral components of signaling cascades mediated by both TGF-β and BMPs in numerous important biological systems.[113] RUNX3 is located at locus 1p36 in humans, a region frequently deleted in many types of cancer, therefore postulated to contain an important tumor suppressor gene. RUNX3 is closely related to gastric cancer, where lack of RUNX3 function was causally related to the genesis and progression of cancer.[114] Dysfunctional RUNX3 is seen in several cancer types (GI cancers, lung cancer, breast cancer, prostate cancer)[115], and RUNX3 can be inactivated in several ways. It could be loss of heterozygosity, mutations, methylation-related transcriptional silencing, and mislocalization of RUNX3 to the cytoplasm. In a previous study in colorectal cancer, expression of RUNX3 in the nucleus has been interpreted as the functional form, whereas exclusive cytoplasmic expression has been interpreted as an inactive form.[115]

1.3.2 SMAD4

SMADs are a group of proteins transducing extracellular signals to the nucleus. In mammalian cells we have 8 different SMADs which can be divided into three classes. (1) The receptor-regulated SMADs (R-SMADs) comprising SMAD1, SMAD2, SMAD3, SMAD5, and SMAD8. SMAD 2 and 3 mainly mediate signaling from TGF-β subfamily members, whereas SMAD1, 5 and 8 mediates signaling from csp subfamily members. (2) The common-mediator SMAD (Co-SMAD), SMAD4, the central mediator of both TGF-β and BMP signaling pathways. (3) The inhibitory SMADs (I-SMADs), comprising SMAD6 and SMAD7, counteracting the effects of R-SMADs. Phosphorylated R-SMADs form

heteromeric complexes with SMAD4 which are translocated to the nucleus, where it positively or negatively regulates gene expression.[116]

The SMAD4 gene is located on chromosome 18q21.[117] It is also known as the deleted in pancreatic cancer 4(dpc4) gene, as approximately half of pancreatic cancers show inactivation of the gene.[118] Genetic alteration of the SMAD4 gene were found in at least 26 cancer types, but more frequently in GI tract cancers (pancreas, colorectal, stomach and esophagus).[119] Inactivation of the SMAD4 gene correlates well with a loss of expression of its protein.[120] Its value as a prognostic factor in colorectal cancer has been assessed earlier. In Dukes C patients, a high expression of was related to improved OS and DFS compared to low expression. This was true whether the patients were given adjuvant chemotherapy or not.[121,122] Similar findings were reported in a CRC cohort comprising Dukes A-C patients, where a low expression of SMAD4 was correlated to poor prognosis.[123]

Several pathways have been reported to interact with the TGF-β/SMAD4 pathway, including MAPK, PI3K/AKT and WNT/β-catenin pathways among others.[119]

1.3.3 IRS-1 and IRS-2

The insulin receptor substrate (IRS) proteins are cytoplasmic adaptor proteins mediating the functional outcomes in insulin signaling.[124] They are involved in normal growth, metabolism, survival, and differentiation. IRS-1 and IRS-2 are primary mediators of insulin-dependent mitogenesis and regulation of glucose metabolism, expressed ubiquitously in humans. IRS-1 is encoded on human chromosome 2q36-37, whereas IRS-2 is encoded on human chromosome 13q34.1.[125,126] IRSs have been found to be oncogenic themselves, but are also required for the transforming ability of several other oncogenes. IRS-1 has been associated with tumor growth and proliferation, whereas IRS-2 is more associated with tumor motility and invasion.[124] IRS-1 and IRS-2 expression is increased in several cancer types (pancreas, prostate, HCC among others), whereas only IRS-1 is increased in breast, ovarian and medulloblastomas. IRS-1 expression has been reported to be decreased in NSCLC and poorly differentiated breast cancers. The IRSs have been characterized as typical cytosolic adaptor proteins, but has been shown to translocate to the nucleus in medulloblastomas and breast cancer.[127]

1.4 Prognostic *vs* predictive biomarkers

A biomarker is a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes or responses to an exposure or intervention. This definition is broad, and may be derived from molecular, histologic, radiographic or physiologic characteristics. The definitions of predictive and prognostic biomarkers were nicely explained by the BEST (Biomarkers, EndpointS, and other Tools) Resource, and the following quotes are from their last updated publication.[128]

"A predictive biomarker is defined by the finding that the presence or change in the biomarker predicts an individual or group of individuals more likely to experience a favorable or unfavorable effect from the exposure to a medical product or environmental agent."

"A prognostic biomarker is used to identify the likelihood of a clinical event, disease recurrence, or disease progression in patients with a disease or medical condition of interest."

In colon cancer you can find examples of both predictive and prognostic biomarkers. An example of a predictive biomarker is the presence of a KRAS/NRAS mutation in metastatic disease. These patients are expected to have no clinical benefit of treatment with an EGFR-inhibitor. Metastatic patients with no KRAS/NRAS mutation (wild-type) are expected to benefit from treatment with an EGFR-inhibitor.[46] The presence of lymph node metastasis is a prognostic biomarker. These patients have an impaired survival compared with patients without lymph node involvement.[129] Hence, these patients are routinely offered adjuvant chemotherapy.

2 AIMS OF THESIS

The ultimate aim of the thesis is to further the understanding of colon cancer. The ultimate aim is comprised of several sub goals.

i. Set up a robust retrospective biobank for colon cancer patients treated with curative intent (stage I-III). Thus, enabling us to pursue prognostic biomarker studies in this patient group.
ii. Set up and conduct biomarker studies. We aimed to separate the biomarker signals from the tumor and stromal compartments utilizing *in situ* hybridization and immunohistochemistry. More specifically we aimed to investigate the prognostic impact of the expression of miR-126 (Paper I), miR-17-5p and miR-20a-5p (Paper II) and some of their downstream effectors, namely IRS-1, IRS-2, RUNX3 and SMAD4 (Paper III).
iii. For miR-17-5p and miR-20a-5p we wanted to validate and elucidate functional aspects through *in vitro* experiments in select colon cancer cell lines (Paper II).

3 MATERIALS AND METHODS

3.1 Data collection

At the beginning of the data collection, a colleague doing his surgical internship and myself were sharing the work collecting the data. We planned to split the workload 50/50. As he was doing this work in his spare-time meanwhile completing his internship, he quickly realized he didn't have time to participate as first intended. Because of this, I ended up collecting most of the data for our clinical database myself.

From 2001 onwards clinical journals were digitized, but analog, handwritten medical journals were the standard prior to that. As we were collecting data from 1998-2007, this included a substantial amount of analog medical journals. Luckily, most hospitals had converted their analog journals to a digitized version by scanning them. In Tromsø and Harstad this work wasn't completed, so I had to go through analog medical journals there. Fortunately, I was able to get away with just one day of travel to Harstad to complete the database – the remaining work could be done from my office in Tromsø.

As previously mentioned, there were some years of standstill between the initial data-collection and the completion of the clinical database. A positive impact was that we ended up with a clinical database with substantially longer follow-up, as all survival-data was updated before the finalization of the database.

3.2 Patient cohort

All patients operated and diagnosed with colon cancer in Northern Norway in the time-period 1998-2007, whose resected tumor specimen was stored in the archive at the Department of Clinical Pathology at the University Hospital of North Norway (UNN), were identified retrospectively (n=861). This included patients who were operated in Tromsø, Narvik, Harstad, Hammerfest, Kirkenes, Stokmarknes, Gravdal, Mo i Rana, and Mosjøen. Inclusion and exclusion criteria are listed in Table 2. Primary tumor blocks were collected, and a database including demographic and clinicopathological data was established. Patients were de-identified, but an access-restrained key was kept facilitating follow-up.

***Table 2**: Inclusion and exclusion criteria.*

INCLUSION CRITERIA

- Age >18 years old
- Histologically verified colon adenocarcinoma
- TNM stage I-III
- Adequate tissue blocks for TMA construction
- Medical records available

EXCLUSION CRITERIA

- Malignancy within last 5 years before colon cancer diagnosis/synchronous malignancy (excluding non-melanoma skin cancer)
- Postoperative mortality <90 days
- TNM stage IV (metastatic disease)
- Missing tissue blocks
- Lost to follow-up

The database was completed December 1, 2017. Of the initially identified 861 patients, a substantial number of these (n=409) were excluded from the database. Of these, 117 patients were excluded due to metastatic disease at the time of diagnosis, 79 had missing tissue-blocks or tissue-blocks were inadequate for TMA construction, 57 were wrongly coded (mainly rectal cancers), 55 were operated outside Northern Norway and were sent for genetic testing, 46 had a prior malignancy within the last 5 years before diagnosis (or synchronous malignancy), 22 died within the first 90 days after surgery, 18 had surgery for recurrent cancer or were treated in a strictly palliative setting, 7 were lost to follow-up (3 of them tourists), 3 appeared twice, 2 were operated at a regional hospital not participating in the study, 2 were operated before the actual time frame, and 1 had missing medical journals. Hence, 452 patients were finally included in the database (Figure 7). Patient characteristics are presented in **Table 6**, section 4.1.

In our material, we decided to omit the rectal cancers and concentrate on colon cancers only. The reason for this, is that a large proportion of rectal cancers undergo neoadjuvant radiochemotherapy prior to surgery. We suspected that this treatment would alter the expression of the prognostic biomarkers we wanted to explore in the tumor tissue.

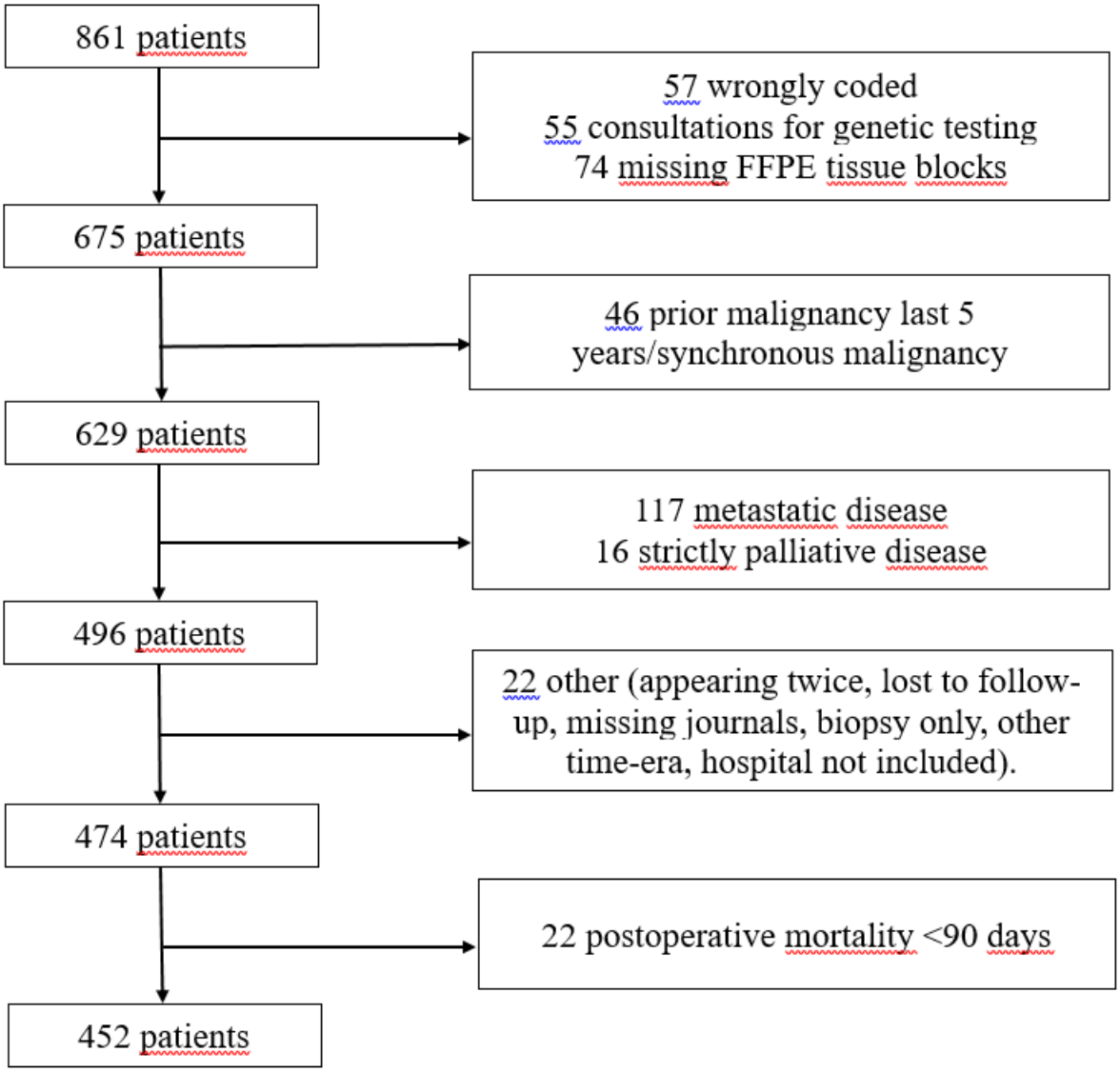

***Figure 9**: Flowchart of the inclusion of patients in the study population.*

3.3 Tissue microarray

Tissue microarray (TMA) is a widely used, high-throughput method of performing molecular analyses. The concept of a multitumor tissue block was introduced by Battifora in 1986 (the "sausage block").[130] The first modern TMA concept and design was described in 1998, where donor tissue cores were spaced evenly and systematically in a grid-like formation in empty, recipient TMA-blocks.[131] TMAs may be produced from formalin-fixed, paraffin-embedded (FFPE) tissue, frozen tissue[132], paraffin-embedded cell lines[133] or cell blocks.[134] Hundreds of cores can be placed on a recipient block, and each recipient block can subsequently be sectioned several hundred times.[135]

3.3.1 TMA construction

All colon cancer cases were histologically reviewed by two experienced pathologists, Vidar Isaksen and Lill-Tove Rasmussen Busund. The most representative paraffin donor blocks were selected for each case, and two areas of neoplastic epithelial cells and two from tumor stroma were marked on hematoxylin and eosin (H&E) slides to guide sampling for the recipient TMA blocks. Areas of necrosis were avoided. Using a tissue arraying instrument (Beecher Instruments, Silver Springs, MD, USA), a thin-walled biopsy needle and stylet created holes in a recipient paraffin block, and 0.6 mm diameter cylindrical tissue cores from the donor block were sampled and transferred to the recipient block at defined array coordinates. Normal tissue distant from the primary tumor was included in the TMAs, serving as control for tissue staining. Non-symmetric arrays were constructed to enable orientation. A total of 15 TMA blocks were constructed, where 12, 2 and 1 included primary tumors, metastatic lymph nodes and controls, respectively. One TMA block containing primary tumors was constructed in a symmetrical manner, making orientation of the slide impossible and was therefore omitted. Multiple 4-µm sections were cut with a Micron microtome (HM355S) and further processed for in situ hybridization analysis. A schematic of TMA construction is presented in **Figure 10**.

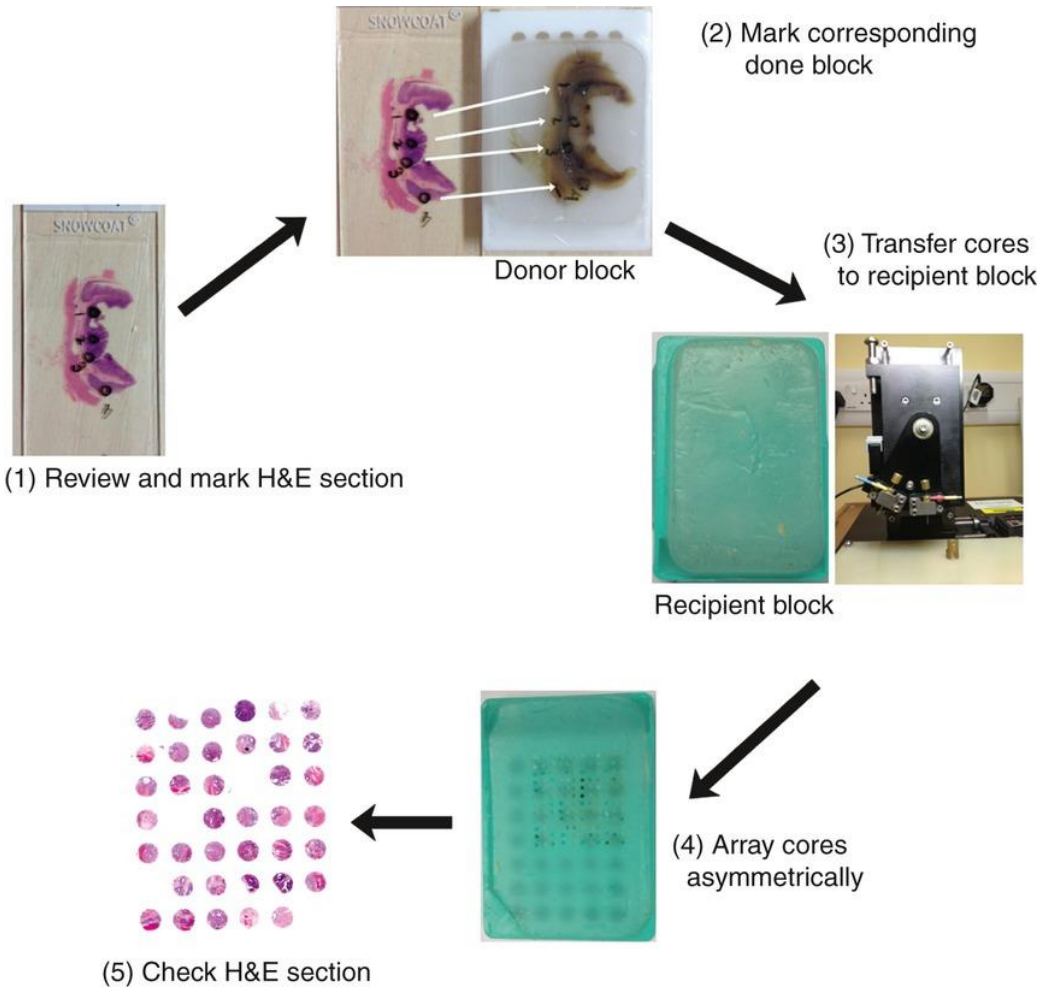

***Figure 10**: A schematic of TMA construction. Representative, small cores are taken from the donor block, and are inserted into recipient blocks. Hundreds of cores can be arrayed onto one recipient block. (Adapted from Ilyas et al with permission.*[136]*)*

3.3.2 TMA considerations

The use of TMAs in cancer research has several advantages compared to working with whole section specimens. Quantities of reagents are reduced, and you know that the staining protocol and conditions are identical for each patient core included in the TMA. As the cores are carefully selected by an experienced pathologist and prepared by a dedicated technician, this makes it possible for a non-specialist to score the core. With the addition of digitization, substantial amounts of tumor cores can be scored fast, given that you have the right classifiers, scripts, and cut-offs. The main challenge would be tumor heterogeneity. TMAs include only a small fraction of a macroscopic tumor, and in case of heterogenous tumors, expression of assessed biomarkers could differ considerably across the tumor. To

overcome this obstacle, larger tissue cores or increasing the number of cores have been suggested, but prior studies have shown that two to four 0.6 mm cores are sufficient.[136] Main advantages and disadvantages are listed in table 3.

***Table 3**: Advantages and disadvantages with TMA technology in cancer research. (Adapted from Tom Dønnem with permission.)*

Advantages	Disadvantages
Time saving, high-throughput method	TMA construction is time consuming, and may be technically challenging
Archived tissue can be used	Less suited for heterogeneous tumors
Increased utilization of limited tissue resources	Some cores are expected to be missing, reducing statistical power
Standardized experimental conditions, reducing variability	Variation through the length of the tissue core
Scoring by less experienced personnel	Not suited for individual diagnosis
Possibility to share with other institutions	
Well suited for scoring by digital pathology	

3.4 Digital pathology

Light microscopy has been the main tool for pathologists for a long time. Now, the field of pathology is being digitized. In our research group, earlier studies on biomarker expression have been conducted by scoring the TMA cores in a manual, semi-quantitative manner (visual interpretation of the expression of the biomarker as absent, low, moderate, or high). In this work, we wanted to use digital pathology. TMA slides were constructed like previously described, but once constructed they were scanned on a Pannoramic 250 FLASH III slide scanner (3DHistech, Budapest, Hungary), yielding a digital representation of the slide. A TMA map was made by the technician constructing the TMAs based on pathology report numbers, and these numbers were de-identified making a new TMA map with unique numbers for all patients in Excel. TMA slides were processed in QuPath v0.1.3 for Paper I

and Paper II, and v.0.3.2 for Paper III. TMAs were de-arrayed and preprocessed according to Bankhead et al.[137] TMA maps were transferred from Excel to QuPath, linking the TMA slides in QuPath to the clinical database. The detailed workflow for the different biomarkers, can be found in the respective papers. The scripts used are available in the supplementary files for the respective articles. In paper III, we used Deep Learning (DL) for classifying the TMA slides.

3.5 Immunohistochemistry

Immunohistochemistry (IHC) is a technique for biomarker detection used in tissue-based diagnostics. In cancer diagnostics, it is most used as an addition to morphologic evaluation as a diagnostic tool. It has been widely adopted for assessment of *in situ* protein expression. IHC as we know it today has been widely used in medical research and diagnostics since the mid 1990's.[138] Briefly, IHC involves sequential application of a primary antibody capable of specifically binding epitopes of a given antigen, a secondary reporter-coupled antibody specific for the primary antibody, an enzyme complex and a chromogenic substrate. The enzymatic activation of the chromogen allows for visualization of the antigen-antibody-complex. The method is inexpensive, widely available, may be performed on archival tissue and makes assessment of biological molecules in cells from different tissue compartments possible. When using IHC on TMAs, you have the advantage of standardized antigen retrieval, temperature, incubation time, washing procedures and reagent concentration – avoiding inter-batch variability.

3.5.1 Antibodies

Antibodies are proteins serving as one of the principal effectors of the adaptive immune system. They are widely used as a diagnostic and research reagent, and even as a therapeutic tool. The selection of a suitable antibody is a crucial step when conducting an IHC based study.[139] Primary antibodies are either mono- or polyclonal.

Polyclonal antibodies are generated by immunizing animals (e.g., rabbits), yielding a mixture of antibodies directed against various epitopes of the given antigen. Polyclonal antibodies generate higher detection sensitivity, although at risk for more cross-reactivity. Consequently, false negative results are rarely seen using polyclonal antibodies.

Monoclonal antibodies are developed from hybrids and supply antibodies against only one antigen epitope, making them homogenous and consistent, resulting in higher specificity.[140] Antibodies used in this thesis were monoclonal and are listed in **Table 6**.

***Table 4**: Single stain IHC protocol.*

Baking	Deparaffination	Antigen retrieval	Primary Antibody	Secondary Multimer	Detection	Counterstain
Tissue 8 min 60°C	Discovery wash 3x12 min 68°C	CC1 40 min 95°C	IRS1 1:800 32 min 36°C	OmniMap anti-Rb HRP 16 min 37°C	ChromoMap DAB/H_2H_2 8/4 min 37°C	Hem II/Bluing 24/8 min 37°C min
Tissue 8 min 60°C	Discovery wash 3x12 min 68°C	CC1 40 min 95°C	IRS2 1:100 32 min 36°C	OmniMap anti-Rb HRP 16min 37°C	ChromoMap DAB/H_2H_2 8/4 min 37°C	Hem II/Bluing 24/8 min 37°C
Tissue 8 min 60°C	Discovery wash 3x12 min 68°C	CC1 40 min 95°C	Smad4 1:200 60 min. 36°C	OmniMap anti-Rb HRP 16 min 37°C	ChromoMap DAB/H_2H_2 8/4 min 37°C	Hem II/Bluing 24/8 min 37°C
Tissue 8 min 60°C	Discovery wash 3x12 min 68°C	CC1 40 min 95°C	RUNX3 1:400 60 min. 36°C	OmniMap anti-Ms HRP 16 min 37°C	ChromoMap DAB/H_2H_2 8/4 min 37°C	Hem II/Bluing 24/8 min 37°C
-	Discovery wash 3x12 min 68°C	CC1 32 min 95°C	CD3 Predil 16 min.	OmniMap anti-Rb HRP 16 min 37°C	ChromoMap DAB/H_2H_2 8/4 min 37°C	Hem II/Bluing24/8 min 37°C
-	Discovery wash 3x12 min 68°C	CC1 64 min 95°C	CD8 Predil 32 min.	OmniMap anti-Rb HRP 16 min 37°C	ChromoMap DAB/H_2H_2 8/4 min 37°C	Hem II/Bluing24/8 min37°C

3.6 *In situ* hybridization

In situ hybridization (ISH) is a technique used to localize and detect specific DNA and RNA sequences in tissue or cells, using labelled complementary DNA or RNA strands (probes). The hybridized probe and sequence may be visualized microscopically, depending

on probe labelling (antigen- or fluorescent-labeled bases).[141] The probes used in this thesis are listed in **Table 5**. Chromogenic ISH (CISH) can be assessed on light microscopy and was the chosen method for this thesis. Fluorescent ISH (FISH) on the other hand, requires assessment on a fluorescence microscope.

Investigation of nucleic acids by *in situ* hybridization has been used for a long time, and was first reported in 1969.[142] To analyze the detailed spatial expression of miRs was for long considered technically challenging due to their small size and the general fragile nature of RNAs. These obstacles were overcome by the invention of using locked nucleic acid (LNA)-modified DNA probes.[143]

By utilizing ISH, we were able to detect the expression of biomarkers in different cells in different compartments. This was crucial as we wanted to explore the prognostic impact of biomarker expression in tumor epithelium *vs* stroma.

***Table 5**: Probes used for ISH.*

Probes	Reference	Company
hsa-miR-126-3p LNA probe	619866-360	Exiqon
hsa-miR-17-5p LNA probe	619852-360	Exiqon
hsa-mir-20a-5p LNA probe	611011-360	Exiqon
MirCURY LNA ISH miRNA ISH buffer and control set. LNA U6 snRNA probe, pos.control LNA Scramble MiR, neg. control	339459	Exiqon

3.7 Cut-off values

For many biomarkers, there exists no common reference standard for assessing their expression. Standardization is difficult due to differences in tissue preparation and antigen

retrieval, among others. The assessment and cut-off determination are left to the investigators, making comparison of results challenging.[144] The expression of our selected biomarkers varies over a continuous scale and is not a matter of positive *vs* negative expression. When assessing the expression level of biomarkers, you need to decide a proper cut-off value. We decided to dichotomize the expression into low *vs* high expression. Using the mean value as a cut-off has been employed in many studies. By using the mean value, you risk losing valuable biological information, leading to possible false negative results (type 2 error). However, the risk of false positive results decreases (type 1 error), and results are more reproducible. For these explorative analyses, we chose to use the optimal cut-off, that is the cut-off giving the most difference in DSS between groups. This data-dependent "optimal" cut-off makes comparison of biomarkers across studies difficult, as it may result in many different optimal cut-offs. The type 1 error rate can be high using this approach.[145]

3.8 *In vitro* studies

3.8.1 Colon cancer cell lines

In paper II we performed *in vitro* experiments investigating the functional aspects of miR-17-5p and miR-20a-5p in colon cancer tumorigenesis. We used two colon cancer cell lines: CACO-2 (ATCC HTB-37) and HT-29 (ATCC-38).[146,147] They are both derived from colon adenocarcinomas and form low-grade/early-stage cancer when grown in nude mice. Their importance is not limited to the study of human cancers.

The CACO-2 cell line was established in 1977 from a primary colon tumor. Over the years, CACO-2 cells have acquired different properties due to different culture conditions and different numbers of passages. The CACO-2 cell line has been widely used as a model of the intestinal epithelial barrier, mimicking the small intestine. It has often been used for pharmacological and toxicological studies.[148]

The HT-29 cell line was established in 1964 from a primary colon tumor. Like CACO-2 it expresses characteristics of mature intestinal cells. It has been an important cell line in studies focused on food digestion and bioavailability, and additionally to study the intestinal epithelial response to bacterial infection.[149]

3.8.2 Cell cultures

The detailed experimental procedures regarding cell culture conditions, cell transfection and in vitro experiments, are presented in paper II. To sum it up briefly: 1) culturing of cell lines in Opti-MEM I (1x) medium without phenol red (cat.#11058-021, Gibco, RF, UK) with 5% fetal bovine serum and 1% Penicillin Streptomycin, 2) incubation in a humidified atmosphere with 5% CO2 at 37°C for 72 h, 3) transfection of cell lines with miR-17-5p or miR-20a-5p mimic in combination with miR negative control. The transfection reagent Lipofectamine RNAiMAX was utilized. Transfection efficiency was estimated to be 80-95%.

3.8.3 Viability assay

Cells were cultured in 96-well plates and incubated with 12 mM of MMT. Formazan crystals were solubilized by adding 0,01 M HCl/SDS and the absorbance was measured in a CLARIOstar plate reader at 570 nm.

3.8.4 Migration/wound healing assay

Cells were grown in 24-well plates, washed with PBS, and incubated in serum free medium with Mitomycin C to avoid cell proliferation. Cells were "wounded" by a sterile pipette, then washed to remove detached cells and debris. Cells were transfected after 4 hours. Photos were taken of wounded areas at 0 and 24 hours in controls and transfected cells. Areas occupied by migrating cells after 24 hours were calculated by subtracting background levels at 0 hours.

3.8.5 Invasion assay

Cells were seeded in ThincertR chambers with polyethylene terephthalate membranes precoated with phenol red-free Matrigel and placed in 24-well plates containing FBS in the lower chamber. In the upper chambers cells were transfected and incubated for 48 hours at 37°C. Chambers were washed, fixed in paraformaldehyde, and stained with crystal violet. Non-invading cells from the upper membrane were removed by a cotton swab. Membranes

containing the invaded cells (under the membrane surface) were photographed. Duplicate images of three random microscope fields were captured, and areas of cell invasion were determined using Image J software.

3.9 Statistical analyses

The initial clinical database was constructed using IBM SPSS, versions 21.0-24.0 (IBM Corp., Armonk, NY, USA). Relevant clinical data from medical journals and important data from histology reports were included in this database. The last update of the database was completed on December 1, 2017.

For paper I, statistical methods were performed using SPSS version 26.0 and R version 3.6.3., while in paper II and III, Rstudio 2021.09.0 build 351 (RStudio PBC) using R version 4.0.4. was used.

The statistical methods used were similar for all three papers (paper I-III). In paper II and III, expression of molecular markers was rescaled to a range between 0 and 1 using max-min scaling. To examine the correlations between the molecular marker expression and clinicopathological variables, the Chi-square ($x^{2)}$ and Fisher´s exact tests were used. The chosen endpoint, disease-specific survival (DSS), was defined as the interval from surgery to the time of colon cancer death. For univariate analyses, the Kaplan-Meier method was used, visualizing associations between molecular marker expression and survival. The log-rank test was used to assess the statistical differences between the survival curves. For the multivariate analyses, they were performed using a backward conditional Cox regression analysis with a probability for stepwise entry and removal at 0.05 and 0.10. P-values <0.05 were considered statistically significant. In paper III, to examine the associations between marker expressions, Pearson correlation was used. *r* values were classified as 0 (negative), 0-0.2 (weak), 0.2-0.3 (weak/moderate), 0.3-0.5 (moderate), 0.5-0.7 (moderate/strong) and > 0.7 (strong). Hierarchical clustering with distance calculated based on the *r* values of their correlations was applied to visualize patterns in the correlation data.

4 MAIN RESULTS

4.1 Patient characteristics

Clinical and pathological variables are presented in **Table 6**. The median age at diagnosis was 74 years (range 30-94 years). Median follow-up of survivors was 173 months (range 119-239 months). There were slightly more women compared to men (53.8 vs 46.2), and approximately 35% of all patients presented with lymph node involvement.

Table 6 *- Frequency table including important clinicopathological variables. For categorical variables, five -year overall survival is given in percent.*

Variable	n (%)	5-y OS
Age at diagnosis		
Median	74.0 years (30-94 years)	N/A
Tumor size		
Median	50 mm (10-180 mm)	N/A
Missing	2	N/A
Gender		
Male	209 (46.2%)	59.8%
Female	243 (53.8%)	62.1%
Relapse		
No	333 (73.7%)	74.5%
Yes	119 (26.3%)	23.5%
pT status		
T1	9 (2.0%)	88.9%
T2	76 (16.8%)	65.8%
T3	320 (70.8%)	61.6%
T4a	26 (5.8%)	46.2%
T4b	21 (4.6%)	42.9%
pN status		
N0	291 (64.4%)	68.4%
N1a	56 (12.4%)	50.0%
N1b	55 (12.2%)	52.7%
N1c	1 (0.2%)	0%
N2a	27 (6.0%)	40.7%
N2b	22 (4.9%)	40.9%
pTNM stage (groups)		
I	72 (15.9%)	70.8%
II	219 (48.5%)	67.6%
III	161 (35.6%)	47.8%
Differentiation		
Well	36 (8.0%)	80.6%
Moderate	329 (72.8%)	59.6%
Poor	75 (16.6%)	60.0%
Undifferentiated	4 (0.9%)	25.0%
Missing	8 (1.8%)	N/A
Site		
Right	227 (50.2%)	63.0%
Transverse	65 (14.4%)	49.2%
Left	21 (4.6%)	66.7%
Sigmoid	136 (30.1%)	61.8%
Missing	3 (0.7%)	N/A
Weight loss		
<10%	248 (54.9%)	68.5%
≥10%	94 (20.8%)	46.8%
Missing	110 (24.3%)	N/A
Performance status (ECOG)		
0	237 (52.4%)	70.0%
1	149 (33.0%)	58.4%
2	54 (11.9%)	33.3%
3	8 (1.8%)	25.0%
Missing	4 (0.9%)	N/A
Adjuvant chemotherapy		
No	365 (80.8%)	60.3%
Yes	87 (19.2%)	67.8%
Postoperative complications		
No	358 (79.2%)	64.5%
Yes	94 (20.8%)	47.9%

4.2 Paper I

In different cancers, miR-126 can act as both an oncogene and as a tumor suppressor. We wanted to explore the prognostic impact of miR-126 in different compartments (tumor *vs* stroma) in resected colon cancer patients. We wanted to assess this utilizing ISH and digital pathology.

4.2.1 Expression and correlations

We observed a moderate/strong correlation between miR-126 expression in tumor and stroma (r = 0.60). A low miR-126 expression in stroma was associated with increasing pathological stage and histological grade.

4.2.2 Univariate analyses

Age, weight loss, pathological stage, histological grade, vascular infiltration, and resection margins were all significantly associated with DSS. We also observed that a high expression of miR-126 in both tumor and stroma were significant indicators of improved DSS ($p < 0.001$ and $p = 0.005$, respectively).

4.2.3 Multivariate analyses

Age, histological grade, and pathological stage were independently associated with DSS. High miR-126 expression in tumor was also independently associated with DSS (HR 0.45, CI 0.27-0.76, $p = 0.002$).

4.3 Paper II

miR-17-5 p and miR-20a-5p are both members of the miR-17-92 cluster, mainly believed to have oncogenic effects. Like paper I, we wanted to explore the prognostic impact of these miRs in the same colon cancer cohort. The method used was identical, utilizing ISH and digital pathology. In addition, we performed cell-line experiments assessing the effects of

miR-17-5p and miR-20a-5p expression on viability, invasion, and migration in two early-stage colon cancer cell lines.

4.3.1 Expression and correlations

High expression of miR-17-5p and miR-20a-5p in tumor tissue was associated with well and moderately differentiated tumors. miR-20a-5p was associated with cancers of the right colon. Correlation between the different miRs were also observed; miR-17-5p in tumor was correlated with miR-17-5p in stroma (weak/moderate, $r = 0.27$) and miR-20a-5p in both tumor and stroma (moderate/strong, $r = 0.52$, and weak, $r = 0.17$, respectively); miR-17-5p in stroma was correlated to miR-20a-5p in both tumor and stroma (weak, $r = 0.16$, and moderate, $r = 0.37$, respectively); miR-20a-5p was correlated with miR-20a-5p in stroma (moderate/strong, $r = 0.65$).

4.3.2 Univariate analyses

High expression of miR-17-5p in tumor was a significant indicator of DSS ($p = 0.002$), as well as high expression of miR-20a-5p in both tumor and stroma ($p = 0.035$ and $p = 0.003$, respectively).

4.3.3 Multivariate analyses

High expression of miR-17-5p in tumor was independently associated with favorable DSS (HR = 0.43, CI 0.26-0.71, $p < 0.001$), and high expression of miR-20a-5p in both tumor (HR = 0.60, CI 0.37-0.97, $p = 0.037$) and stroma (HR 0.63, CI 0.42-0.95, $p = 0.027$).

4.4 Paper III

In paper III we wanted to assess proposed targets of the previously explored miRs; miR-126, miR-17-5p and/or miR-20a-5p. More specifically, we wanted to explore the prognostic impact of IRS-1, IRS-2, RUNX3, and SMAD4 in both stromal and tumor tissue.

4.4.1 Expression and correlations

SMAD4 and RUNX3 were expressed both in nucleus and cytoplasm. IRS-1 and IRS-2 were expressed in the cytoplasm. IRS-1, IRS-2 and SMAD4 were evenly expressed in tumor epithelial cells, spindle shaped cells/stromal cells and immune cells. RUNX3 was expressed mainly in tumor infiltrating lymphocytes (TILs), and in some other cells. Cytoplasmic expression of SMAD4 was correlated with pathological stage, while RUNX3 expression was correlated with site. The epithelial and stromal expression for each marker showed moderate to strong correlations. Extensive correlations between IRS-1, IRS-2, RUNX3 and SMAD4 in tumor and stroma were observed ($0.15 < r < 0.60$). CD3+ and CD8+ TILs density showed moderate/strong correlations with stromal RUNX3 ($0.35 < r < 0.60$).

4.4.2 Univariate analyses

Several factors were all significant predictors of a favorable DSS: 1) Increased expression of SMAD4 in nucleus and cytoplasm in the tumor epithelial compartment, and in cytoplasm in the stromal compartment. 2) Increased expression of RUNX3 in the nucleus or cytoplasm in both the tumor epithelial and stromal compartments. 3) Increased expression of IRS-1 in stromal cytoplasm.

4.4.3 Multivariate analyses

Increased tumor epithelial expression of SMAD4 in cytoplasm (HR 0.58, 95% CI 0.43-0.80, $p < 0.001$) and RUNX3 in nucleus (HR 0.62, 95% CI 0.45-0.84, $p = 0.002$) were independent positive predictors of DSS. In the stromal compartment, increased expression of IRS-1 in cytoplasm (HR 0.64, 95% CI 0.47-0.87, $p = 0.005$), SMAD4 in cytoplasm (HR 0.67, 95% CI 0.5-0.91, $p = 0.009$), and RUNX3 in cytoplasm (HR 0.62, 95% CI 0.44-0.87, $p = 0.006$) were independent predictors of a favorable DSS.

4.4.4 Co-expression analyses

Patients with increased/preserved SMAD4 and RUNX3 expression in either the tumor epithelium (HR 0.34, 95% CI 0.18-0.66, $p = 0.001$) or the stromal compartment (HR 0.34,

95% CI 02-0.57, $p < 0.001$) had significantly better prognoses compared to those with decreased/lost expression.

5 DISCUSSION

5.1 A summary of strengths and weaknesses

There are several strengths in this study. The patient cohort is a large, consecutive colon cancer patient cohort treated with curative intent. This makes the risk of selection bias low, as we have included patients operated at all hospitals in Northern-Norway (except Bodø) during the specified period. We also have a long follow-up of all patients, with a median follow-up of survivors of 173 months. In addition, relevant clinicopathological variables were obtained by an experienced clinician directly from the patients' medical journals.

For the TMAs, duplicate cores were collected for both tumor and stromal tissue. The TMA collection and the ISH and IHC procedures were performed by dedicated and experienced technicians.

The assessment of protein and miR expression was performed using digital pathology. Historically, this has been performed using light microscopy scoring the expression in a semi-quantitative manner, usually with two scorers. The introduction of digital pathology makes this process reproducible, and it eliminates the interobserver variability. With time, it will also be less time-consuming, as you are able to score several TMA-slides in one operation.

When choosing cut-off values, we ended up using optimal cutoffs, thereby reducing the chance of false negative results (type 2 errors).

Regarding weaknesses, the retrospective design of the study is one. In the collection of the clinical data, this limited the collection of variables to what is found in the regular medical journals. In the case of a prospective design, any variable of interest would have been clarified at the beginning, and more variables of interest could have been added. The retrospective design also limited the information found in the pathological reports. These reports were adequate at the time of surgery (1998-2007), but today they would have been classified as insufficient. This can be highlighted by looking at the information on vascular infiltration, which was a statistically significant prognosticator for DSS in univariate analyses. Information about the presence of vascular infiltration or not was introduced as a mandatory part of the histological report midway through the investigated period. Consequently, about

half of the tumors are missing this information. We considered complementing the histology reports with this information by reevaluation by a pathologist in missing cases. Unfortunately, we didn't find a pathologist with sufficient time available to do so, meaning this information is missing for 52% of tumors. Using optimal cut-offs for our analyses, increases the chance of false positive results (type 1 errors).

5.2 General discussion

The quality of the database is crucial for good research. Retrospective collection of data has some limitations, introducing potential information bias. Data are extracted from medical journals, which may have several variables of interest missing. Cause of death can be biased by subjective interpretation. This information should be easily available from the Norwegian Cause of Death Registry, but based on prior discussions within our research group, this information has not been considered reliable, and we decided to omit this. We only had medical journals from the involved hospitals available (including all radiographic examinations) but assumed that patients with a relapse of their disease would have been in contact with their local hospital. Regarding death, this information is available for all the patients. Patients too old or in severely reduced performance status may have been taken care of by their general practitioner in a strictly palliative setting, but the number of such patients is probably negligible. As far as no information about relapse was found in hospital journals by the time of death (or in autopsy findings), cause of death was classified as "other cause".

As previously mentioned, Norway has one of the highest incidences of colon cancer in the world relatively to population. The reason for this is largely unknown, but environmental exposures, genetic susceptibility and Western lifestyle are probably influential. The large number of colon cancer patients gives us a unique possibility to perform research on this group of patients. In our patient cohort, we included resected colon cancer patients, stage I-III, operated over a 10-year period. Whether these patients are offered adjuvant chemotherapy to eradicate potential microscopic disease post-surgery depends on prognostication according to the TNM-classification. This prognostication is imprecise, and still there are numerous relapses among patients in different stages. In our cohort, we wanted to explore the prognostic impact of several biomarkers not currently used in the clinical setting. Regarding our project, this was a retrospective study without treatment interventions, meaning we were solely looking for prognostic biomarkers. Could we find one biomarker, or more, that was able to better stratify the patients' risk of recurrence, complementing the existing TNM-stratification, guiding the clinical oncologist upon decision-making regarding adjuvant chemotherapy.

Why do we need to alter adjuvant therapy? Obviously, we want to improve survival for the patients if possible. This could be the case for stage II patients without high-risk features, who would not be candidates for adjuvant chemotherapy as of today's guidelines. If they were proven to be at high risk of recurrence with the supplementary precision a new

biomarker provides, they could be offered adjuvant chemotherapy, thereby reducing the risk of recurrent disease. But improved survival isn't always what we are trying to achieve. Adjuvant chemotherapy is a potential harmful treatment as all drugs, especially chemotherapeutic drugs, come with potential side effects. For adjuvant chemotherapy in colon cancer, polyneuropathy because of platinum derivates, and cardiotoxicity because of fluoropyrimidines, are common and unwanted side-effects.[150,151] The last study to implement a change in the guidelines for adjuvant chemotherapy, was in fact a deintensification in stage III patients with T1-3N1-disease.[38] This study was a non-inferiority study, and it proved that three months of treatment was non-inferior to six months of treatment for this group of patients. This will be beneficial for the patients, as they will receive less treatment and experience fewer side-effects. It will also reduce costs. Not only the costs of the omitted 3 months of chemotherapy, but also the costs for the health-care system managing the potential side-effects.

In our first paper we were looking at the prognostic impact of miR-126. In our material, we concluded that a high expression of miR-126 in tumor tissue is an independent positive prognostic factor in stage I-III colon cancer. Regarding miR-126 expression in tumor, 5-year disease-specific survival (DSS) increased from 75% to 88% for low *vs* high expression, respectively. This corroborates findings from previous studies. In most cancers, miR-126 is considered a tumor suppressor. When assessing the prognostic impact of miR-126 expression, a high expression was related to improved prognosis for non-small cell lung cancer, breast cancer, pancreatic cancer and colorectal cancer.[152] On the contrary, in gastric cancer cell lines, over-expression of miR-126 promotes tumor growth.[153] This highlights some of the complexity of the prognostic value of miR-expression analyses. Different miRs have many targets, and these targets can be influenced by several miRs as well as other factors. It has been shown that not all validated target mRNAs are affected by a specific miR in every cell type.[154] The function of the miR relies on both the cell type and the cellular environment. miRs simultaneously exert competing oncogenic and tumor-suppressive effects, and the question is whether the miR has a net oncogenic or net tumor suppressive effect.[63]

When comparing our results with earlier studies, several aspects must be kept in mind. In our patient cohort, we present exclusively colon cancer patients. In many studies used for comparison, they have used colorectal cancer cohorts. Looking at three review articles on miR-126, half of the cited articles assessed colorectal cancers (11 out of 22).[152,155,156] Of interest, a previous article showed that the expression of miR-126 differs between colon and

rectal cancers.[157] This highlights that direct comparisons must be interpreted with caution. Another aspect to consider is the stage of the disease. All our patients were treated with curative intent, meaning stage I-III colon cancer, and metastatic disease was considered an exclusion criterium. Several of the other studies included metastatic patients (stage IV). It appears likely that expression in non-metastatic patients differ from metastatic patients. Previous studies have shown that the expression of miR-126 is significantly lower in patients with multiple metastatic lesions compared to patients with a single metastatic lesion.[158] On the contrary, Ebrahimi et al. showed similar levels of expression in primary lesions compared with metastatic lesions.[159] We also need to take demographic variations into consideration. There are both genetic and environmental differences between people from different continents. In colorectal cancer, comparing African American and non-Hispanic Caucasian patients, miR-expression profiles differed significantly between the two groups.[160] Comparable results were found in prostate cancer patients.[161] Numerous miR-expression studies come from Asian research groups, and given the ethnical differences shown in the beforementioned studies, caution should be taken when comparing different demographic groups.

The prognostic value of miRs as single biomarkers has been questioned. To improve the prognostic value, several miRs have been integrated into panels. Zhang et al. investigated six miRs combined as a prognostic tool for disease recurrence in stage II colon cancer (miR-21-5p, miR-20a, miR-103a-3p, miR-106b-5p, miR-143-5p, miR-215).[162] In a cohort of 735 Chinese patients, patients were classified into low *vs* high risk of recurrence based on the expression of these miRs. 5-year DFS was 89% *vs* 60%, 85% *vs* 57%, and 85% *vs* 54% (all groups highly statistically significant) in the initial training group, in the internal training set and an independent validation set, respectively. The six-miR-based classifier had a better prognostic value than known clinicopathological risk factors and mismatch repair status. In an ad hoc analysis, patients in the high-risk group showed favorable response to adjuvant chemotherapy, making it a potential tool for deciding which stage II patients to offer adjuvant chemotherapy. This experiment was tested in a material of 71 white patients by Caritg *et al.*[163] Interestingly, only three of the examined miRs were significantly associated with DFS in the white population (miR-103a-3p, miR-143-5p and miR-215). Elevated levels of miR-103a and low levels of miR-143-5p correlated to worse prognosis in both cohorts. Elevated levels of miR-215 was associated to impaired prognosis in the white cohort but was found to predict a better outcome in the Chinese cohort. This could support hypotheses that divergent results are

due to ethnic differences among populations, as shown previously.[160] The remaining three miRs were independent prognostic markers on multivariate analyses and were equally good as six to classify patients into low- and high-risk groups.

In paper II, we were looking at the prognostic impact of two members of the proposed oncogenic miR-17~92 cluster, namely miR-17-5p and miR-20a-5p.[164] Surprisingly, we did not observe an oncogenic effect when these miRs were over-expressed. On the contrary, we found an improved survival among patients with high expression of miR-17-5p and miR-20a-5p, suggesting a tumor suppressive effect. These results contrasts findings in previous studies in gastrointestinal cancers, where over-expression of miR-17-5p was related to poor OS and DFS.[89] For miR-20a-5p, over-expression was seen in the majority of colon cancer patients, and results showed a tendency towards impaired survival, but non-significantly.[165] To investigate their functional aspects, we decided to use the cell lines CACO-2 and HT-29, both derived from primary colon adenocarcinomas. Although their use is not restricted to cancer research, we believe that these cell lines represent early-stage colon in a better way than choosing a cell line from a metastatic lesion. When the investigated miRs were over-expressed in these cell lines, we did not observe any signs of increased aggressiveness or metastatic capacity, corroborating our findings where increased expression of the miRs related to improved survival. To get even more robust results, more cell lines could have been included, but this is a trade-off concerning price and time-consumption, among others.

In our first two papers, we were assessing the prognostic impact of three different miRs (miR-126, miR-17-5p, and miR-20a-5p) utilizing *in situ* hybridization. For our third paper, we wanted to explore the prognostic impact of some of the proposed targets for these miRs utilizing immunohistochemistry.

Following a literature search, we decided to explore the expressions of RUNX3, SMAD4, IRS-1 and IRS-2. Based on previous studies in different cancer forms, miR-126 was associated with IRS-1, miR-17-5p associated with RUNX3, and miR-20a-5p associated with both RUNX3 and SMAD4.[100,108,166–170] In a study on gastric cancer, Song et al. demonstrated that miR-17-5p was upregulated in gastric cancer, and that RUNX3 was downregulated in gastric cancer tissues.[167] Their finding that RUNX3 was downregulated in gastric cancer, thereby suppressing proliferation and invasiveness, correlates well with our findings on RUNX3 in colon cancer which relates to better prognosis. Their finding regarding upregulation of miR-17-5p on the other hand, also promoting proliferation and invasion via

suppressing apoptosis, contrasts our finding that a high expression of miR-17-5p was beneficial in colon cancer. In a study in colorectal cancer, Cheng et al. were assessing expression of miR-20a-5p and SMAD4.[171] In both cell-line and nude mice experiments, they demonstrated that a high expression of miR-20a-5p resulted in downregulation of SMAD4. The net effect was increased invasion and metastatic capacity in colorectal cancer cells. Their results indicate a positive effect of high expression of SMAD4, corroborating our findings and the fact that SMAD4 is a tumor suppressor in colon cancer. Their conclusion that miR-20a-5p is an onco-miR and that high expression predicts a poor prognosis for colorectal cancer patients, is opposite to our findings where we found high expression of miR-20a-5p to be related with improved prognosis for colon cancer patients. miR-20a-5p is also associated with RUNX3. In a study in human hepatocellular carcinoma, Chen et al. demonstrated that over-expression of miR-20a-5p resulted in downregulation of RUNX3, contributing to increased proliferation and migration of HCC cells. Once again supporting the tumor suppressive effect of RUNX3 but contrasting findings with an oncogenic effect of miR-20a-5p.

In routine clinical practice, surgically removed tissues are fixed with formalin, embedded in paraffin, and sectioned with a microtome yielding 4-5 μm-thick paraffin sections. Dewaxed sections are stained with hematoxylin and eosin, and potentially used for specific purposes (*in situ* hybridization, immunohistochemistry, special stains etc.).[172] Thereafter, a pathologist examines the section by the use of light microscopy. The pathologist then confirms a histological diagnosis based on visual recognition, semi-quantification, and integration of multiple morphological features in the context of the medical history of the patient. In the field of oncology, the histopathological diagnosis is delivered to the oncologist, with the appropriate treatment being initiated based on this.

Assessing slides by light microscopy is a time-consuming activity for the pathologists, and the requirements for the histopathological reports are getting more complex. The cancer incidence is steadily increasing, mainly because of the aging population.[1] This results in increasing amounts of biopsies and surgical specimens in the years to come, further increasing the workload for the pathologists. To ease the burden on the pathologists, introducing digital pathology might be helpful. Digital pathology is useful in many settings, like quantifying tumor infiltrating immune cells, assessing tumor differentiation grade, counting mitoses among others, with the main target of predicting patient outcome.[173–175] By introducing machine learning techniques, maybe these intermediate proxies for outcome could be

bypassed. Could machine learning directly learn the prognostic relevant features in microscopy images of the tumor? In colorectal cancer, several groups have explored this hypothesis. Bychkov et al. used a TMA-based colorectal cancer material stained with hematoxylin and eosin, and wanted to explore if a deep learning algorithm could convert these small areas of tumor tissue to a predicted outcome for the patients.[176] Patients were stratified into low- and high-risk groups. They demonstrated that such an approach can outperform a visual histological prognostic assessment based on TMA material as well as based on histological grading on whole slide tissue sections. Similarly, Skrede et al. also explored deep learning in a colorectal cancer cohort.[177] One important difference is that they used whole slide images that were divided into tiles, reducing the challenge of tumor heterogeneity. They also reported a deep learning model that could be clinically useful in dividing patients into poor *vs* good prognosis. It will be interesting to see whether these results can be reproduced in a prospective trial that is planned.

The introduction of deep learning models in the healthcare system is controversial. The lack of explanation of how the software works, as it is considered a "black-box" lacking interpretability, is a big challenge for its implementation.[178] In pathology, the assistance of deep learning algorithms improved the histopathologic interpretation of lymph node metastases in breast cancer. When the pathologists used an algorithm-assisted approach, where a deep learning algorithm identified areas likely to contain tumor, this approach outperformed either the algorithm or the pathologist alone.[179] So, this kind of an approach can aid the pathologists in their decision-making, as they would use less time finding the areas of interest in the tumors, but still a trained pathologist is very much needed.

6 CONCLUSIONS AND IMPLICATIONS FOR FURTHER RESEARCH

Being diagnosed with cancer is a terrifying experience for most patients. The word *cancer* itself is intimidating, and often associated with death. Luckily, prognoses have improved for cancer patients in general, and overall, 3 out of 4 cancer patients are now cured from their disease. But even for the ones who get cured, the risk of recurrence will forever haunt them. Every time an ache appears in your body, the same question pops up in your mind; "Could this be the first sign of cancer recurrence?". This is psychologically demanding, and it could be difficult to find the balance between not thinking too much about the risk of recurrence, and keeping the risk of recurrence in the back of your head in order to take proper action when indicated. As medical oncologists, our job is to reduce this risk of recurrence as much as possible - at a reasonable prize. To be successful at this, we need good biomarkers – both prognostic and predictive.

To tailor the treatment for colon cancer patients, the risk stratification needs to be improved. Today's risk stratification, relying mostly on the AJCC-TNM classification, has limitations differentiating within large subgroups (stage II vs stage III). Adjuvant chemotherapy is routinely offered to patients fit for chemotherapy in stage II disease with high-risk features, and stage III disease. Chemotherapy regimens are similar, consisting of 5-FU based chemotherapy, with or without the addition of Oxaliplatin. Treatment duration ranges from 3-6 months. While there are predictive biomarkers in metastatic colon cancer (targeted therapies depending on mutational status), only mismatch repair deficiency (dMMR) matters in the choice of adjuvant treatment. Patients with dMMR-tumors respond poorly to 5-FU chemotherapy, and should not get 5-FU as monotherapy. Hence, we need to rely on prognostic biomarkers for decision making regarding adjuvant treatment.

In this work, we show promising results for new prognostic biomarkers in stage I-III colon cancer. We found that a high expression of miR-126, miR-17-5p and miR-20a-5p predicts improved disease specific survival for these patients. Additionally, an increased expression of RUNX3 and SMAD4 are related to improved disease specific survival in the same cohort. This information can help us to identify the correct treatment for each patient. This could be to offer adjuvant chemotherapy to stage II patients without high-risk features who wouldn't be offered this treatment as of today's guidelines. It could also mean to identify

stage III patients with a low risk of recurrence, who potentially could end up with 3 months of chemotherapy instead of 6 months, or even completely without chemotherapy.

Our results are based on observations from retrospective studies on archived tissue. For this to be implemented in routine clinical practice, it needs validation in large prospective clinical trials. As stated in Paper III, results regarding co-expression of RUNX3 and lymphocytes, particularly in the stromal compartment, are very interesting. This will be further elucidated using multiplex immunohistochemistry in the future, and can perhaps be of importance in other cancer types, not only in colon cancer.

Works cited

1. Cancer Registry of Norway. *Cancer in Norway 2021 - Cancer incidence, mortality, survival and prevalence in Norway.* (2021).
2. Cancer, I. A. for R. on. GLOBOCAN 2020. (2020). Available at: https://gco.iarc.fr/today/data/factsheets/cancers/39-All-cancers-fact-sheet.pdf.
3. Tebala, G. D. History of colorectal surgery: A comprehensive historical review from the ancient Egyptians to the surgical robot. *Int. J. Colorectal Dis.* **30**, 723–748 (2015).
4. Global Cancer Observatory. Available at: https://gco.iarc.fr/. (Accessed: 12th January 2021)
5. Cancer registry of Norway. *Cancer incidence, mortality, survival and prevalence in Norway: Clustering of cancer*. (2019).
6. Araghi, M. *et al.* Changes in colorectal cancer incidence in seven high-income countries: a population-based study. *Lancet Gastroenterol. Hepatol.* **4**, 511–518 (2019).
7. Vuik, F. E. R. *et al.* Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. *Gut* **68**, 1820 (2019).
8. Giovannucci, E. Modifiable risk factors for colon cancer. *Gastroenterol. Clin. North Am.* **31**, 925–943 (2002).
9. Platz, E. A. *et al.* Proportion of colon cancer risk that might be preventable in a cohort of middle-aged US men. *Cancer Causes Control* **11**, 579–588 (2000).
10. Lu, L., Mullins, C. S., Schafmayer, C., Zeißig, S. & Linnebacher, M. A global assessment of recent trends in gastrointestinal cancer and lifestyle-associated risk factors. *Cancer Commun.* **41**, 1137 (2021).
11. Gausman, V. *et al.* Risk Factors Associated With Early-Onset Colorectal Cancer. *Clin. Gastroenterol. Hepatol.* **18**, 2752-2759.e2 (2020).
12. Jasperson, K. W., Tuohy, T. M., Neklason, D. W. & Burt, R. W. Hereditary and Familial Colon Cancer. *Gastroenterology* **138**, 2044–2058 (2010).
13. Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).
14. Keum, N. N. & Giovannucci, E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 713–732 (2019).
15. Guttmacher, A. E., Collins, F. S., Lynch, H. T. & De La Chapelle, A. Hereditary Colorectal Cancer. *https://doi.org/10.1056/NEJMra012242* **348**, 919–932 (2003).
16. Hamilton, W., Round, A., Sharp, D. & Peters, T. J. Clinical features of colorectal cancer before diagnosis: A population-based case-control study. *Br. J. Cancer* **93**, 399–405 (2005).
17. Helsedirektoratet. Nasjonalt handlingsprogram med retningslinjer for dianostikk, behandling og oppfølging av kreft i tykktarm og endetarm. (2019). Available at: https://www.helsedirektoratet.no/retningslinjer/kreft-i-tykktarm-og-endetarm-

handlingsprogram/IS 2849 Nasjonalt handlingsprogram kreft i tykktarm og endetarm.pdf/_/attachment/inline/4a5fa48e-8d76-4618-98b3-43af5a85b76e:4c4a29f71e7a68ff93a19dd82848f36a49ab.

18. Jecius, H. *et al.* Emergent colectomy for colorectal cancer: A comparative analysis of open vs. minimally invasive approach. *Am. J. Surg.* (2022). doi:10.1016/J.AMJSURG.2022.10.019
19. Holtedahl, K. Challenges in early diagnosis of cancer: the fast track. *http://www.manuscriptmanager.com/sjphc* **38**, 251–252 (2020).
20. Health, T. N. D. of. Cancer patient pathway for colorectal cancer. (2016). Available at: https://www.helsedirektoratet.no/nasjonale-forlop/tykk-og-endetarmskreft.
21. Cancer Registry of Norway. ColorectalScreen Norway. (2022). Available at: https://www.kreftregisteret.no/screening/tarmscreening/.
22. Fitzpatrick-Lewis, D. *et al.* Screening for Colorectal Cancer: A Systematic Review and Meta-Analysis. *Clinical Colorectal Cancer* **15**, 298–313 (2016).
23. Brierley, J. D., Gospodarowicz, M. K. & Wittekind, C. *TNM Classification of Malignant Tumours, 8th Edition*. (2016).
24. Dienstmann, R., Salazar, R. & Tabernero, J. Personalizing colon cancer adjuvant therapy: Selecting optimal treatments for individual patients. *J. Clin. Oncol.* **33**, 1787–1796 (2015).
25. Mehrkhani, F., Nasiri, S., Donboli, K., Meysamie, A. & Hedayat, A. Prognostic factors in survival of colorectal cancer patients after surgery. *Color. Dis.* **11**, 157–161 (2009).
26. McClelland, D. & Murray, G. I. A Comprehensive Study of Extramural Venous Invasion in Colorectal Cancer. *PLoS One* **10**, e0144987 (2015).
27. Becerra, A. Z. *et al.* Evaluating the Prognostic Role of Elevated Preoperative Carcinoembryonic Antigen Levels in Colon Cancer Patients: Results from the National Cancer Database. *Ann. Surg. Oncol.* **23**, 1554–1561 (2016).
28. Lugli, A. *et al.* Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Mod. Pathol.* **30**, 1299–1311 (2017).
29. Ferlitsch, M. *et al.* Colorectal polypectomy and endoscopic mucosal resection (EMR): European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline. *Endoscopy* **49**, 270–297 (2017).
30. Smith, A. J. *et al.* Guideline for optimization of colorectal cancer surgery and pathology. *J. Surg. Oncol.* **101**, 5–12 (2010).
31. Hohenberger, W., Weber, K., Matzel, K., Papadopoulos, T. & Merkel, S. Standardized surgery for colonic cancer: complete mesocolic excision and central ligation – technical notes and outcome. *Color. Dis.* **11**, 354–364 (2009).
32. Amri, R., Bordeianou, L. G., Sylla, P. & Berger, D. L. Association of Radial Margin Positivity With Colon Cancer. *JAMA Surg.* **150**, 890–898 (2015).

33. Laurie, J. A. *et al.* Surgical adjuvant therapy of large-bowel carcinoma: An evaluation of levamisole and their combination of levamisole and fluorouracil. *J. Clin. Oncol.* **7**, 1447–1456 (1989).
34. Moertel, C. G. Fluorouracil plus Levamisole as Effective Adjuvant Therapy after Resection of Stage III Colon Carcinoma: A Final Report. *Ann. Intern. Med.* **122**, 321 (1995).
35. André, T. *et al.* Oxaliplatin, Fluorouracil, and Leucovorin as Adjuvant Treatment for Colon Cancer. *N. Engl. J. Med.* **350**, 2343–2351 (2004).
36. André, T. *et al.* Adjuvant Fluorouracil, Leucovorin, and Oxaliplatin in Stage II to III Colon Cancer: Updated 10-Year Survival and Outcomes According to BRAF Mutation and Mismatch Repair Status of the MOSAIC Study. *J. Clin. Oncol.* **33**, 4176–87 (2015).
37. Böckelman, C., Engelmann, B. E., Kaprio, T., Hansen, T. F. & Glimelius, B. Risk of recurrence in patients with colon cancer stage II and III: A systematic review and meta-analysis of recent literature. *Acta Oncol. (Madr).* **54**, 5–16 (2015).
38. Grothey, A. *et al.* Duration of Adjuvant Chemotherapy for Stage III Colon Cancer. *N. Engl. J. Med.* **378**, 1177 (2018).
39. van der Geest, L. G. M. *et al.* Nationwide trends in incidence, treatment and survival of colorectal cancer patients with synchronous metastases. *Clin. Exp. Metastasis* **32**, 457–465 (2015).
40. Cook, A. D., Single, R. & McCahill, L. E. Surgical resection of primary tumors in patients who present with stage IV colorectal cancer: An analysis of surveillance, epidemiology, and end results data, 1988 to 2000. *Ann. Surg. Oncol.* **12**, 637–645 (2005).
41. Riihimaki, M., Hemminki, A., Sundquist, J. & Hemminki, K. Patterns of metastasis in colon and rectal cancer. *Sci. Rep.* **6**, (2016).
42. Panettieri, E. *et al.* Survival by Number and Sites of Resections of Recurrence after First Curative Resection of Colorectal Liver Metastases. *J. Gastrointest. Surg.* **26**, 2503–2511 (2022).
43. Oweira, H., Mehrabi, A., Reissfelder, C. & Abdel-Rahman, O. A Real-World, Population-Based Analysis of the Outcomes of Colorectal Cancer Patients with Isolated Synchronous Liver or Lung Metastases Treated with Metastasectomy. *World J. Surg.* **44**, 1604–1611 (2020).
44. Milosevic, M. *et al.* Pulmonary Metastasectomy in Colorectal Cancer: updated analysis of 93 randomized patients – control survival is much better than previously assumed. *Color. Dis.* **22**, 1314–1324 (2020).
45. Van Cutsem, E. *et al.* ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann. Oncol.* **27**, 1386–1422 (2016).
46. Price, T. J. *et al.* Panitumumab versus cetuximab in patients with chemotherapy-refractory wild-type KRAS exon 2 metastatic colorectal cancer (ASPECCT): A randomised, multicentre, open-label, non-inferiority phase 3 study. *Lancet Oncol.* **15**, 569–579 (2014).

47. Kabbinavar, F. F. *et al.* Combined analysis of efficacy: The addition of bevacizumab to fluorouracil/leucovorin improves survival for patients with metastatic colorectal cancer. *J. Clin. Oncol.* **23**, 3706–3712 (2005).
48. Venook, A. P. *et al.* Effect of First-Line Chemotherapy Combined With Cetuximab or Bevacizumab on Overall Survival in Patients With KRAS Wild-Type Advanced or Metastatic Colorectal Cancer: A Randomized Clinical Trial. *JAMA* **317**, 2392–2401 (2017).
49. Kopetz, S. *et al.* Encorafenib, Binimetinib, and Cetuximab in BRAF V600E–Mutated Colorectal Cancer . *N. Engl. J. Med.* **381**, 1632–1643 (2019).
50. Grothey, A. *et al.* Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet* **381**, 303–312 (2013).
51. Mayer, R. J. *et al.* Randomized Trial of TAS-102 for Refractory Metastatic Colorectal Cancer. *N. Engl. J. Med.* **372**, 1909–1919 (2015).
52. Pfeiffer, P. *et al.* TAS-102 with or without bevacizumab in patients with chemorefractory metastatic colorectal cancer: an investigator-initiated, open-label, randomised, phase 2 trial. *Lancet Oncol.* **21**, 412–420 (2020).
53. Jin, Z. & Sinicrope, F. A. Mismatch Repair-Deficient Colorectal Cancer: Building on Checkpoint Blockade. *J. Clin. Oncol.* **71**, (2022).
54. Chalabi, M. *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat. Med. 2020 264* **26**, 566–576 (2020).
55. Verschoor, Y. L. *et al.* Neoadjuvant nivolumab, ipilimumab, and celecoxib in MMR-proficient and MMR-deficient colon cancers: Final clinical analysis of the NICHE study. *https://doi.org/10.1200/JCO.2022.40.16_suppl.3511* **40**, 3511–3511 (2022).
56. Le, D. T. *et al.* Phase II open-label study of pembrolizumab in treatment-refractory, microsatellite instability–high/mismatch repair–deficient metastatic colorectal cancer: KEYNOTE-164. in *Journal of Clinical Oncology* **38**, 11–19 (American Society of Clinical Oncology, 2020).
57. Kostouros, A. *et al.* Large intestine embryogenesis: Molecular pathways and related disorders (Review). *Int. J. Mol. Med.* **46**, 27 (2020).
58. Hui, C. *et al.* Rectosigmoid Cancer—Rectal Cancer or Sigmoid Cancer? *Am. J. Clin. Oncol.* **45**, 333 (2022).
59. Cercek, A. *et al.* PD-1 Blockade in Mismatch Repair-Deficient, Locally Advanced Rectal Cancer. *N. Engl. J. Med.* **386**, 2363–2376 (2022).
60. Zhang, B., Pan, X., Cobb, G. P. & Anderson, T. A. microRNAs as oncogenes and tumor suppressors. *Developmental Biology* **302**, 1–12 (2007).
61. Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes

small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–854 (1993).

62. 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.* **99**, 15524–15529 (2002).
63. Svoronos, A. A., Engelman, D. M. & Slack, F. J. OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer. *Cancer Res.* **76**, 3666 (2016).
64. Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231–235 (2004).
65. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
66. Gregory, R. I., Chendrimada, T. P., Cooch, N. & Shiekhattar, R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631–640 (2005).
67. Meijer, H. A., Smith, E. M. & Bushell, M. Regulation of miRNA strand selection: follow the leader? *Biochem. Soc. Trans.* **42**, 1135–1140 (2014).
68. Bartel, D. P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* **116**, 281–297 (2004).
69. Jevsinek Skok, D. *et al.* Genome-wide in silico screening for microRNA genetic variability in livestock species. *Anim. Genet.* **44**, 669–677 (2013).
70. Lin, S. & Gregory, R. I. MicroRNA biogenesis pathways in cancer. *Nature Reviews Cancer* **15**, 321–333 (2015).
71. Bobbili, M. R., Mader, R. M., Grillari, J. & Dellago, H. OncomiR-17-5p: alarm signal in cancer? *Oncotarget* **8**, 71206–71222 (2017).
72. Mogilyansky, E. & Rigoutsos, I. The miR-17/92 cluster: A comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death and Differentiation* **20**, 1603–1614 (2013).
73. Kosaka, N., Iguchi, H. & Ochiya, T. Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis. *Cancer Sci.* **101**, 2087–2092 (2010).
74. AHMED, F. E. *et al.* Diagnostic MicroRNA Markers to Screen for Sporadic Human Colon Cancer in Stool: I. Proof of Principle. *Cancer Genomics Proteomics* **10**, (2013).
75. Volinia, S. *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2257–2261 (2006).
76. Motoyama, K. *et al.* Over- and under-expressed microRNAs in human colorectal cancer. *Int. J. Oncol.* **34**, 1069–1075 (2009).
77. Yu, G. *et al.* Prognostic values of the miR-17-92 cluster and its paralogs in colon cancer. *J. Surg. Oncol.* **106**, 232–237 (2012).
78. Li, X. M., Wang, A. M., Zhang, J. & Yi, H. Down-regulation of miR-126 expression in colorectal cancer and its clinical significance. *Med. Oncol.* **28**, 1054–1057 (2011).

79. Tanzer, A. & Stadler, P. F. Molecular evolution of a microRNA cluster. *J. Mol. Biol.* **339**, 327–335 (2004).
80. He, L. *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **435**, 828–833 (2005).
81. Ventura, A. *et al.* Targeted Deletion Reveals Essential and Overlapping Functions of the miR-17∼92 Family of miRNA Clusters. *Cell* **132**, 875–886 (2008).
82. Nagtegaal, I. D., Quirke, P. & Schmoll, H. J. Has the new TNM classification for colorectal cancer improved care? *Nature Reviews Clinical Oncology* **9**, 119–123 (2012).
83. Gong, A. Y. *et al.* miR-17-5p targets the p300/CBP-associated factor and modulates androgen receptor transcriptional activity in cultured prostate cancer cells. *BMC Cancer* **12**, 492 (2012).
84. Moi, L., Braaten, T., Al-Shibli, K., Lund, E. & Busund, L. T. R. Differential expression of the miR-17-92 cluster and miR-17 family in breast cancer according to tumor type; Results from the Norwegian Women and Cancer (NOWAC) study. *J. Transl. Med.* **17**, 334 (2019).
85. Yang, X. *et al.* Both mature miR-17-5p and passenger strand miR-17-3p target TIMP3 and induce prostate tumor growth and invasion. *Nucleic Acids Res.* **41**, 9688 (2013).
86. Wei, Q., Li, Y.-X., Liu, M., Li, X. & Tang, H. MiR-17-5p targets TP53INP1 and regulates cell proliferation and apoptosis of cervical cancer cells. *IUBMB Life* **64**, 697–704 (2012).
87. Li, H., Gupta, S., Du, W. W. & Yang, B. B. MicroRNA-17 inhibits tumor growth by stimulating T-cell mediated host immune response. *Oncoscience* **1**, 531–539 (2014).
88. Wang, M. *et al.* Circulating miR-17-5p and miR-20a: Molecular markers for gastric cancer. *Mol. Med. Rep.* **5**, 1514–1520 (2012).
89. Wang, Z. *et al.* Prognostic value of miR-17-5 p in gastrointestinal cancers: a systematic review and meta-analysis. *Onco. Targets. Ther.* **11**, 5991–5999 (2018).
90. Knudsen, K. N. *et al.* MicroRNA-17 is the most up-regulated member of the miR-17-92 cluster during early colon cancer evolution. *PLoS One* **10**, (2015).
91. Yang, F. *et al.* miR-17-5p Promotes migration of human hepatocellular carcinoma cells through the p38 mitogen-activated protein kinase-heat shock protein 27 pathway. *Hepatology* **51**, 1614–1623 (2010).
92. Qu, Y. *et al.* MiR-17-5p regulates cell proliferation and migration by targeting transforming growth factor-β receptor 2 in gastric cancer. *Oncotarget* **7**, 33286 (2016).
93. Gu, J. *et al.* GFRα2 prompts cell growth and chemoresistance through down-regulating tumor suppressor gene PTEN via Mir-17-5p in pancreatic cancer. *Cancer Lett.* **380**, 434–441 (2016).
94. Fang, L. *et al.* MicroRNA-17-5p promotes chemotherapeutic drug resistance and tumour metastasis of colorectal cancer by repressing PTEN expression. *Oncotarget* **5**, 2974–2987 (2014).
95. O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839–843 (2005).

96. Sokolova, V. *et al.* The Effects of miR-20a on p21: Two Mechanisms Blocking Growth Arrest in TGF-β-Responsive Colon Carcinoma. *J. Cell. Physiol.* **230**, 3105–3114 (2015).
97. Xu, X., Zhu, S., Tao, Z. & Ye, S. High circulating miR-18a, miR-20a, and miR-92a expression correlates with poor prognosis in patients with non-small cell lung cancer. *Cancer Med.* **7**, 21–31 (2018).
98. Fan, M. Q. *et al.* Decrease expression of microRNA-20a promotes cancer cell proliferation and predicts poor survival of hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* **32**, 21 (2013).
99. Chang, C. C. *et al.* MicroRNA-17/20a functions to inhibit cell migration and can be used a prognostic marker in oral squamous cell carcinoma. *Oral Oncol.* **49**, 923–931 (2013).
100. Cheng, D. *et al.* MicroRNA-20a-5p promotes colorectal cancer invasion and metastasis by downregulating Smad4. *Oncotarget* **7**, 45199–45213 (2016).
101. Huang, G., Chen, X., Cai, Y., Wang, X. & Xing, C. miR-20a-directed regulation of BID is associated with the TRAIL sensitivity in colorectal cancer. *Oncol. Rep.* **37**, 571–578 (2017).
102. Wang, S. *et al.* The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. *Dev. Cell* **15**, 261–71 (2008).
103. Sun, Y. *et al.* miR-126 inhibits non-small cell lung cancer cells proliferation by targeting EGFL7. *Biochem. Biophys. Res. Commun.* **391**, 1483–1489 (2010).
104. Hu, Y. *et al.* Prognostic significance of differentially expressed miRNAs in esophageal cancer. *Int. J. Cancer* **128**, 132–143 (2011).
105. Sasahira, T. *et al.* Downregulation of miR-126 induces angiogenesis and lymphangiogenesis by activation of VEGF-A in oral cancer. *Br. J. Cancer 2012 1074* **107**, 700–706 (2012).
106. Hamada, S. *et al.* MiR-126 Acts as a Tumor Suppressor in Pancreatic Cancer Cells via the Regulation of ADAM9. *Mol. Cancer Res.* **10**, 3–10 (2012).
107. Miko, E. *et al.* miR-126 inhibits proliferation of small cell lung cancer cells by targeting SLC7A5. *FEBS Lett.* **585**, 1191–1196 (2011).
108. Zhang, J. *et al.* The cell growth suppressor, mir-126, targets IRS-1. *Biochem. Biophys. Res. Commun.* **377**, 136–140 (2008).
109. Feng, R. *et al.* miR-126 functions as a tumour suppressor in human gastric cancer. *Cancer Lett.* **298**, 50–63 (2010).
110. Otsubo, T. *et al.* MicroRNA-126 Inhibits SOX2 Expression and Contributes to Gastric Carcinogenesis. *PLoS One* **6**, e16617 (2011).
111. Watahiki, A. *et al.* MicroRNAs Associated with Metastatic Prostate Cancer. *PLoS One* **6**, e24950 (2011).
112. Resnick, K. E. *et al.* The detection of differentially expressed microRNAs from the serum of ovarian cancer patients using a novel real-time PCR platform. *Gynecol. Oncol.* **112**, 55–59 (2009).
113. Ito, Y. & Miyazono, K. RUNX transcription factors as key targets of TGF-β superfamily

signaling. *Curr. Opin. Genet. Dev.* **13**, 43–47 (2003).

114. Li, Q. L. *et al.* Causal Relationship between the Loss of RUNX3 Expression and Gastric Cancer. *Cell* **109**, 113–124 (2002).
115. Soong, R. *et al.* The expression of RUNX3 in colorectal cancer is associated with disease stage and patient outcome. *Br. J. Cancer* **100**, 676 (2009).
116. Attisano, L. & Wrana, J. L. Signal transduction by the TGF-β superfamily. *Science (80-.).* **296**, 1646–1647 (2002).
117. Hahn, S. A. *et al.* DPC4, A Candidate Tumor Suppressor Gene at Human Chromosome 18q21.1. *Science (80-.).* **271**, 350–353 (1996).
118. Miyaki, M. & Kuroki, T. Role of Smad4 (DPC4) inactivation in human cancer. *Biochem. Biophys. Res. Commun.* **306**, 799–804 (2003).
119. Zhao, M., Mishra, L. & Deng, C. X. The role of TGF-β/SMAD4 signaling in cancer. *Int. J. Biol. Sci.* **14**, 111 (2018).
120. Wilentz, R. E. *et al.* Immunohistochemical labeling for dpc4 mirrors genetic status in pancreatic adenocarcinomas : a new marker of DPC4 inactivation. *Am. J. Pathol.* **156**, 37–43 (2000).
121. Alazzouzi, H. *et al.* SMAD4 as a Prognostic Marker in Colorectal Cancer. *Clin. Cancer Res.* **11**, 2606–2611 (2005).
122. Alhopuro, P. *et al.* SMAD4 Levels and Response to 5-Fluorouracil in Colorectal Cancer. *Clin. Cancer Res.* **11**, 6311–6316 (2005).
123. Isaksson-Mettävainio, M., Palmqvist, R., Forssell, J., Stenling, R. & Öberg, Å. SMAD4/DPC4 Expression and Prognosis in Human Colorectal Cancer. *Anticancer Res.* **26**, 507–510 (2006).
124. Mardilovich, K., Pankratz, S. L. & Shaw, L. M. Expression and function of the insulin receptor substrate proteins in cancer. *Cell Commun. Signal.* **7**, 14 (2009).
125. Sun, X. J. *et al.* Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. *Nat. 1991 3526330* **352**, 73–77 (1991).
126. Sun, X. J. *et al.* Role of IRS-2 in insulin and cytokine signalling. *Nat. 1995 3776545* **377**, 173–177 (1995).
127. Reiss, K., Del Valle, L., Lassak, A. & Trojanek, J. Nuclear IRS-1 and Cancer. *J. Cell. Physiol.* **227**, 2992 (2012).
128. Silver Spring (MD): Food and Drug Administration (US); Bethesda (MD): National Institutes of Health (US). BEST (Biomarkers, EndpointS, and other Tools) Resource. *BEST (Biomarkers , EndpointS , other Tools) Resour.* (2016).
129. Goldstein, N. S., Sanford, W., Coffey, M. & Layfield, L. J. Lymph node recovery from colorectal resection specimens removed for adenocarcinoma. Trends over time and a recommendation for a minimum number of lymph nodes to be recovered. *Am. J. Clin. Pathol.* **106**, 209–216 (1996).

130. Battifora, H. The multitumor (sausage) tissue block: novel method for immunohistochemical antibody testing. *Lab. Invest.* (1986).
131. Kononen, J. *et al.* Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat. Med. 1998 47* **4**, 844–847 (1998).
132. Fejzo, M. S. & Slamon, D. J. Frozen Tumor Tissue Microarray Technology for Analysis of Tumor RNA, DNA, and Proteins. *Am. J. Pathol.* **159**, 1645–1650 (2001).
133. Andersson, A. C. *et al.* Analysis of Protein Expression in Cell Microarrays: A Tool for Antibody-based Proteomics. *J. Histochem. Cytochem.* **54**, 1413 (2006).
134. Wen, C. H., Su, Y. C., Wang, S. L., Yang, S. F. & Chai, C. Y. Application of the Microarray Technique to Cell Blocks. *Acta Cytol.* **51**, 42–46 (2007).
135. Kallioniemi, O. P., Wagner, U., Kononen, J. & Sauter, G. Tissue microarray technology for high-throughput molecular profiling of cancer. *Hum. Mol. Genet.* **10**, 657–662 (2001).
136. Ilyas, M. *et al.* Guidelines and considerations for conducting experiments using tissue microarrays. *Histopathology* **62**, 827–839 (2013).
137. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
138. Taylor, C. R. An exaltation of experts: Concerted efforts in the standardization of immunohistochemistry. *Hum. Pathol.* **25**, 2–11 (1994).
139. de Matos, L. L., Trufelli, D. C., de Matos, M. G. L. & Pinhal, M. A. da S. Immunohistochemistry as an Important Tool in Biomarkers Detection and Clinical Practice. *Biomark. Insights* **5**, 9 (2010).
140. Lipman, N. S., Jackson, L. R., Trudel, L. J. & Weis-Garcia, F. Monoclonal Versus Polyclonal Antibodies: Distinguishing Characteristics, Applications, and Information Resources. *ILAR J.* **46**, 258–268 (2005).
141. Jensen, E. Technical Review: In Situ Hybridization. *Anat. Rec.* **297**, 1349–1353 (2014).
142. Gall, J. G., Lou, M., Kline, P. & Giles, N. H. FORMATION AND DETECTION OF RNA-DNA HYBRID MOLECULES IN CYTOLOGICAL PREPARATIONS. *Proc. Natl. Acad. Sci. U. S. A.* **63**, 378 (1969).
143. Kloosterman, W. P., Wienholds, E., de Bruijn, E., Kauppinen, S. & Plasterk, R. H. A. In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat. Methods 2006 31* **3**, 27–29 (2005).
144. Seidal, T., Balaton, A. J. & Battifora, H. Interpretation and quantification of immunostains. *Am. J. Surg. Pathol.* **25**, 1204–1207 (2001).
145. Altman, D. G., Lausen, B., Sauerbrei, W. & Schumacher, M. Dangers of using 'optimal' cutpoints in the evaluation of prognostic factors. *J. Natl. Cancer Inst.* **86**, 829–835 (1994).
146. Caco-2 [Caco2] | ATCC. Available at: https://www.atcc.org/products/htb-37. (Accessed: 2nd February 2022)

147. HT-29 | ATCC. Available at: https://www.atcc.org/products/htb-38. (Accessed: 2nd February 2022)
148. Lea, T. Caco-2 Cell Line. *Impact Food Bioact. Heal. Vitr. Ex Vivo Model.* 103–111 (2015). doi:10.1007/978-3-319-16104-4_10
149. Martínez-Maqueda, D., Miralles, B. & Recio, I. HT29 Cell Line. *Impact Food Bioact. Heal. Vitr. Ex Vivo Model.* 113–124 (2015). doi:10.1007/978-3-319-16104-4_11
150. Dyhl-Polk, A. *et al.* Incidence and risk markers of 5-fluorouracil and capecitabine cardiotoxicity in patients with colorectal cancer. *Acta Oncol. (Madr).* **59**, 475–483 (2020).
151. Soveri, L. M. *et al.* Long-term neuropathy and quality of life in colorectal cancer patients treated with oxaliplatin containing adjuvant chemotherapy. *Acta Oncol. (Madr).* **58**, 398–406 (2019).
152. Dong, Y. *et al.* Prognostic significance of miR-126 in various cancers: A meta-analysis. *Onco. Targets. Ther.* **9**, 2547–2555 (2016).
153. Yang, Z., Wang, R., Zhang, T. & Dong, X. MicroRNA-126 regulates migration and invasion of gastric cancer by targeting CADM1. *Int. J. Clin. Exp. Pathol.* **8**, 8869 (2015).
154. Díaz, R. *et al.* Deregulated expression of miR-106a predicts survival in human colon cancer patients. *Genes, Chromosom. Cancer* **47**, 794–802 (2008).
155. Ebrahimi, F., Gopalan, V., Smith, R. A. & Lam, A. K. Y. MiR-126 in human cancers: Clinical roles and current perspectives. *Experimental and Molecular Pathology* **96**, 98–107 (2014).
156. Hu, M., Xiong, S., Chen, Q., Zhu, S. & Zhou, X. Novel role of microRNA-126 in digestive system cancers: From bench to bedside (review). *Oncology Letters* **17**, 31–41 (2019).
157. Li, X. *et al.* Identification of aberrantly expressed miRNAs in rectal cancer. *Oncol. Rep.* **28**, 77–84 (2012).
158. Hansen, T. F., Sørensen, F. B., Lindebjerg, J. & Jakobsen, A. The predictive value of microRNA-126 in relation to first line treatment with capecitabine and oxaliplatin in patients with metastatic colorectal cancer. *BMC Cancer* **12**, 83 (2012).
159. Ebrahimi, F. *et al.* Deregulation of miR-126 expression in colorectal cancer pathogenesis and its clinical significance. *Exp. Cell Res.* **339**, 333–341 (2015).
160. Bovell, L. C. *et al.* The prognostic value of MicroRNAs varies with patient race/ethnicity and stage of colorectal cancer. *Clin. Cancer Res.* **19**, 3955–3965 (2013).
161. Calin, G. A. & Croce, C. M. MicroRNA signatures in human cancers. *Nat. Rev. Cancer 2006 611* **6**, 857–866 (2006).
162. Zhang, J. X. *et al.* Prognostic and predictive value of a microRNA signature in stage II colon cancer: A microRNA expression analysis. *Lancet Oncol.* **14**, 1295–1306 (2013).
163. Caritg, O. *et al.* Identifying High-Risk Stage II Colon Cancer Patients: A Three-MicroRNA-Based Score as a Prognostic Biomarker. *Clin. Colorectal Cancer* **15**, e175–e182 (2016).
164. Olive, V., Li, Q. & He, L. mir-17-92: A polycistronic oncomir with pleiotropic functions.

Immunol. Rev. **253**, 158–166 (2013).

165. Moody, L., Dvoretskiy, S., An, R., Mantha, S. & Pan, Y. X. The efficacy of miR-20a as a diagnostic and prognostic biomarker for colorectal cancer: A systematic review and meta-analysis. *Cancers* **11**, (2019).
166. Zhou, Y. *et al.* Down-Regulation of miR-126 Is Associated with Colorectal Cancer Cells Proliferation, Migration and Invasion by Targeting IRS-1 via the AKT and ERK1/2 Signaling Pathways. *PLoS One* **8**, (2013).
167. Song, J., Liu, Y., Wang, T., Li, B. & Zhang, S. MiR-17-5p promotes cellular proliferation and invasiveness by targeting RUNX3 in gastric cancer. *Biomed. Pharmacother.* **128**, 110246 (2020).
168. Zhang, Y. *et al.* Cancer-associated fibroblasts-derived exosomal miR-17-5p promotes colorectal cancer aggressive phenotype by initiating a RUNX3/MYC/TGF-β1 positive feedback loop. *Cancer Lett.* **491**, 22–35 (2020).
169. Bai, X., Han, G., Liu, Y., Jiang, H. & He, Q. MiRNA-20a-5p promotes the growth of triple-negative breast cancer cells through targeting RUNX3. *Biomed. Pharmacother.* **103**, 1482–1489 (2018).
170. Chen, Y. *et al.* MicroRNA-20a-5p targets RUNX3 to regulate proliferation and migration of human hepatocellular cancer cells. *Oncol. Rep.* **36**, 3379–3386 (2016).
171. Cheng, D. *et al.* MicroRNA-20a-5p promotes colorectal cancer invasion and metastasis by downregulating Smad4. *Oncotarget* **7**, 45199 (2016).
172. Slaoui, M. & Fiette, L. Histopathology procedures: from tissue sampling to histopathological evaluation. *Methods Mol. Biol.* **691**, 69–82 (2011).
173. Turkki, R., Linder, N., Kovanen, P. E., Pellinen, T. & Lundin, J. Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. *J. Pathol. Inform.* **7**, (2016).
174. Kallen, H., Molin, J., Heyden, A., Lundstrom, C. & Astrom, K. Towards grading gleason score using generically trained deep convolutional neural networks. *Proc. - Int. Symp. Biomed. Imaging* **2016-June**, 1163–1167 (2016).
175. Wang, H. *et al.* Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. *https://doi.org/10.1117/1.JMI.1.3.034003* **1**, 034003 (2014).
176. Bychkov, D. *et al.* Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci. Reports 2018 81* **8**, 1–11 (2018).
177. Skrede, O. J. *et al.* Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet* **395**, 350–360 (2020).
178. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell. 2019 15* **1**, 206–215 (2019).

179. Steiner, D. F. *et al.* Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *Am. J. Surg. Pathol.* **42**, 1636 (2018).

scientific reports

Check for updates

OPEN

High expression of microRNA-126 relates to favorable prognosis for colon cancer patients

Hallgeir Selven[1,2]✉, Lill-Tove Rasmussen Busund[3,4], Sigve Andersen[1,2], Roy M. Bremnes[1,2] & Thomas Karsten Kilvær[1,2]

miR-126 has been identified both as a tumor suppressor and an oncogene in different types of cancer. The aim of this study was to investigate the prognostic impact of miR-126-expression in colon cancer patients. Tumor tissue from 452 patients operated for stage I–III colon cancer was retrospectively collected and tissue microarrays were constructed. miR-126 expression was evaluated by in situ hybridization and analyzed using digital pathology. To isolate the compartment specific contribution of miR-126, tumor and adjacent tumor stroma were considered separately. In univariate analyses, high expression of miR-126 in tumor and stroma was related to increased disease-specific survival ($p < 0.001$ and $p = 0.005$, respectively). In multivariate analyses, high miR-126 expression in tumor remained a significant independent predictor of improved disease-specific survival (HR = 0.42, CI 0.23–0.75, $p = 0.004$). Within different TNM-stages there was a tendency towards the same results, but with statistically significant results in stage II only ($p = 0.007$). High expression of miR-126 is an independent positive prognostic factor in stage I–III colon cancer. This finding may be used to identify patients in need of adjuvant chemotherapy.

In the US, colon cancer is the 4th most common cancer among women and men, separately and combined. Estimates show that more than 100,000 patients will be diagnosed with colon cancer in the US in 2020. Moreover, the estimated life time risk of contracting colon cancer in the US is 4%[1]. Globally, colon cancer is the 4th most common cancer type, and the 5th leading cause of cancer-related death[2].

Prognostication of colon cancer patients relies on the TNM-system and histopathological criteria according to the American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC)[3]. Clinical management, including the need for adjuvant treatment, is assessed according to the TNM staging system[4]. However, contrary to other forms of cancer, the TNM staging system for colon cancer is rather imprecise, and the risk of recurrence varies significantly for patients within the same pathological stage[5]. This lack of precision has encouraged researchers to search for more precise prognostic and predictive biomarkers.

MicroRNAs (miRNAs or miRs), first identified in 1993, are single-stranded, non-coding RNAs, approximately 22 nucleotides long, regulating gene expression at the post-transcriptional level. Many miRNAs exist in the human genome, and each miRNA can potentially regulate hundreds of mRNAs, making them important mediators of cellular processes (differentiation, proliferation, apoptosis, stress response etc.). In cancer, miRNAs regulate molecular pathways by targeting oncogenes and tumor suppressors. Moreover, they play significant roles in cancer-stem-cell biology, angiogenesis, epithelial-mesenchymal-transition, metastasis and drug resistance, among others[6,7].

miR-126 mediates developmental angiogenesis in vivo and is the most highly enriched miR in endothelial cells. It represses intracellular inhibitors of angiogenic signaling (VEGF-pathway), thus enhancing the pro-angiogenic actions of VEGF and FGF, leading to blood-vessel formation. In zebrafish, knockdown of miR-126, results in loss of vascular integrity and hemorrhage during embryonic development[8,9].

miR-126 is aberrantly expressed in most cancer types, including cancers of the gastrointestinal tract, genital tracts, breast cancer, thyroid cancer, lung cancer and acute myeloid leukemia[10]. Previous studies have shown a loss of miR-126 expression in colon cancer cell lines compared to normal colon epithelium. Reconstitution of miR-126 results in significant growth reduction[11]. Hence, miR-126 is considered a suppressor of colon cancer development.

[1]Department of Oncology, University Hospital of North Norway, 9038 Tromsö, Norway. [2]Department of Clinical Medicine, UiT The Arctic University of Norway, Tromsö, Norway. [3]Department of Medical Biology, UiT The Arctic University of Norway, Tromsö, Norway. [4]Department of Clinical Pathology, University Hospital of North Norway, Tromsö, Norway. ✉email: hallgeir.selven@uit.no

Further, corroborating the results from cell line studies, Hansen et al. demonstrated that patients with metastatic colorectal cancer (mCRC) responding to 1st line chemotherapy presented with a significantly higher median miR-126 expression in tumor associated vasculature *vs* non-responders. This translated to a significantly enhanced median progression-free survival (PFS) (11.5 months *vs* 6.0 months) for patients with high *vs* low miR-126 expressing tumors[12]. In patients operated for stage II colon cancer, low expression of miR-126 correlated to established, negative prognostic factors (T4 and high malignancy grade among others). Patients with high miR-126 expression had a significantly improved overall survival compared to patients with low miR-126 expression[13].

Other studies have shown that miR-126 is detectable in plasma, and that an increase in plasma miR-126 may be predictive of tumor response in colon cancer patients receiving palliative chemotherapy[14]. These factors may provide an attractive, non-invasive method to evaluate treatment response.

Several studies have shown high miR-126 expression to be a positive prognostic factor in colon cancer patients. But most of these studies used qPCR, excluding the possibilities to detect specific expression in the various tumor compartments. We sought to explore the prognostic impact of miR-126 expression in both the tumor epithelial and surrounding stromal cells utilizing in situ hybridization. The study was conducted on primary tumors from 452 stage I–III colon cancer patients. We hypothesized that miR-126 is a clinically relevant biomarker for this group of patients.

Results

Patient characteristics. A total number of 452 patients were included in this study. The main patient characteristics are summarized in Table 1. Median age at surgery was 74 years (range 30–94), median follow-up of survivors was 173 months (range 119–239). There were 243 females (53.8%) and 209 males (46.2%). Median tumor size was 50 mm (range 10–180 mm). Nine (2%), 76 (16.8%), 320 (70.8%) and 47 (10.4%) patients were categorized as pT1-4, respectively. Lymph node-positive disease was present in 161 patients (35.6%). According to the pTNM-stage, 72 (15.9%), 219 (48.5%) and 161 (35.6%) patients were diagnosed with stage I-III disease, respectively. A total of 87 patients (19.2%) received adjuvant chemotherapy. Nordic-FLv (5-FU/Leukovorin bolus schedule) was the chosen regimen until 2004 when it was shown that adding Oxaliplatin gave superior results for patients < 70 years. From 2004, patients < 70 years were consequently offered treatment with Nordic FLOX (Oxaliplatin/5-FU/Leukovorin bolus schedule)[15–17]. Thus, 69 patients (79.3%) were given Nordic-FLv and five patients (5.7%) Nordic FLOX. Eleven patients (12.6%) were initially administered Nordic FLOX, but were later converted to Nordic-FLv because of unacceptable toxicity. Two patients (2.3%) were given adjuvant radio-/chemotherapy. At the end of follow-up, 119 patients had verified recurrent disease (26.3%) and 313 (69.2%) were dead, either due to colon cancer (108, 34.5%) or other causes (205, 65.5%).

Expression of miR-126 and its correlations. miR-126 was expressed in tumor epithelial cells as well as stromal cells including spindle shaped cells (likely fibroblasts, endothelial cells and vascular smooth muscle cells) and immune cells (Fig. 1). miR-126 expression in tumor and stroma was highly correlated ($r = 0.60$). Table 2B shows associations between miR-126-expression in both tumor and stroma and clinicopathological variables. Low miR-126-expression in stroma was significantly associated with increasing pStage ($p = 0.027$) as well as histological grade ($p = 0.004$).

Univariate analysis. Univariate survival analyses according to clinicopathological variables and miR-126 are summarized in Table 2A and visualized in Fig. 2. Age ($p = 0.013$), weight loss ($p = 0.012$), pathological stage ($p < 0.001$), histological grade ($p < 0.001$), vascular infiltration ($p < 0.001$), resection margins ($p = 0.005$), miR-126 expression in tumor ($p < 0.001$) and miR-126 expression in stroma ($p = 0.005$) were all significant indicators of DSS in the total population.

Analyses in subgroups according to pTNM stage, revealed that high expression of miR-126 in both tumor and stroma showed a tendency towards being a positive predictor of DSS within each TNM stage (S1 Fig). This effect was significant in stage II patients ($p = 0.007$ and 0.009 for tumor and stroma, respectively).

Multivariate analysis. Multivariate analyses are summarized in Table 3. Age (HR 1.03, CI95% 1.01–1.05, $p = 0.002$) and miR-126-expression in tumor (HR 0.45, CI 0.27–0.76, $p = 0.002$) were independently associated with DSS. Histological grade also showed statistical significance overall ($p = 0.032$), but with non-significant results within the subgroups (moderate (HR 0.76, CI 0.35–1.68, $p = 0.5$), poor (HR 0.63, CI 0.26–1.57, $p = 0.322$), or undifferentiated (HR 4.02, CI 0.99–16.36, $p = 0.052$)). Pathological stage was statistically significant overall ($p < 0.001$), but in subgroups only stage III was statistically significant (stage II: HR 2.12, CI 0.82–5.53, $p = 0.123$, stage III: HR 7.35, CI 2.89–18.72, $p < 0.001$).

Discussion

In the presented material, DSS was significantly associated with previously well-known prognostic factors including age, weight loss, pathological stage, histological grade, vascular infiltration and resection margins[18]. In addition, low expression of miR-126 in both tumor and stroma was significantly associated with shorter DSS. Overall, patients with high expression of miR-126 in cancer cells had a 5-year DSS of 88% compared with 75% for low expression. Likewise, patients with high expression of miR-126 in tumor stroma had a 5-year DSS of 88% compared with 76% for low expression. To our knowledge, this is the first paper to demonstrate compartment-specific prognostic impact of miR-126 in cancer and stromal cells using in situ hybridization in a large, homogenous cohort of primary tumors from colon cancer patients treated with curative intent. Moreover, we introduce an open source pipeline for miR expression analyses that may be automated and adapted to other use cases.

Variable	n (%)	5-y OS
Age at diagnosis		
Median	74.0 years (30–94 years)	N/A
Tumor size		
Median	50 mm (10–180 mm)	N/A
Missing	2	N/A
Gender		
Male	209 (46.2%)	59.8%
Female	243 (53.8%)	62.1%
Relapse		
No	333 (73.7%)	74.5%
Yes	119 (26.3%)	23.5%
pT status		
T1	9 (2.0%)	88.9%
T2	76 (16.8%)	65.8%
T3	320 (70.8%)	61.6%
T4a	26 (5.8%)	46.2%
T4b	21 (4.6%)	42.9%
pN status		
N0	291 (64.4%)	68.4%
N1a	56 (12.4%)	50.0%
N1b	55 (12.2%)	52.7%
N1c	1 (0.2%)	0%
N2a	27 (6.0%)	40.7%
N2b	22 (4.9%)	40.9%
pTNM stage (groups)		
I	72 (15.9%)	70.8%
II	219 (48.5%)	67.6%
III	161 (35.6%)	47.8%
Differentiation		
Well	36 (8.0%)	80.6%
Moderate	329 (72.8%)	59.6%
Poor	75 (16.6%)	60.0%
Undifferentiated	4 (0.9%)	25.0%
Missing	8 (1.8%)	N/A
Site		
Right	227 (50.2%)	63.0%
Transverse	65 (14.4%)	49.2%
Left	21 (4.6%)	66.7%
Sigmoid	136 (30.1%)	61.8%
Missing	3 (0.7%)	N/A
Weight loss		
< 10%	248 (54.9%)	68.5%
≥ 10%	94 (20.8%)	46.8%
Missing	110 (24.3%)	N/A
Performance status (ECOG)		
0	237 (52.4%)	70.0%
1	149 (33.0%)	58.4%
2	54 (11.9%)	33.3%
3	8 (1.8%)	25.0%
Missing	4 (0.9%)	N/A
Adjuvant chemotherapy		
No	365 (80.8%)	60.3%
Yes	87 (19.2%)	67.8%
Postoperative complications		
No	358 (79.2%)	64.5%
Yes	94 (20.8%)	47.9%

Continued

Variable	n (%)	5-y OS
miR-126 expression tumor		
High	140 (31.0%)	67.9%
Low	300 (66.4%)	58.0%
Missing	12 (2.7%)	N/A
miR-126 expression stroma		
High	109 (24.1%)	63.3%
Low	341 (75.4%)	60.4%
Missing	2 (0.4%)	N/A

Table 1. Frequency table summarizing the median and range and total number of important continuous and cathegorical clinicopathological variables, respectively. For categorical variables five-year overall survival is given in percent.

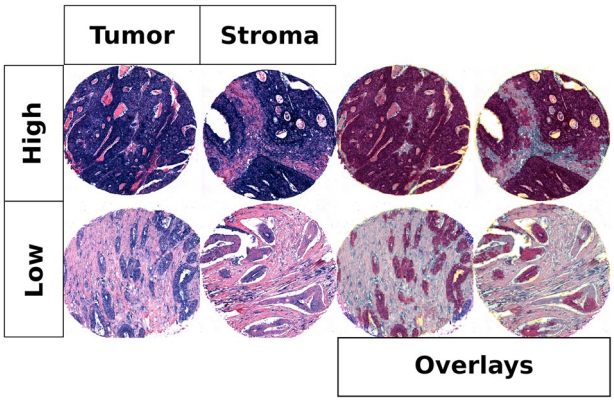

Figure 1. High and low scores for miR-126 in tumor and stroma with and without overlays.

A meta-analysis, published in 2016, assessed the effectiveness of miR-126 as a prognostic biomarker for various cancers[19]. In small cell lung cancer, miR-126 overexpression induced delayed G1 phase of the cancer cell cycle. Likewise, in breast cancer cells, cell cycle progression was inhibited from the G1/G0 phase to S phase by miR-126. In pancreatic cancer cells, low expression of miR-126 led to increased epithelial-mesenchymal transition, a process involved in metastasis. In some gastric cancer cell lines, however, overexpression of miR-126 resulted in tumor growth by regulating its downstream target genes. According to this meta-analysis, higher miR-126 expression predicted better OS in digestive system and respiratory system cancers.

miR-126 normally functions as a growth suppressor in colon cells. In colon cancer cell lines, miR-126 is frequently downregulated compared to normal colon epithelium. Guo et al. showed that miR-126 regulates phosphatidylinositol 3-kinase (PI3K)-signaling by targeting the PI3K regulatory subunit beta (p85β) through translational repression[11]. Downregulation of miR-126 leads to upregulation of PI3K-signalling, resulting in tumor growth. Restoration of miR-126 yielded a marked reduction in p85β. Chemokine receptor 4 (CXCR4) is highly expressed in various types of cancer, and is considered important for mobilization, migration, proliferation and survival of different cell types. CXCR4 is a target for miR-126-mediated repression, and this repression may inhibit migration and invasion of colon cancer cells. The Ras homolog gene family, member A (RhoA), is associated with invasion and poor prognosis in colorectal cancer. Rho function by downstream signaling via PI3K and ROCK (Rho-associated coiled-coil-containing protein kinase) among others. miR-126 has been shown to act as a tumor suppressor via RhoA/ROCK inhibition in cancer cells. Yuan et al. demonstrated that colon cancer cell invasion and migration was inhibited by miR-126, in vivo and in vitro, by down-regulating CXCR4 and inactivating the RhoA signaling pathway[20].

In 2014, Liu et al. published an article on the prognostic impact of miR-126 expression in colorectal cancer with similar results to ours. By using qPCR they found approximately the same DSS in patients with high expression of miR-126, but even worse prognosis for patients with low expression of miR-126[21]. A possible explanation for this worse outcome could be the inclusion of patients with metastatic disease in their study, while we included only patients treated with curative intent.

Hansen et al. assessed the prognostic value of miR-126 and microvessel density (MVD) in patients with stage II colon cancer[13]. Their study included 560 patients, and the primary endpoints were recurrence-free cancer specific survival (RF-CSS) and overall survival (OS). qPCR was performed to analyze miR-126 expression. Like in our study, they found that low expression of miR-126 correlated to histological grade and a worse OS ($p = 0.03$). In their multivariate analysis, a borderline impact on OS was found ($p = 0.051$).

	A					B					
						miR126 in tumor			miR126 in stroma		
	N (%)	5 Year	Median	HR (95%CI)	P	Low	High		Low	High	P
Age					0.013			0.358			0.130
≤ 65	110 (24)	88	NA	1.00		78	30		89	20	
> 65	342 (76)	76	NA	1.83 (1.21–2.77)		222	110		252	89	
Gender					0.388			0.751			1.000
Female	243 (54)	78	NA	1.00		165	74		184	59	
Male	209 (46)	81	NA	0.85 (0.58–1.23)		135	66		157	50	
Weight_loss					0.012			0.092			0.621
< 10%	248 (55)	82	NA	1.00		157	83		192	55	
≥ 10%	94 (21)	68	NA	1.74 (1.06–2.87)		69	22		70	24	
Missing	110 (24)										
ECOG_status					0.424			0.997			0.993
0	237 (52)	82	NA	1.00		157	75		178	57	
1	149 (33)	76	NA	1.31 (0.86–1.99)		98	45		113	36	
2	54 (12)	74	NA	1.18 (0.62–2.28)		36	17		40	14	
3	8 (2)	50	47	2.31 (0.31–16.98)		6	2		6	2	
Missing	4 (1)										
Site					0.951			0.824			0.700
Sigmoid	227 (50)	79	NA	1.00		152	69		171	56	
Transversum	65 (14)	78	NA	0.94 (0.52–1.68)		45	19		47	18	
Left	21 (5)	80	NA	1.15 (0.47–2.78)		14	5		18	3	
Right	136 (30)	79	NA	1.09 (0.71–1.67)		87	46		102	32	
Missing	3 (1)										
pStage					< 0.001			0.388			0.027
1	72 (16)	94	NA	1.00		46	25		53	19	
2	219 (48)	89	NA	2.27 (1.35–3.81)		142	72		156	63	
3	161 (36)	59	NA	8.04 (4.58–14.11)		112	43		132	27	
Differentiation					< 0.001			0.311			0.006
Well	36 (8)	89	NA	1.00		19	16		27	9	
Moderate	329 (73)	78	NA	1.47 (0.76–2.82)		224	98		259	68	
Poor	75 (17)	79	NA	1.45 (0.67–3.15)		49	22		48	27	
Undifferentiated	4 (1)	25	10	11.34 (0.45–283.15)		3	1		1	3	
Missing	8 (2)										
Vasc+					< 0.001			0.305			0.575
No	199 (44)	83	NA	1.00		121	69		151	48	
Yes	19 (4)	45	30	5.22 (1.66–16.37)		14	4		15	3	
Missing	234 (52)										
Resection margins					0.005			0.519			0.630
0 mm	31 (7)	54	68	1.00		23	6		26	5	
< 1 mm	42 (9)	70	NA	0.56 (0.2–1.54)		31	11		31	11	
1–2 mm	35 (8)	88	NA	0.3 (0.11–0.85)		21	14		25	10	
2–10 mm	121 (27)	85	NA	0.33 (0.14–0.77)		81	35		94	26	
10–50 mm	155 (34)	80	NA	0.4 (0.17–0.95)		100	50		112	43	
> 50 mm	46 (10)	80	NA	0.3 (0.11–0.81)		29	17		36	9	
Missing	22 (5)										
miR126 in tumor					< 0.001						
Low	300 (66)	75	NA	1.00							
High	140 (31)	88	NA	0.43 (0.29–0.64)							
Missing	12 (3)										
miR126 in stroma					0.005						
Low	341 (75)	76	NA	1.00							
High	109 (24)	88	NA	0.46 (0.29–0.71)							
Missing	2 (0)										

Table 2. (A) Clinicopathological variables and miR-126 in tumor and stroma as predictors of disease-specific survival for colon cancer patients (univariate analyses, log-rank test, n = 452), (B) dichotomized miR-126 in tumor and stroma and their distribution over and correlation with clinicopathological variables (chi-square and Fisher's exact tests). ECOG, eastern cooperative oncology group.

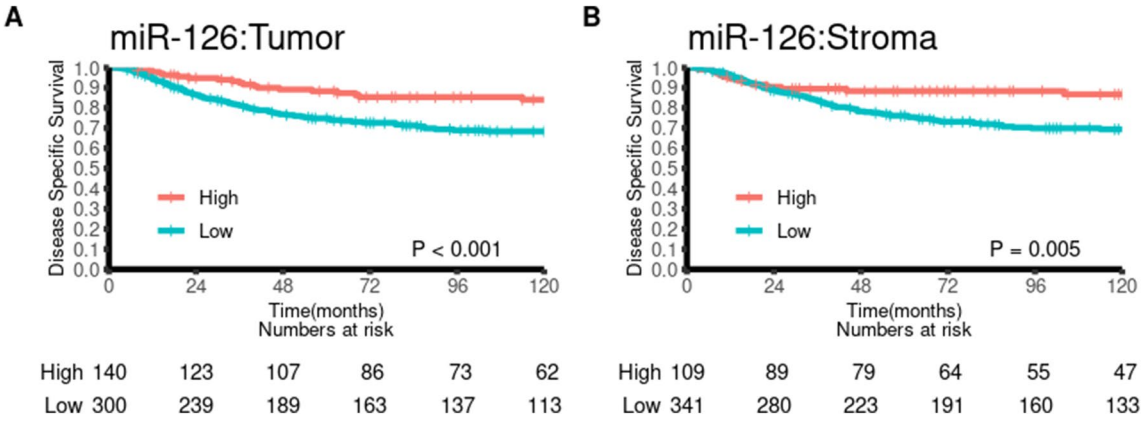

Figure 2. Disease-specific survival curves for tumor (**A**) and stromal (**B**) expression of miR-126, using the optimal cut-offs for each marker.

	(A)		95.0% CI		(B)		95.0% CI	
	p	HR	Lower	Upper	p	HR	Lower	Upper
Age at diagnosis	**0.002**	1.03	1.01	1.05	**0.003**	1.03	1.01	1.05
Differentiation	**0.032**				**0.027**			
Well differentiated		1.00				1.00		
Moderately differentiated	0.500	0.76	0.35	1.68	0.706	0.86	0.39	1.89
Poorly differentiated	0.322	0.63	0.26	1.57	0.653	0.81	0.33	2.01
Undifferentiated	0.052	4.02	0.99	16.36	**0.020**	5.27	1.30	21.46
pTNM	**< 0.001**				**< 0.001**			
pTNM-stage I		1.00				1.00		
pTNM-stage II	0.123	2.12	0.82	5.53	0.122	2.12	0.82	5.51
pTNM-stage III	**< 0.001**	7.35	2.89	18.72	**< 0.001**	6.76	2.66	17.18
miR-126 tumor								
Low		1.00						
High	**0.002**	0.46	0.27	0.76				
miR-126 stroma								
Low						1.00		
High					**0.049**	0.55	0.30	1.00

Table 3. Multivariate models for miR-126 in (A) tumor and (B) stroma (cox proportional hazards test, n = 452). Statistically significant results indicated in bold.

In our stage I–III material, patients with high expression of miR-126 in both tumor and stroma, show a better 5-year DSS than patients with low expression. Several studies have assessed the association between miR-126-expression and metastatic colorectal cancer (stage IV). Hansen et al. published a study in 2012 looking at the predictive value of miR-126 in patients treated with first line palliative chemotherapy[12]. They found that responders to chemotherapy had a significantly higher median miR-126 expression than non-responders ($p < 0.001$). Median PFS was 11.5 *vs* 6.0 months ($p < 0.001$) and median OS 26.2 *vs* 16.8 months ($p = 0.002$) in high *vs* low expression tumors. The same group showed similar results with increased PFS for patients with high miR-126 expression ($p = 0.005$) when assessing the clinical outcome of patients with metastatic colorectal cancer treated with chemotherapy combined with Bevacizumab (anti-VEGF-A)[22]. It is, however, questionable to compare data from metastatic patients with those treated with curative intent. However, a study by Ebrahimi et al. showed similar levels of miR-126 expression in primary tumors and metastatic lesions[23]. Caution must be taken when correlating these data, as the results from the metastatic patients also included patients with primary rectal cancer.

One of the strengths of this paper is the unselected study population from Northern Norway with a relatively large number of patients. The study has an extended follow-up and includes relevant clinicopathological variables. Although our clinical database comprises patients operated at least 13 years ago, it still is a representative cohort when comparing to the most contemporary data from the Norwegian cancer registry and colorectal cancer statistics from the United States[1,24]. Median age at diagnosis is in the early 70s in Norway, and approximately 1 in 3 patients presents with regional disease (stage III). In the United States, median age at diagnosis has dropped to 69 years, possibly as a result of colon cancer screening programs. Weaknesses of our study are the retrospective design and the former inadequate histological reports when compared to today's standards. Vascular infiltration was not routinely assessed in this time-period and neither was possibly relevant mutations (KRAS, NRAS,

BRAF) nor MSI-status. However, in Norway, MSI assessment is only conducted in patients below 60 years of age at the time of diagnosis, and the previously mentioned mutations only in the case of metastatic disease. In the patient cohort treated with curative intent, this information is missing for a majority of the patients and was consequently omitted from statistical analyses.

Previous studies assessing the prognostic impact of miR-126 expression in colon cancer were mainly performed using qPCR. In the previously mentioned meta-analysis from 2016, ISH was applied in three studies, qRT-PCR in 14 studies[19]. When using RNA extracts from whole tumors you get a mixture of neoplastic tumor cells and tumor-related stromal cells, and this method does not give information about miR-126 expression in the various tumor compartments. By using in situ hybridization in our study, we are able to precisely identify the miR-126 expression in different compartments and cell types. Assessing miR-expression by in situ hybridization has previously been evaluated in a semi-quantitative manner. By using digital pathology for this, we omit inter-observer variability, and results are likely to be more accurate and reproducible. This is particularly useful when analyzing the intensity of the positive, blue staining of miR-126 in this material.

We found statistically significant correlations between miR-126 expression in stroma and pathological stage and histological grade ($p = 0.027$ and $p = 0.004$ respectively). However, the true clinical significance of these findings must be considered uncertain due to the low number of patients with undifferentiated tumors ($n = 4$) that may have contributed to the statistical significance. Colon cancer cell line experiments were inconclusive on this issue. Li et al. found a correlation between miR-126 expression and histological grade, but not pathological stage[25]. Ebrahimi et al., on the other hand, found a correlation between miR-126 expression and pathological stage, but not with histological grade[23].

Adjuvant chemotherapy is administered to eradicate potential microscopic disease post surgery. It has been shown to increase 5-year DSS and OS for both stage II and III colon cancer patients[26]. The effect is more pronounced in stage III, and as a consequence, Norwegian patients with stage III colon cancer are routinely offered adjuvant chemotherapy unless they have contraindications. For stage II patients the effect is not as clear-cut[18]. Here, adjuvant chemotherapy is offered on an individual basis to high-risk patients (perforated tumors, pT4 or low number of examined lymph nodes). miR-126 expression may aid the oncologist in treatment-decision-making for stage II colon cancer patients and help select patients likely to benefit from potentially harmful adjuvant chemotherapy. However, before clinical implementation can commence, a standardized assessment of miR-126 must be established and prospective clinical trials conducted.

Conclusion

We have shown that a high expression of miR-126 is an independent positive prognostic factor for stage I–III colon cancer patients. An improved 5-year DSS was seen for patients with a high expression of miR-126 in both cancer cells and stromal cells. Our results largely confirms the results from previous studies. Further, we complement the knowledge of miR-126 expression in colon cancer by investigating its expression in different tumor compartments and by introducing digital pathology. A potential clinical implication of our findings, may be to use miR-126 expression to select the stage II colon cancer patients most likely to benefit from adjuvant chemotherapy.

Material and methods

Study population. Patients who underwent radical surgery for colon cancer, in various hospitals in Northern-Norway in the time-period 1998–2007, were eligible for inclusion in this study (Fig. 3). Exclusion criteria were metastatic disease/non-radical surgery, prior malignancy within the last 5 years before colon cancer diagnosis or other synchronous malignancies. A list of patients diagnosed with colon cancer in the time-period was procured from the Department of Clinical Pathology at the University Hospital of North Norway. Of a total of 861 identified patients, 409 patients were excluded due to: metastatic disease at the time of diagnosis (117), tissue-blocks missing or inadequate for TMA construction (79), wrongly coded as colon cancer (57, mainly rectal cancer), genetic testing on patients operated outside of Northern Norway (55), prior malignancies within the last 5 years before colon cancer diagnosis (46), mortality within the first 90 days after surgery (22), surgery for recurrent colon cancer or treated in a strictly palliative setting (18), lost to follow-up (7, three of them tourists), appearing twice on the list (3), surgery at a regional hospital not participating in the study (2), operated before the actual time frame (2) or a missing medical journal (1). Hence, 452 patients were included in this study. Follow-up was completed December 1, 2017.

Tissue microarray construction. All colon cancer cases were histologically reviewed by two pathologists, and the most representative areas of tumor without necrosis were selected. A 0.6 mm-diameter stylet was used to sample a total of 4 cores securing both tumor tissue and tumor stroma from each included patient. The TMAs were assembled using a tissue-arraying instrument (Beecher Instruments, Silver Springs, MD, USA). The detailed methodology has previously been reported[27].

In situ hybridization (ISH). The method used for in situ hybridization is based on a protocol developed by Jorgensen et al., and adjusted for automatic ISH on the Ventana Discovery-Ultra platform (Roche, Tucson, USA)[28]. Double-DIG labeled miRCURY LNA detection probes (Exiqon AS, Denmark) were used to visualize miR-126-3p. A scrambled probe and U6 were used as negative and positive controls, respectively. Further, a multi organ TMA section was used as an additional control. Optimizations regarding temperatures, times, and concentrations were done for each probe and reagent.

Slides were baked at 60 °C overnight, and then transferred to the Discovery-Ultra for ISH. Sections were then deparaffinized by heating the slide to 68 °C and incubating for three cycles of 12 min. Antigen retrieval was

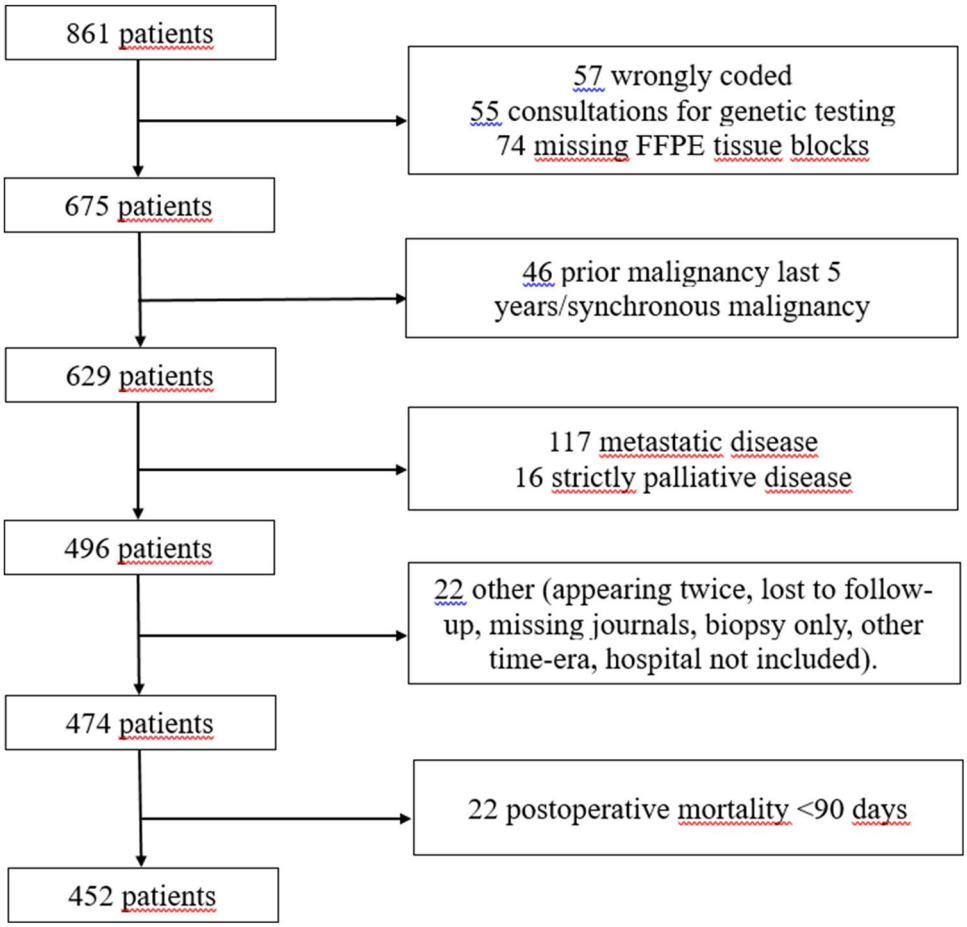

Figure 3. Flowchart of the inclusion of patients in the study.

performed by heating the slides to 95 °C and subsequent treatment with Discovery Cell Conditioning Solution (CC1) for 40 min. The LNA detection probe for miR-126-3p was diluted to a concentration of 2.0 nM, and added manually. The hybridization reaction was carried out at 51 °C for 60 min. Two stringency washes with 2.0X SSC at 51 °C with 8 min incubation before each wash were performed. The slides were incubated with blocking solution for 16 min to block against unspecific bindings. Alkaline phosphatase-conjugated anti DIG (Anti-DIG-AP) was incubated for 20 min for immunologic detection. Substrate enzymatic reactions were carried out with NBT/BCIP for 60 min to give a blue precipitate to detect the miR. The slides were counterstained with Nuclear Fast Red to visualize the nuclei. Slides were dehydrated through an increasing gradient of ethanol solutions and xylene, and mounted with Histokitt mounting medium.

In situ hybridization scoring/QuPath. TMA slides were digitized using a Pannoramic 250 Flash III (3DHistech, Budapest, Hungary) slide scanner, and processed in QuPath v.0.1.3 (Queen's University, Belfast, Northern Ireland) in an Ubuntu 20.04 environment. TMA slides were de-arrayed and preprocessed according to Bankhead et al.[29]. Tissue within each TMA core was identified using simple tissue detection and tiled into 10×10 µm tiles. Image features were calculated for each tile and used to train a Random Forest model. Each tile was classified as either tumor, stroma, necrosis or other. After classification, tiles were converted into continuous areas and the mean intensity of miR-126 within tumor and stroma were calculated. The scripts used to process the TMAs are included in the supplementary file.

All possible dichotomized cut-offs were evaluated (S2 Fig). For any subsequent analyses the optimal cut-off was chosen.

Statistical methods. Statistical tests were performed using the statistical packages SPSS version 26.0 or R version 3.6.3. χ^2 test or Fisher's exact tests were used to examine the association between molecular marker expression and clinicopathological parameters. DSS (disease-specific survival) was defined as the interval from surgery to the time of colon cancer death. For the univariate analyses, the Kaplan–Meier method was used to visualize associations between molecular marker expression and survival. The log-rank test was used to assess the statistical significance of the differences between the survival curves. Multivariate analyses were performed

using a backward conditional Cox regression analysis with a probability for stepwise entry and removal at 0.05 and 0.10, respectively. A p-value < 0.05 was considered statistically significant.

Ethics declaration. This study was approved by the Regional Committee for Medical and Health Research Ethics North (REK Nord, protocol ID: 2011/2151) and the need for patient consent waived. The reporting of clinicopathological variables, survival data and biomarker expression was conducted in accordance with the REMARK guidelines[30].

Data availability

Data will be shared upon reasonable request to the corresponding author.

Received: 11 January 2021; Accepted: 5 April 2021
Published online: 05 May 2021

References

1. Siegel, R. L. *et al.* Colorectal cancer statistics, 2020. *CA. Cancer J. Clin.* **70**, 145–164 (2020).
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **68**, 394–424 (2018).
3. Brierley, J. D., Gospodarowicz, M. K. & Wittekind, C. *TNM Classification of Malignant Tumours, 8th ed* (2016).
4. Labianca, R. *et al.* Early colon cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **24**, vi64–vi72 (2013).
5. Nagtegaal, I. D., Quirke, P. & Schmoll, H. J. Has the new TNM classification for colorectal cancer improved care?. *Nat. Rev. Clin. Oncol.* **9**, 119–123 (2012).
6. Zhang, B., Pan, X., Cobb, G. P. & Anderson, T. A. microRNAs as oncogenes and tumor suppressors. *Dev. Biol.* **302**, 1–12 (2007).
7. Kong, Y. W., Ferland-McCollough, D., Jackson, T. J. & Bushell, M. microRNAs in cancer management. *Lancet Oncol.* **13**, e249–e258 (2012).
8. Fish, J. E. *et al.* miR-126 regulates angiogenic signaling and vascular integrity. *Dev. Cell* **15**, 272–284 (2008).
9. Wang, S. *et al.* The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. *Dev. Cell* **15**, 261–271 (2008).
10. Ebrahimi, F., Gopalan, V., Smith, R. A. & Lam, A. K. Y. MiR-126 in human cancers: Clinical roles and current perspectives. *Exp. Mol. Pathol.* **96**, 98–107 (2014).
11. 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 Chromosom. Cancer* **47**, 939–946 (2008).
12. Hansen, T. F., Sørensen, F. B., Lindebjerg, J. & Jakobsen, A. The predictive value of microRNA-126 in relation to first line treatment with capecitabine and oxaliplatin in patients with metastatic colorectal cancer. *BMC Cancer* **12**, 83 (2012).
13. Hansen, T. F. *et al.* The prognostic value of microRNA-126 and microvessel density in patients with stage II colon cancer: Results from a population cohort. *J. Transl. Med.* **12**, 254 (2014).
14. Hansen, T. F., Carlsen, A. L., Heegaard, N. H. H., Sørensen, F. B. & Jakobsen, A. Changes in circulating microRNA-126 during treatment with chemotherapy and bevacizumab predicts treatment response in patients with metastatic colorectal cancer. *Br. J. Cancer* **112**, 624–629 (2015).
15. André, T. *et al.* Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N. Engl. J. Med.* **350**, 2343–2351 (2004).
16. André, T. *et al.* Adjuvant fluorouracil, leucovorin, and oxaliplatin in stage II to III colon cancer: updated 10-year survival and outcomes according to BRAF mutation and mismatch repair status of the MOSAIC study. *J. Clin. Oncol.* **33**, 4176–4187 (2015).
17. Yothers, G. *et al.* Oxaliplatin as adjuvant therapy for colon cancer: Updated results of NSABP C-07 trial, including survival and subset analyses. *J. Clin. Oncol.* **29**, 3768–3774 (2011).
18. Böckelman, C., Engelmann, B. E., Kaprio, T., Hansen, T. F. & Glimelius, B. Risk of recurrence in patients with colon cancer stage II and III: A systematic review and meta-analysis of recent literature. *Acta Oncol.* **54**, 5–16 (2015).
19. Dong, Y. *et al.* Prognostic significance of miR-126 in various cancers: A meta-analysis. *Onco. Targets. Ther.* **9**, 2547–2555 (2016).
20. Yuan, W. *et al.* MicroRNA-126 inhibits colon cancer cell proliferation and invasion by targeting the chemokine (C-X-C motif) receptor 4 and Ras homolog gene family, member A, signaling pathway. *Oncotarget* **7**, 60230–60244 (2016).
21. Liu, Y. *et al.* Low expression of microRNA-126 is associated with poor prognosis in colorectal cancer. *Genes Chromosom. Cancer* **53**, 358–365 (2014).
22. Hansen, T. F. *et al.* MicroRNA-126 and epidermal growth factor-like domain 7-an angiogenic couple of importance in metastatic colorectal cancer. Results from the Nordic ACT trial. *Br. J. Cancer* **109**, 1243–1251 (2013).
23. Ebrahimi, F. *et al.* Deregulation of miR-126 expression in colorectal cancer pathogenesis and its clinical significance. *Exp. Cell Res.* **339**, 333–341 (2015).
24. Cancer registry of Norway. *Cancer Incidence, Mortality, Survival and Prevalence in Norway: Clustering of Cancer* (2019).
25. Li, X. M., Wang, A. M., Zhang, J. & Yi, H. Down-regulation of miR-126 expression in colorectal cancer and its clinical significance. *Med. Oncol.* **28**, 1054–1057 (2011).
26. Gill, S. *et al.* Pooled analysis of fluorouracil-based adjuvant therapy for stage II and III colon cancer: Who benefits and by how much?. *J. Clin. Oncol.* **22**, 1797–1806 (2004).
27. Bremnes, R. M. *et al.* High-throughput tissue microarray analysis used to evaluate biology and prognostic significance of the E-cadherin pathway in non-small-cell lung cancer. *J. Clin. Oncol.* **20**, 2417–2428 (2002).
28. Jørgensen, S., Baker, A., Møller, S. & Nielsen, B. S. Robust one-day in situ hybridization protocol for detection of microRNAs in paraffin samples using LNA probes. *Methods* **52**, 375–381 (2010).
29. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
30. Sauerbrei, W., Taube, S. E., McShane, L. M., Cavenagh, M. M. & Altman, D. G. Reporting recommendations for tumor marker prognostic studies (REMARK): An abridged explanation and elaboration. *J. Natl Cancer Inst.* **110**, 803–811 (2018).

Acknowledgements

The authors want to thank Mona Irene Pedersen for doing the labwork for the ISH-procedure and for proofreading on methodology. We would also like to thank Magnus L Persson for preparing the TMA-slides, and Olav Magnus Liavåg for helping to collect clinical data for our database. The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway.

Author contributions

All authors contributed to study design. H.S. collected the clinical data. H.S. and T.K.K. analyzed and interpreted the data. H.S. wrote the manuscript. Figures and tables were prepared by H.S. and T.K.K. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-87985-3.

Correspondence and requests for materials should be addressed to H.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

scientific reports

Check for updates

OPEN

High expression of miR-17-5p and miR-20a-5p predicts favorable disease-specific survival in stage I-III colon cancer

Hallgeir Selven[1,2]✉, Sigve Andersen[1,2], Mona I. Pedersen[2], Ana Paola Giometti Lombardi[3], Lill-Tove Rasmussen Busund[3,4] & Thomas Karsten Kilvær[1,2]

In many types of cancer, microRNAs (miRs) are aberrantly expressed. The aim of this study was to explore the prognostic impact of miR-17-5p and miR-20a-5p in colon cancer. Tumor tissue from 452 stage I-III colon cancer patients was retrospectively collected and tissue microarrays constructed. miR-17-5p and miR-20a-5p expression was evaluated by in situ hybridization and analyzed using digital pathology. Cell line experiments, using HT-29 and CACO-2, were performed to assess the effect of miR-17-5p and miR-20a-5p over expression on viability, invasion and migration. In multivariate analyses, high miR-17-5p expression in tumor (HR = 0.43, CI 0.26–0.71, $p < 0.001$) and high expression of miR-20a-5p in tumor (HR = 0.60, CI 0.37–0.97, $p = 0.037$) and stroma (HR = 0.63, CI 0.42–0.95, $p = 0.027$) remained independent predictors of improved disease-specific survival. In cell lines, over expression of both miRs resulted in mitigated migration without any significant effect on viability or invasion. In conclusion, in stage I-III colon cancer, high expression of both miR-17-5p and miR-20a-5p are independent predictors of favorable prognosis.

Colon cancer is the 4th most and 5th most common cause of cancer and cancer related deaths, respectively[1]. In 2020, it is estimated that 1 150 000 patients experienced a de novo colon cancer and that 575 000 succumbed to their disease. Despite improved diagnostics and treatment, and decreasing incidence, the mortality of colon cancer remains high. To further tailor the treatment of these patients, the development of novel prognostic and predictive biomarkers is important[2,3].

MicroRNAs (miRs or miRNAs) are short, non-coding RNAs, approximately 22 nucleotides long. They regulate gene expression at the post-transcriptional level. An estimated 30% of human genes are regulated by miRs[4]. miRs influence diverse biological mechanisms including apoptosis, growth, differentiation and proliferation[5]. In cancer, miRs act as both oncogenes and tumor suppressors[6]. The miR-17 ~ 92 cluster, located at chromosomal locus 13q31.3, comprises six tandem stem-loop hairpin structures that yield six mature miRs (miR-17, miR-18a, miR-19a, miR-19b, miR-20a and miR-92a)[7]. There are two miR-17 ~ 92 cluster paralogs in mammals: miR-106b ~ 25 located on chromosome 7 and miR-106a ~ 363 located on the X-chromosome, comprising an additional 6 and 3 mature miRs. Collectively, these 15 mature miRs are grouped into four families, namely the miR-17, miR-18, miR-19 and miR-92[8]. The polycistronic structure of miR cluster genes differs from most protein coding genes, as multiple miRs can be produced within a single pri-miR transcript. Each of these can act independently[9]. miR-17 ~ 92 is predominantly related to cell cycle regulation. In normal development it is involved in lung and heart maturation and hematopoiesis, where it promotes cell proliferation and survival[10]. The cluster was first discovered in 2005, when it was found to act with c-MYC to promote tumorigenesis in B-cell lymphomas[11]. Over expression of the miR-17 ~ 92 cluster has been observed in multiple tumor types including hematological malignancies (B-cell lymphomas) and solid tumors (breast, lung, CRC, pancreas and prostate)[11,12].

The miR-17 family comprise miRs 17 and 20a in the miR-17 ~ 92 cluster, 106a and 20b in the miR-106a ~ 363 cluster and 106b and 93 in the miR-106b ~ 25 cluster[8]. Hence, miR-17 and miR-20a are the only members of the miR-17 family situated in the miR-17 ~ 92 cluster. Of interest, the expression level of mature miRs belonging to the same cluster are not equivalent[13]. An early discovery was that c-MYC activates expression of the

[1]Department of Oncology, University Hospital of North Norway, 9038 Tromso, Norway. [2]Department of Clinical Medicine, UiT The Arctic University of Norway, Tromso, Norway. [3]Department of Medical Biology, UiT The Arctic University of Norway, Tromso, Norway. [4]Department of Clinical Pathology, University Hospital of North Norway, Tromso, Norway. ✉email: hallgeir.selven@uit.no

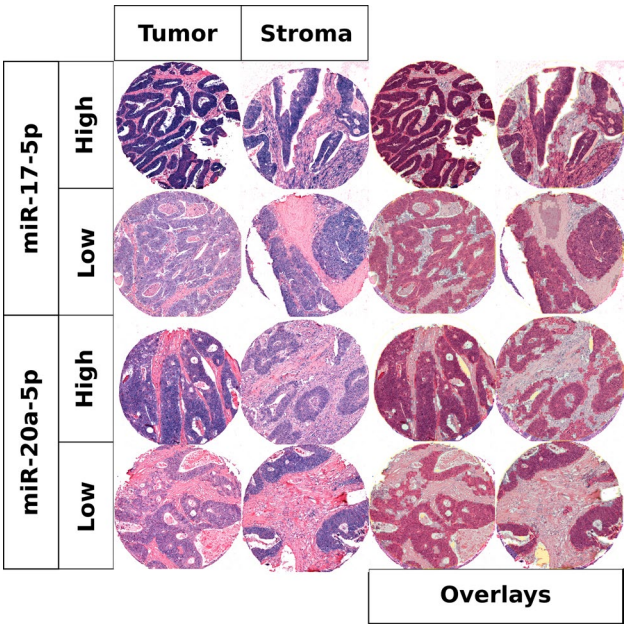

Figure 1. High and low scores for miR-17-5p and miR-20a-5p in tumor and stroma with and without overlays.

miR-17 ~ 92 cluster by binding directly to its locus on chromosome 13. c-MYC also targets transcription factor E2F1, promoting cell cycle progression. E2F1 is negatively regulated by both miR-17 and miR-20a. Thus, on one side c-MYC activates members of the E2F family of transcription factors, and on the other side limits their translation[14]. Subsequent studies identified that miR-17-5p and miR-20a-5p expression is suppressed by p53 and NKX3.1, stimulated by MXI1 and STAT, and that their expression suppress known regulators of cell death, cell cycle regulation, hypoxia, angiogenesis, and proliferation[15]. In gastrointestinal cancers (gastric, CRC and HCC), miR-17 has been related to increased cell proliferation, migration and invasion and reduced overall survival[16]. In contrast, miR-17 was shown to act as a tumor suppressor in breast-, cervical- and prostate cancer[17–19]. miR-20a regulates cell proliferation and cancer progression, and is dysregulated in both solid and hematopoietic cancers[15]. miR-20a was related to poor survival in lung- and gastric cancer, among others[20,21]. Whereas in oral squamous cell carcinoma and hepatocellular carcinoma, miR-20a acted as a tumor suppressor[22,23].

Previous studies in gastrointestinal cancers reported that both miR-17 and miR-20a frequently are over expressed and that miR-17 was associated with an unfavorable prognosis[16,24]. However, to our knowledge these studies included patients with advanced disease, combined separate entities such as colon and rectal adenocarcinoma and did not distinguish between expression in tumor epithelial cells and tumor stroma. Therefore, we explore the prognostic impact of miR-17-5p and miR-20a-5p expression in tissue from colon cancer patients treated with curative intent. We supplement our results with functional experiments in select early-stage colon cancer cell lines. And hypothesize that these miRs are clinically relevant biomarkers in localized colon cancer.

Results

Patient characteristics. The patient characteristics have previously been reported[25]. Briefly, 452 patients were included in the study. There was a minor female predominance (53.8% vs 46.2%), and the median age at surgery was 74 years (range 30–94). Seventy-two (15.9%), 219 (48.5%) and 161 (35.6%) patients were diagnosed with pTNM stage I-III, respectively. Median follow-up of survivors was 173 months. At the end of follow-up, 119 patients had recurrent disease and 313 patients were dead, either due to colon cancer (108) or other causes (205).

Expression of miR-17-5p and miR-20a-5p and their correlations. miR-17-5p and miR-20a-5p were mostly expressed in tumor epithelial cells and to a varying degree in stromal cells including spindle shaped cells (likely fibroblasts, endothelial cells and vascular smooth muscle cells) and immune cells (Fig. 1).

Correlations between miRs and clinicopathological variables are presented in Table 1. High expression of both miR-17-5p and miR-20a-5p in tumor tissue was associated with well and moderately differentiated tumors. In addition, miR-20a-5p in tumor was associated with cancers of the right colon. Between-miR correlations were as follows: miR-17-5p in tumor was correlated with mir-17-5p in stroma ($r = 0.27$), miR-20a-5p in tumor ($r = 0.52$) and stroma ($r = 0.17$); miR-17-5p in stroma was correlated to miR-20a-5p in tumor ($r = 0.16$) and stroma ($r = 0.37$); miR-20a-5p in tumor was correlated with miR-20a-5p in stroma ($r = 0.65$).

Cell line experiments. HT-29 and CACO-2 cell lines were tested for viability using MTT assays and invasion using transwell assays. No differences in viability or invasion were observed when either miR-17-5p or miR-20a-5p was over expressed compared to controls (Figs. 2 and 3). For migration analyses, using wound

	miR-17-5p in tumor			miR-17-5p in stroma			miR-20a-5p in tumor			miR-20a-5p in stroma		
	Low	High	*p*	Low	High	*p*	Low	High	*p*	Low	High	*p*
Age			0.115			0.947			0.742			0.472
≤65	80	27		26	83		76	31		27	80	
>65	213	110		83	252		244	89		72	266	
Gender			0.547			1.000			0.930			0.665
Female	154	77		59	181		170	65		51	189	
Male	139	60		50	154		150	55		48	157	
Weight_loss			**0.027**			0.818			0.055			0.507
<10%	151	81		62	180		172	71		52	194	
≥10%	70	19		22	71		73	16		23	68	
ECOG_status			0.827			0.093			0.526			0.066
0	150	76		47	185		163	68		45	188	
1	100	41		45	102		106	38		42	104	
2	35	16		14	39		42	11		9	45	
3	6	2		3	5		7	1		3	5	
Site			0.101			0.133			**0.023**			0.150
Sigmoid	158	58		48	177		165	56		40	183	
Transversum	42	20		16	49		53	11		17	48	
Left	12	6		3	16		16	5		7	14	
Right	79	52		42	90		83	48		34	99	
pStage			0.085			**0.006**			0.306			0.165
1	42	28		10	61		45	24		10	60	
2	139	70		48	168		160	57		49	170	
3	112	39		51	106		115	39		40	116	
Hist_grade			**0.015**			0.463			**0.026**			0.722
Well	22	13		10	26		20	14		6	29	
Moderate	203	111		81	240		230	93		74	251	
Poor	59	12		14	61		61	12		15	59	
Undifferentiated	3	1		0	4		3	0		1	2	
Vasc+			0.281			0.323			0.131			0.066
No	128	60		49	146		149	45		43	152	
Yes	14	3		7	11		17	1		8	10	
Resection margins			0.234			0.255			0.571			0.609
0 mm	25	4		12	19		24	5		6	25	
<1 mm	31	10		9	33		31	11		9	33	
1-2 mm	22	11		6	29		22	13		4	31	
2-10 mm	76	38		28	89		82	35		26	91	
10-50 mm	94	52		34	117		111	38		39	113	
>50 mm	30	15		15	31		33	13		11	35	

Table 1. Dichotomized miR-17-5p and miR-20a-5p in tumor and stroma and their distribution over and correlation with clinicopathological variables (chi-square and Fisher's exact tests). *ECOG*, Eastern cooperative oncology group, *pStage*, pathological stage, *Hist_grade*, histological grade, *Vasc+*, vascular infiltration. Statistically significant values in bold.

healing assays, reduced migration rates after miR-17-5p and miR-20a-5p over expression was observed for both cell lines (Fig. 4). These findings were statistically significant for both miR-17-5p ($p<0.001$) and miR-20a-5p ($p=0.029$) in the CACO-2 cell line. For the HT-29 cell line, the results for miR-20a-5p were statistically significant ($p=0.017$) and borderline significant for miR-17-5p ($p=0.052$).

Univariate analyses. Univariate survival analyses of clinicopathological variables were presented previously[25]. In brief, age, weight loss, pathological stage, histological grade, vascular infiltration, and resection margins were significant indicators of DSS. Univariate analyses of the investigated markers are presented in Table 2 and Fig. 5. High miR-17-5p expression in tumor was a significant indicator of DSS ($p=0.002$), while high miR-20a-5p expression was a significant indicator of DSS in tumor ($p=0.035$) and stroma ($p=0.003$).

Multivariate analyses. Multivariate analyses are summarized in Table 3. All significant variables from the univariate analyses were entered into the initial analyses. In the final models, high miR-17-5p expression in

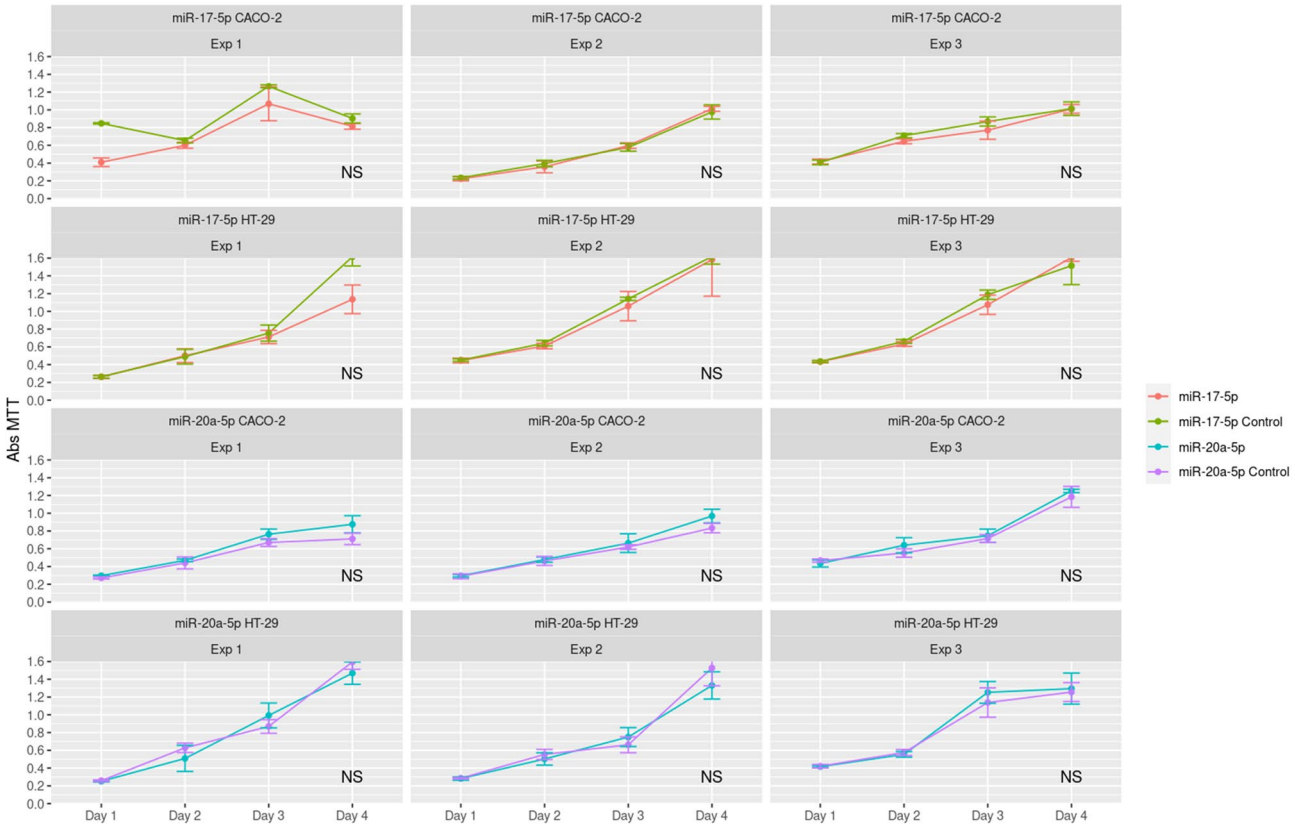

Figure 2. MTT assays comparing viability in cells transfected with either miR-17-5p (rows 1 and 2) or miR-20a-5p (rows 3 and 4) with control in the CACO-2 (rows 1 and 3) and HT-29 (rows 2 and 4) cell lines.

tumor was independently associated with favorable DSS (HR = 0.43, CI 0.26–0.71, p < 0.001), while high miR-20a-5p expression remained an independent predictor of favorable DSS in both tumor and stroma, (HR = 0.60, CI 0.37–0.97, p = 0.037) and (HR = 0.63, CI 0.42–0.95, p = 0.027), respectively.

Discussion

We present results showing favorable prognosis for colon cancer patients with high expression of miR-17-5p or miR-20a-5p. Patients with high expression of miR-17-5p in tumor tissue had a 5-year DSS of 86% compared with 75% for low expression. Patients with high expression of miR-20a-5p in tumor and stromal tissue had a 5-year DSS of 86% and 82% *vs* 77% and 71% for low expression, respectively (Table 2 and Fig. 1). Our results diverge from previous findings in similar cohorts and highlights important considerations for future studies on miRs in gastrointestinal cancer patients.

Contrasting our findings, most previous studies in gastrointestinal cancers related high miR-17-5p expression to impaired prognosis[26]. However, the number of high-quality studies exploring the prognostic value of miR-17-5p expression in localized colon cancer is limited. A meta-analysis from 2018, investigated the prognostic impact of miR-17-5p expression in gastrointestinal cancers[16]. Pooled analyses suggested that high expression of miR-17-5p predicted both poor overall survival (HR = 1.86, CI 1.55–2.25, p < 0.001) and poor disease-free survival (HR = 1.43, CI 1.01–2.03, p = 0.046). Of interest, ~ half of the identified publications reported non-significant results and/or were based on miR expression in serum/plasma, all studies included patients with advanced disease and several distinct cancers were represented[27]. Similar to our study, three studies report the prognostic impact of miR-17-5p expression in tissue from CRC patients[27–29]. Contrary to our study, these studies included stage I-IV patients, used overall survival as endpoint and did not distinguish between colon and rectal cancer. Both Ma and Fang et al. reported that high levels of miR-17-5p in tumor tissue, identified using ISH, was associated with impaired survival in Asian CRC patients[28,29]. Diaz et al. did not observe a survival difference in European patients[27]. This latter observation may suggest a demographic difference in CRC patients as proposed by others[30]. Moreover, although all three studies reported a similar percentage of miR-17-5p high patients, both Ma and Fang et al. observed increased miR-17-5p expression in stage III and IV patients[28,29]. Similar to our study, Diaz et al. observed a decline in miR-17-5p expression with increasing stage[27]. These results further corroborate the notion of a demographic difference for these biomarkers. In addition, neither Ma and Fang nor Diaz properly address the potential confounder introduced with patients with metastatic disease or rectal cancer.

Several previous studies stated that miR-20a-5p is upregulated both in feces and tumor tissue from colon cancer patients [12,31,32]. In our study, we observed that high expression of miR-20a-5p, both in tumor and stromal tissue, was related to a favorable disease-specific survival. Our findings contradict a recent meta-analysis

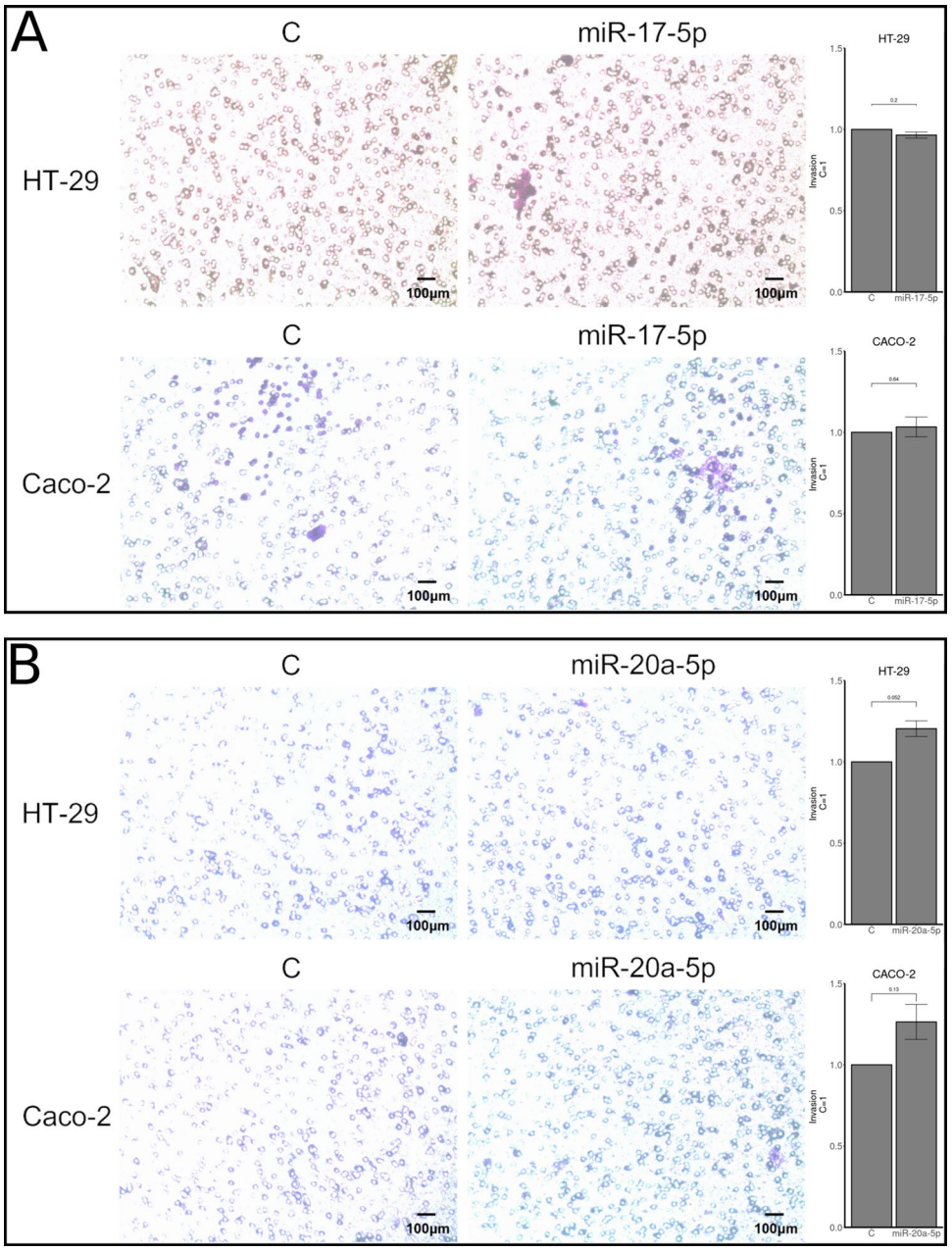

Figure 3. Transwell assays as measures of invasion/migration in CACO-2 and HT-29 cell lines transfected with miR-17-5p (panel A) or miR-20a-5p (panel B). Results are plotted as a mean of 3 experiments + /− SEM and relative to control (C = 1).

assessing the efficacy of miR-20a as a diagnostic and prognostic biomarker for colorectal cancer[24]. The meta-analysis, comprised of thirty-two studies, six including colon cancer patients only, concluded that miR-20a-5p expression was associated with impaired overall survival. However, similar to studies in miR-17-5p, several issues including differences in methodology, patient demographics and study endpoints precludes direct comparison with our study. Of interest, Signs et al. explored the impact of miR-20a-5p expression in the stromal compartment of colitis-associated cancer. They observed that stromal miR-20a-5p expression was higher in normal colon compared to a colitic or cancerous colon. Further, low levels of miR-20a-5p correlated with low levels of the inflammatory and oncogenic chemokine CXCL8 secreted by stromal fibroblasts. Stromal downregulation of miR-20a expression appeared to occur prior to epithelial upregulation. This suggests that downregulated miR-20a-5p expression in fibroblasts in the colitic field is responsible for the upregulation of CXCL8 responsible for tumorigenesis in colitis-associated cancer[33]. These findings are in line with our results, where high stromal expression of miR-20a-5p correlates to better outcome for the patients.

To further elucidate the role of miR-17-5p and miR-20a-5p, we investigated their functional aspects in two colon cancer cell lines. HT-29 and CACO-2 are known to form low-grade/early-stage cancer when grown in

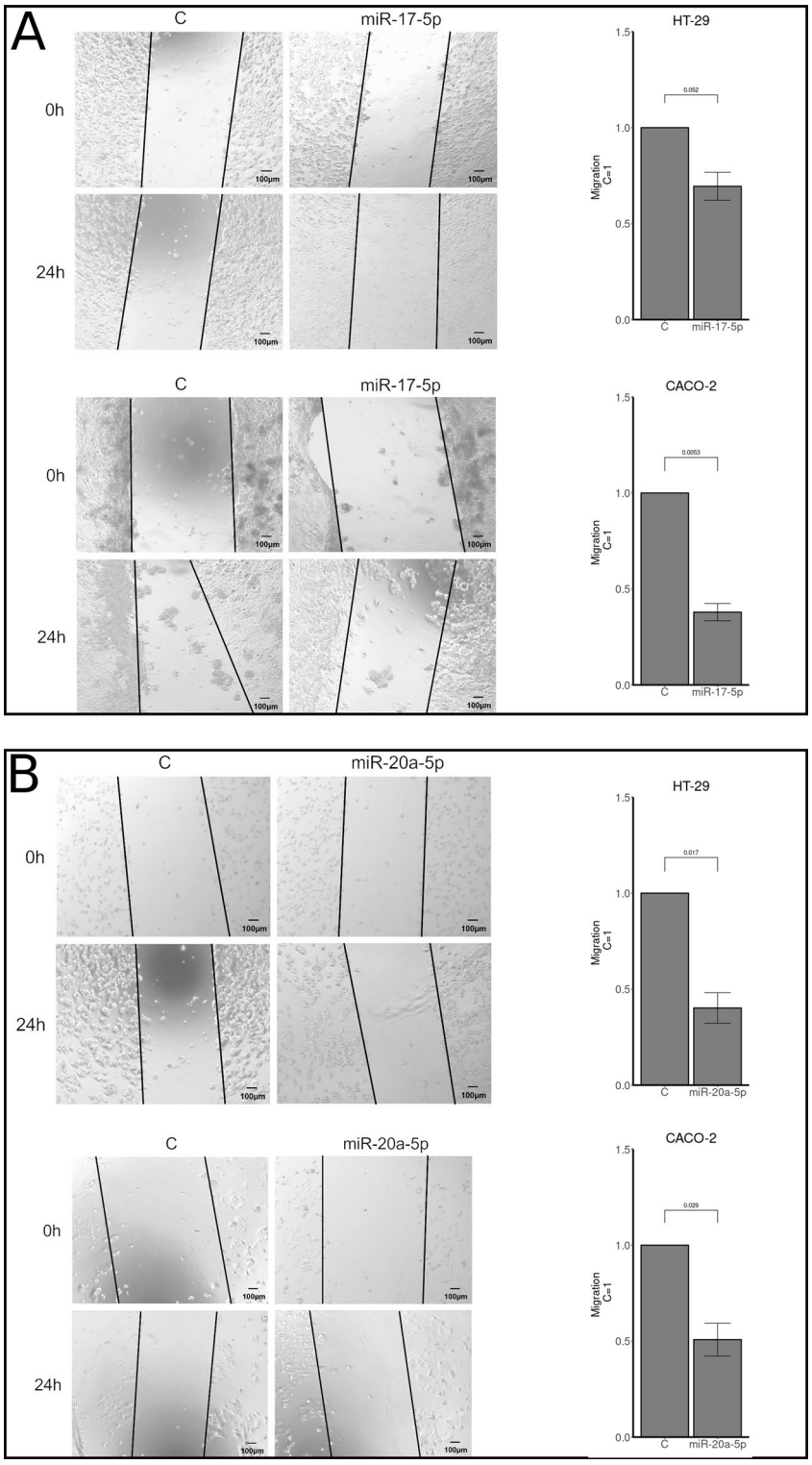

Figure 4. Wound healing assays as measures of migration in CACO-2 and HT-29 cell lines transfected with miR-17-5p (panel A) or miR-20a-5p (panel B). Results are plotted as a mean of 3 experiments + /− SEM and relative to control (C = 1).

nude mice and are thus likely representative of localized colon cancer[34,35]. Interestingly, over expression of the miRs did not impact viability or invasion and mitigated migration in both cell lines. These results strengthen our findings in patients with localized colon cancer, where over expression of miR-17-5p and miR-20a-5p predicted

	N	5Y	M	HR (95% CI)	*p*
miR-17-5p in tumor					**0.002**
Low	293(65)	75	NA	1.000	
High	137(30)	86	NA	0.48(0.32–0.72)	
Missing	22(5)				
miR-17-5p in stroma					0.053
Low	109(24)	73	NA	1.000	
High	335(74)	81	NA	0.67(0.42–1.05)	
Missing	8(2)				
miR-20a-5p in tumor					**0.035**
Low	315(70)	77	NA	1.000	
High	125(28)	86	NA	0.6(0.39–0.92)	
Missing	12(3)				
miR-20a-5p in stroma					**0.003**
Low	99(22)	71	NA	1.000	
High	346(77)	82	NA	0.55(0.35–0.88)	
Missing	7(2)				

Table 2. Univariate analyses of miR-17-5p and miR-20a-5p in tumor and stroma (log-rank, n = 452). *N*, number, *5Y*, 5-year survival, *M*, median survival, *p*, *p*-value, *HR*, hazard ratio, *CI*, confidence interval, *NA*, not applicable. Statistically significant values in bold.

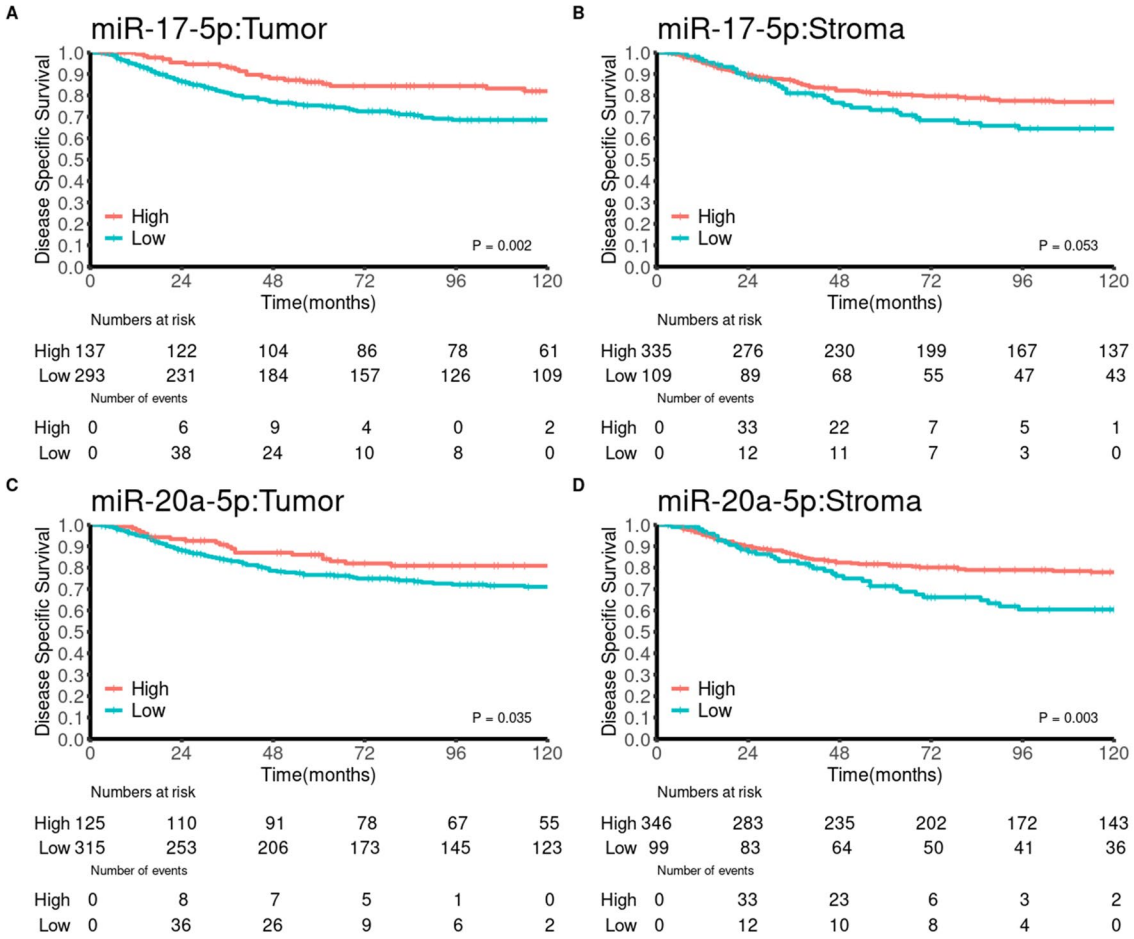

Figure 5. Disease-specific survival curves for expression of miR-17-5p and miR-20a-5p using the optimal cut-offs for each marker.

	I		II		III	
	HR (95% CI)	*p*	HR (95% CI)	*p*	HR (95% CI)	*p*
miR-17-5p in tumor						
Low	1					
High	0.43(0.26–0.71)	**<0.001**				
Missing						
miR-20a-5p in tumor						
Low			1			
High			0.60(0.37–0.97)	**0.037**		
Missing						
miR-20a-5p in stroma						
Low					1	
High					0.63(0.42–0.95)	**0.027**
Missing						
Age at diagnosis	1.03(1.01–1.05)	**0.002**	1.03(1.01–1.05)	**0.009**	1.02(1.01–1.04)	**0.013**
Differentiation						
Well differentiated	1					
Moderately differentiated	0.91(0.42–2.01)	0.822				
Poorly differentiated	0.61(0.25–1.51)	0.286				
Undifferentiated	7.97(1.93–32.95)	**0.004**				
pTNM						
pTNM-stage I	1		1		1	
pTNM-stage II	2.32(0.89–6.03)	0.085	2.18(0.85–5.58)	0.105	2.16(0.84–5.55)	0.108
pTNM-stage III	8.27(3.26–21.01)	**<0.001**	7.85(3.16–19.55)	**<0.001**	7.58(3.04–18.88)	**<0.001**

Table 3. Multivariate models for miR-17-5p in tumor (I) and for miR-20a-5p in tumor (II) and stroma (III), (cox proportional hazards tests, n = 452). *p*, *p*-value, *HR*, hazard ratio, *CI*, confidence interval, *pTNM*, pathological tumor-node-metastasis. Statistically significant values in bold.

better outcome for the patients. Corroborating our findings, several groups observed that over expression and/or suppression of miR-17-5p and miR-20a-5p subsequently mitigated and promoted migration in colon cancer cells[36,37]. Of particular interest, Ast et al. investigated the role of miR-17-5p and tumor-stromal cell interaction in the setting of CRC carcinogenesis[38]. By co-culturing colon cancer cells and colon fibroblasts transfected with a miR-17 mimic, they noticed significantly reduced cell invasion. Increased expression of miR-17 also significantly reduced the invasive activity of fibroblasts. However, other groups report that over expression of miR-17-5p and miR-20a-5p increases proliferation, migration and invasion in colon cancer cell lines, thus highlighting the complex role of these miRs in distinct settings[28,29,39].

The conflicting results hamper both interest and implementation of use of miR-17-5p and miR-20a-5p as biomarkers in colon cancer. Nevertheless, according to clinicaltrials.gov, Wu Song and co-workers are validating a signature of six miRs to predict chemotherapy response in stage II colon cancer[40]. This trial is based on their previous work[41], but no results are as of yet presented from their trial. Regardless of the outcome from the ongoing validation by Wu Song, our data indicate that positive results need to be validated and not automatically extrapolated to other demographic groups.

Conclusion

In conclusion, we have shown that high expression of miR-17-5p in tumor tissue and high expression of miR-20a-5p in both tumor and stroma are independent indicators of favorable disease-specific survival for localized colon cancer. Our findings contradict previous studies in colorectal cancer, and highlights that potential differences in methodology, patient demographics and endpoints may highly influence the prognostic value of these biomarkers. Further, although data from several pre-clinical studies and our cell line studies corroborates our findings, contrasting results exists also in this domain. Due to these contradictions, prospective trials resolving these issues have to be conducted before clinical implementation of miR-17-5p and miR-20a-5p as prognostic or predictive biomarkers in colon cancer can be considered.

Materials and methods

Study population. Patients who underwent radical surgery for colon cancer, in various hospitals in Northern Norway from 1998–2007, were eligible for inclusion. Initially, 861 patients were identified. Of these, 409 patients were excluded, mainly due to metastatic disease/prior malignancy within the last 5 years before diagnosis, missing tissue blocks/inadequate tissue for TMA construction or miscoding (mainly rectal cancer). Hence, 452 patients were included in the study. Follow-up was completed December 1, 2017. Detailed information about the study population was previously published[25].

Tissue Microarray construction. All colon cancer cases were reviewed by two pathologists, and the most representative areas of tumor without necrosis were selected. A 0.6 mm-diameter stylet was used to sample a total of 4 cores securing both tumor and stroma from each included patient. The TMAs were assembled using a tissue-arraying instrument (Beecher Instruments, Silver Springs, MD, USA). The detailed methodology has previously been reported[42].

In situ hybridization (ISH). The microRNA in situ hybridization method was performed on the Ventana Discovery Ultra platform for IHC and ISH. The protocol was developed by Roche, (Tucson, USA), based on the manual protocol previously published by Jorgensen et al.[43].

Double-DIG labeled miRCURY LNA detection probes and control probes from Exiqon (Exiqon AS, Denmark) was used to define the expression level of miR-20a-5p and miR-17-5p in colon cancer FFPE tissue. Detection kits and buffers purchased from Roche gave the chromogenic visualization of the microRNAs.

Slides were baked at 60 °C overnight, and then transferred to the Discovery Ultra for ISH staining. Sections were deparaffinized at 68 °C for three cycles in Ventana EZ buffer. Heat retrieval was performed at 95 °C with Discovery Cell Conditioning Solution (CC1) for 40 min to make access for the probes. Optimized concentrations of probe controls and target miR probes were manually applicated, miR-20a-5p, 50 nM, and miR-17-5p, 20 nM. The hybridization reaction was carried out for 60 min at 54 °C for miR-17-5p and 40 °C for miR-20a-5p followed by two stringency washes with 2.0X SSC buffer. Possible unspecific bindings were blocked with AB blocking solution for 16 min. Alkaline phosphatase-conjugated anti DIG (Anti-DIG-AP) was incubated for 20 min for immunologic detection. Substrate enzymatic reactions were carried out with NBT/BCIP for 60–120 min to give a blue precipitate. The slides were counterstained with Nuclear Fast Red for contrast staining. Slides were dehydrated through an increasing gradient of ethanol solutions to xylene and mounted with Histokitt mounting medium.

Good sensitivity level of the ISH method and minimal RNA degradation in tissue was confirmed by U6, snRNA control probe at a concentration of 1.5 nM. 10 nM scramble miR negative control indicated no unspecific staining from reagents or tissues. The level of microRNA expression in other tissues than colon cancer was confirmed by a TMA multi tissue control. Optimizations regarding temperatures, times, and concentrations were done for each probe and reagent.

In situ hybridization scoring/QuPath. The details of the digital workflow is described thoroughly in our previous paper[25]. In brief, TMA slides were digitized and processed in QuPath v.0.1.3 according to Bankhead et al.[44]. Tissue within each TMA core was identified and tiled. Image features were used to train a Random Forest model. Each tile was classified as either tumor, stroma, necrosis or other. After classification, tiles were converted into continuous areas and the mean intensity of miR-17-5p and miR-20a-5p within tumor and stroma was calculated. The scripts used to process the TMAs are included in the supplementary file.

All possible dichotomized cut-offs were evaluated. For any subsequent analyses, the optimal cut-off was chosen.

Cell line experiments. *Cell cultures.* The functional aspects of miR-17-5p and miR-20a-5p were tested in two colon cancer cell lines: CACO-2 (ATCC HTB-37) and HT-29 (ATCC HTB-38) both derived from colon adenocarcinoma. They have been authenticated and recently tested negative for mycoplasma contamination. The cell lines were tested for viability, migration and invasion in the absence and presence of miR-17-5p and miR-20a-5p as previously described by Stoen et al.[45,46]. The most important steps of each assay are referred below.

Cell culture and transfection. Cells were transiently transfected with either 10 μM has-miR-17-5p Pre-miR miRNA Precursor (catalog# PM12412, Thermo Fisher Scientific, USA) or has-miR-20a-5p Pre-miR miRNA Precursor (catalog# AM17100, Thermo Fisher Scientific, USA), alongside the Cy3 Dye-Labeled Pre-miR Negative Control #1 (catalog# AM17120, Thermo Fisher Scientific, USA) using the transfection reagent Lipofectamine RNAiMAX (catalog#13,778,075, Thermo Fisher Scientific, USA). Transfected Cy3 Dye-Labeled Pre-miR Negative Control emits fluorescent light when exposed to UV-light, and using a fluorescence microscope, the transfection efficiency was estimated to 80–95%.

Viability assay. Cells were cultured in 96-well plates and incubated with 12 mM of [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide] (MTT, 5 mg/ml) (cat.# M6494, Invitrogen, OR, USA). Formazan crystals were solubilized by addition of 0,01 M HCl/SDS (cat.# 28,312, Thermo Scientific, IL, USA) and the absorbance was measured in the CLARIOstar plate reader (BMG Labtech, Ortenberg, Germany) at 570 nm.

Migration/wound healing assay. Cells were grown in a 24-well plate, washed with PBS and incubated in a serum free medium with mitomycin C (10 μg/L) to avoid cell proliferation. The cells were "wounded" using a 200 μl sterile pipette tip and then washed to remove detached cells and debris. After 4 h the cells were transfected. To measure wound healing in controls and transfected cells, photographs of the same areas of the wound were taken at 0 and 24 h. Images were captured using a Nikon Eclipse TS100 inverted optical microscope and analyzed by Micrometrics SE Premium 4 software. Areas occupied by migrating cells after 24 h were calculated by subtracting the background levels at 0 h.

Invasion assay. Cells were seeded in ThincertR chambers (Greiner Bio-one, Kremsmünster, Austria) with polyethylene terephthalate membranes (8 mm pore size) pre-coated with 50 mL of phenol red-free Matrigel (Gibco). These chambers were placed in 24-well plates containing culture medium with 10% FBS in the lower

chamber. Cells in the upper chambers were transfected and incubated for 48 h at 37 °C. The chambers were washed thoroughly with 10 mM PBS, fixed in 4% paraformaldehyde for 30 min, and stained with 0.2% crystal violet for 10 min. Non-invading cells, from the membrane upper surface, were removed using a cotton swab. The membranes containing the invaded cells (under the membrane surface), were photographed. Images of three random microscope fields were captured in duplicate, using an inverted optical microscope Nikon Eclipse TS100. The areas of cell invasion were determined by Image J software.

Statistical methods. Statistical tests were performed in Rstudio 2021.09.0 build 351 (RStudio PBC) using R version 4.0.4. DSS (disease-specific survival) was defined as the interval from surgery to the time of colon cancer death. Before analyses, expression of all miRs were rescaled to a range between 0 and 1 using max–min scaling. For univariate analyses, the Kaplan–Meier method was used to visualize associations between molecular marker expression and survival. The log-rank test was used to assess the statistical significance of the differences between the survival curves. Multivariate analyses were performed using a backward conditional Cox regression analysis with a probability for stepwise entry and removal at 0.05 and 0.10, respectively. A p-value < 0.05 was considered statistically significant.

Ethics declaration. The study was conducted according to the guidelines of the Declaration of Helsinki, and was approved by the Regional Committee for Medical and Health Research Ethics North (REK Nord, protocol ID: 2011/2151). The need for informed consent was waived by REK Nord due to the retrospective nature of the study. The reporting of clinicopathological variables, survival data and biomarker expression was conducted in accordance with the REMARK guidelines[47].

Data availability

Data will be shared upon reasonable request to the corresponding author.

Received: 17 February 2022; Accepted: 15 April 2022
Published online: 30 April 2022

References

1. Global Cancer Observatory. Available at: https://gco.iarc.fr/. (Accessed: 12th January 2021)
2. Edwards, B. K. *et al.* Annual report to the nation on the status of cancer, 1975–2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. *Cancer* **116**, 544–573 (2010).
3. Fitzpatrick-Lewis, D. *et al.* Screening for colorectal cancer: A systematic review and meta-analysis. *Clin. Colorectal Cancer* **15**, 298–313 (2016).
4. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
5. Bartel, D. P. MicroRNAs: Target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
6. Zhang, B., Pan, X., Cobb, G. P. & Anderson, T. A. microRNAs as oncogenes and tumor suppressors. *Dev. Biol.* **302**, 1–12 (2007).
7. Tanzer, A. & Stadler, P. F. Molecular evolution of a microRNA cluster. *J. Mol. Biol.* **339**, 327–335 (2004).
8. Mendell, J. T. miRiad roles for the miR-17-92 cluster in development and disease. *Cell* **133**, 217–222 (2008).
9. Olive, V., Jiang, I. & He, L. Mir-17-92, a cluster of miRNAs in the midst of the cancer network. *Int. J. Biochem. Cell Biol.* **42**, 1348–1354 (2010).
10. Ventura, A. *et al.* Targeted deletion reveals essential and overlapping functions of the miR-17~92 family of miRNA clusters. *Cell* **132**, 875–886 (2008).
11. He, L. *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **435**, 828–833 (2005).
12. Volinia, S. *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U.S.A* **103**, 2257–2261 (2006).
13. Yu, J. *et al.* Human microRNA clusters: Genomic organization and expression profile in leukemia cell lines. *Biochem. Biophys. Res. Commun.* **349**, 59–68 (2006).
14. O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839–843 (2005).
15. Mogilyansky, E. & Rigoutsos, I. The miR-17/92 cluster: A comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. *Cell Death Differ.* **20**, 1603–1614 (2013).
16. Wang, Z. *et al.* Prognostic value of miR-17-5 p in gastrointestinal cancers: a systematic review and meta-analysis. *Onco. Targets. Ther.* **11**, 5991–5999 (2018).
17. Fan, M., Sethuraman, A., Brown, M., Sun, W. & Pfeffer, L. M. Systematic analysis of metastasis-associated genes identifies miR-17-5p as a metastatic suppressor of basal-like breast cancer. *Breast Cancer Res. Treat.* **146**, 487–502 (2014).
18. Wei, Q., Li, Y.-X., Liu, M., Li, X. & Tang, H. MiR-17-5p targets TP53INP1 and regulates cell proliferation and apoptosis of cervical cancer cells. *IUBMB Life* **64**, 697–704 (2012).
19. Ottman, R., Levy, J., Grizzle, W. E. & Chakrabarti, R. The other face of miR-17-92a cluster, exhibiting tumor suppressor effects in prostate cancer. *Oncotarget* **7**, 73739–73753 (2016).
20. Xu, X., Zhu, S., Tao, Z. & Ye, S. High circulating miR-18a, miR-20a, and miR-92a expression correlates with poor prognosis in patients with non-small cell lung cancer. *Cancer Med.* **7**, 21–31 (2018).
21. Wang, M. *et al.* Circulating miR-17-5p and miR-20a: Molecular markers for gastric cancer. *Mol. Med. Rep.* **5**, 1514–1520 (2012).
22. Chang, C. C. *et al.* MicroRNA-17/20a functions to inhibit cell migration and can be used a prognostic marker in oral squamous cell carcinoma. *Oral Oncol.* **49**, 923–931 (2013).
23. Fan, M. Q. *et al.* Decrease expression of microRNA-20a promotes cancer cell proliferation and predicts poor survival of hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* **32**, 21 (2013).
24. Moody, L., Dvoretskiy, S., An, R., Mantha, S. & Pan, Y. X. The efficacy of miR-20a as a diagnostic and prognostic biomarker for colorectal cancer: A systematic review and meta-analysis. *Cancers* **11**, 1111 (2019).
25. Selven, H., Busund, L.-T.R., Andersen, S., Bremnes, R. M. & Kilvær, T. K. High expression of microRNA-126 relates to favorable prognosis for colon cancer patients. *Sci. Rep.* **11**, 9592 (2021).
26. Huang, C., Yu, M. & Yao, X. MicroRNA-17 and the prognosis of human carcinomas: A systematic review and meta-analysis. *BMJ Open* **8**, e018070 (2018).

27. Díaz, R. *et al.* Deregulated expression of miR-106a predicts survival in human colon cancer patients. *Genes Chromosom. Cancer* **47**, 794–802 (2008).
28. Fang, L. *et al.* MicroRNA-17-5p promotes chemotherapeutic drug resistance and tumour metastasis of colorectal cancer by repressing PTEN expression. *Oncotarget* **5**, 2974–2987 (2014).
29. Ma, Y. *et al.* Elevated oncofoetal miR-17-5p expression regulates colorectal cancer progression by repressing its target gene P130. *Nat. Commun.* **3**, 1–12 (2012).
30. Bovell, L. C. *et al.* The prognostic value of MicroRNAs varies with patient race/ethnicity and stage of colorectal cancer. *Clin. Cancer Res.* **19**, 3955–3965 (2013).
31. Ahmed, F. E. *et al.* Diagnostic microRNA markers to screen for sporadic human colon cancer in stool: I. Proof of principle. *Cancer Genom. Proteomics* **10**, 93–113 (2013).
32. Schetter, A. J. *et al.* MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *J. Am. Med. Assoc.* **299**, 425–436 (2008).
33. Signs, S. A. *et al.* Stromal miR-20a controls paracrine CXCL8 secretion in colitis and colon cancer. *Oncotarget* **9**, 13048 (2018).
34. HT-29|ATCC. Available at: https://www.atcc.org/products/htb-38. (Accessed: 2nd February 2022)
35. Caco-2 [Caco2]|ATCC. Available at: https://www.atcc.org/products/htb-37. (Accessed: 2nd February 2022)
36. Kim, T. W. *et al.* MicroRNA-17–5p regulates EMT by targeting vimentin in colorectal cancer. *Br. J. Cancer* **123**, 1123–1130 (2020).
37. Jiang, H. *et al.* Quantitatively controlling expression of miR-17~92 determines colon tumor progression in a mouse tumor model. *Am. J. Pathol.* **184**, 1355–1368 (2014).
38. Ast, V. *et al.* MiR-192, miR-200c and miR-17 are fibroblast-mediated inhibitors of colorectal cancer invasion. *Oncotarget* **9**, 35559–35580 (2018).
39. Xi, X. P. *et al.* MicroRNA-17 induces epithelial-mesenchymal transition consistent with the cancer stem cell phenotype by regulating CYP7B1 expression in colon cancer. *Int. J. Mol. Med.* **38**, 499–506 (2016).
40. microRNAs Tool for Stratifying Stage II Colon Cancer—Tabular View—ClinicalTrials.gov. Available at: https://clinicaltrials.gov/ct2/show/record/NCT02635087?term=mir-20&draw=2&rank=3&view=record. (Accessed: 9th February 2022)
41. Zhang, J. X. *et al.* Prognostic and predictive value of a microRNA signature in stage II colon cancer: A microRNA expression analysis. *Lancet Oncol.* **14**, 1295–1306 (2013).
42. Bremnes, R. M. *et al.* High-throughput tissue microarray analysis used to evaluate biology and prognostic significance of the E-cadherin pathway in non-small-cell lung cancer. *J. Clin. Oncol.* **20**, 2417–2428 (2002).
43. Jørgensen, S., Baker, A., Møller, S. & Nielsen, B. S. Robust one-day in situ hybridization protocol for detection of microRNAs in paraffin samples using LNA probes. *Methods* **52**, 375–381 (2010).
44. Bankhead, P. *et al.* QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **7**, 16878 (2017).
45. Stoen, M. J. *et al.* High expression of miR-17–5p in tumor epithelium is a predictor for poor prognosis for prostate cancer patients. *Sci. Rep.* **11**, 1–12 (2021).
46. Stoen, M. J. *et al.* Overexpression of mir-20a-5p in tumor epithelium is an independent negative prognostic indicator in prostate cancer—A multi-institutional study. *Cancers (Basel)* **13**, 4096 (2021).
47. Sauerbrei, W., Taube, S. E., McShane, L. M., Cavenagh, M. M. & Altman, D. G. Reporting recommendations for tumor marker prognostic studies (REMARK): An abridged explanation and elaboration. *J. Natl Cancer Inst.* **110**, 803–811 (2018).

Acknowledgements

We would like to thank Magnus L Persson for preparing the TMA-slides.

Author contributions

All authors contributed to study design. H.S collected the clinical data. L.T.B collected and revised the pathological specimens. M.I.P conducted the ISH experiments. A.P.L conducted the cell line experiments. H.S, T.K.K and S.A analyzed and interpreted the data. H.S and T.K.K wrote the manuscript. All authors revised the manuscript.

Funding

Open Access funding provided by UiT The Arctic University of Norway. The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-11090-2.

Correspondence and requests for materials should be addressed to H.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

cancers

MDPI

Article

High Expression of IRS-1, RUNX3 and SMAD4 Are Positive Prognostic Factors in Stage I–III Colon Cancer

Hallgeir Selven [1,2], Lill-Tove Rasmussen Busund [3,4], Sigve Andersen [1,2], Mona Irene Pedersen [2], Ana Paola Giometti Lombardi [2] and Thomas Karsten Kilvaer [1,2,*]

[1] Department of Oncology, University Hospital of North Norway, 9038 Tromsø, Norway
[2] Department of Clinical Medicine, UiT The Arctic University of Norway, 9038 Tromsø, Norway
[3] Department of Pathology, University Hospital of North Norway, 9038 Tromsø, Norway
[4] Department of Medical Biology, UiT The Arctic University of Norway, 9038 Tromsø, Norway
* Correspondence: thomas.k.kilvar@uit.no; Tel.: +47-905-24-635

Simple Summary: We studied the expression of several protein biomarkers in both stromal and tumor tissue from colon cancer patients. High expression of IRS1 in stromal tissue and RUNX3 and SMAD4 in both stromal and tumor tissue were positive prognostic factors. Of particular interest, RUNX3 expression in stromal tissue was associated with the density of tumor-infiltrating lymphocytes. This finding may be important for understanding the prognostic impact of lymphocytes and predicting and increasing the efficacy of immunotherapy in colon cancer.

Abstract: Colon cancer is a common malignancy and a major contributor to human morbidity and mortality. In this study, we explore the expression and prognostic impact of IRS-1, IRS-2, RUNx3, and SMAD4 in colon cancer. Furthermore, we elucidate their correlations with miRs 126, 17-5p, and 20a-5p, which are identified as potential regulators of these proteins. Tumor tissue from 452 patients operated for stage I–III colon cancer was retrospectively collected and assembled into tissue microarrays. Biomarkers' expressions were examined by immunohistochemistry and analyzed using digital pathology. In univariate analyses, high expression levels of IRS1 in stromal cytoplasm, RUNX3 in tumor (nucleus and cytoplasm) and stroma (nucleus and cytoplasm), and SMAD4 in tumor (nucleus and cytoplasm) and stromal cytoplasm were related to increased disease-specific survival (DSS). In multivariate analyses, high expression of IRS1 in stromal cytoplasm, RUNX3 in tumor nucleus and stromal cytoplasm, and high expression of SMAD4 in tumor and stromal cytoplasm remained independent predictors of improved DSS. Surprisingly, with the exception of weak correlations ($0.2 < r < 0.25$) between miR-126 and SMAD4, the investigated markers were mostly uncorrelated with the miRs. However, weak to moderate/strong correlations ($0.3 < r < 0.6$) were observed between CD3 and CD8 positive lymphocyte density and stromal RUNX3 expression. High expression levels of IRS1, RUNX3, and SMAD4 are positive prognostic factors in stage I–III colon cancer. Furthermore, stromal expression of RUNX3 is associated with increased lymphocyte density, suggesting that RUNX3 is an important mediator during recruitment and activation of immune cells in colon cancer.

Keywords: colon cancer; RUNX3; SMAD4; biomarker; prognosis

Citation: Selven, H.; Busund, L.-T.R.; Andersen, S.; Pedersen, M.I.; Lombardi, A.P.G.; Kilvaer, T.K. High Expression of IRS-1, RUNX3 and SMAD4 Are Positive Prognostic Factors in Stage I–III Colon Cancer. *Cancers* **2023**, *15*, 1448. https://doi.org/10.3390/cancers15051448

Academic Editor: Giovanni Monteleone

Received: 9 February 2023
Revised: 21 February 2023
Accepted: 22 February 2023
Published: 24 February 2023

1. Introduction

The colon comprises one fifth of the digestive tract's length. However, despite its small lengthwise contribution, the colon harbors more than 40% of digestive tract cancers. With an estimated incidence of 106,000, colon cancer will be the fifth most common malignancy in the USA in 2022. Moreover, it will be the second most common cause of cancer-related deaths [1]. Despite the increasing knowledge of colon cancer etiology, patients' prognoses have not improved significantly over the last decade. Consequently, novel prognostic and

predictive biomarkers are needed to (1) identify patients at high risk of colon cancer death and (2) to select the right candidates for adjuvant and novel treatments.

In our previous works, we found that high expression levels of miRs 126, 17-5p, and 20a-5p were positive prognosticators in early stage colon cancer patients [2,3]. In an attempt to further elucidate the role of miRs 126, 17-5p, and 20a-5p, we chose to focus on IRS-1 and 2, RUNX3, and SMAD4. These proteins are four members of three distinct protein families thought to be regulated by these miRs [4–6]. They are directly involved in important signaling pathways including the ERK, PI3K/AKT, TGF-β pathways, among others [7–10]. Moreover, they are already implicated in colon cancer development through cell line experiments and a few prognostic studies [11–13]. IRSs are important in insulin signaling and maintain cellular functions such as growth, survival, and metabolism [14]. There are six known IRS-substrates, IRS-1 to IRS-6, of which IRS-1 and IRS-2 are widely expressed in humans [15]. IRSs can be oncogenic and induce malignant transformation. In addition, IRSs are required to facilitate the transforming ability of other oncogenes, depending on IRS tyrosine phosphorylation. IRS-1 overexpression with subsequent ERK 1/2 pathway activation in fibroblasts was shown to promote cellular proliferation. Previous studies have shown that IRS-1 and 2 are overexpressed in hepatocellular carcinoma (HCC) and pancreatic cancer [16–19]. However, IRS-1 expression was decreased in squamous cell lung cancer. Moreover, IRS-1 is constitutively activated in several sarcomas and in breast cancer [20], whereas IRS-2 is constitutively activated in patients with the hereditary condition multiple endocrine neoplasia type 2 (MEN2) [10].

RUNX3 is a protein belonging to the runt domain family of transcription factors involved in mammalian developmental pathways [21]. The RUNX3 gene is localized at the 1p36 locus, a region believed to harbor one or more tumor suppressor genes, as loss of heterozygosity in this region is observed in gastric, breast, ovarian, and colon cancers [22]. Activation of TGF-β phosphorylates SMAD3, which associates with SMAD4 and enters the nucleus. In the nucleus, SMAD3/SMAD4 forms complexes with RUNX3, thereby mediating the suppressive effects of TGF-β [8]. The tumor suppressive effect of RUNX3 can be inactivated in several ways (mutations, methylation-related transcriptional silencing, and mis-localization to the cytoplasm) [23]. Furthermore, RUNX3 is an important mediator during the development of both CD4+ and CD8+ cytotoxic T-lymphocytes (CTL), and is thus likely pivotal in the development of anti-tumor immunity [24]. The polarization of CD4+ towards CTLs may be especially important in the gut epithelium [25].

SMAD is a family of transcription factors, acting as mediators of the TGF-β signaling cascade. There are three functional classes of SMAD proteins in mammals: the receptor-regulated SMADs (SMAD1, 2, 3, 5, and 8); the co-mediator SMAD (SMAD4); and the inhibitory SMADs (SMAD6 and 7). The TGF-β/SMAD4 signaling pathway controls a wide range of cellular processes including proliferation, differentiation, apoptosis, migration, and cancer initiation and progression [26]. In early tumorigenesis, the TGF-β/SMAD4 signaling pathway acts as a tumor suppressor, inducing cell cycle arrest and apoptosis. However, as cells progress through tumorigenesis, they become refractory to TGF-β-mediated growth inhibition and respond by stimulating pathways resulting in TGF-β-mediated tumor progression. Numerous other pathways interact with the TGF-β/SMAD4 pathway, including MAPK, ERK, and PI3K/AKT [27–29]. SMAD4's role in cancer was first discovered in 1996, where SMAD4 gene alterations were shown to be closely related to pancreatic cancer [30]. Loss of SMAD4 has also been reported in cholangiocarcinomas and colorectal cancer, among others [31,32].

In this study, we apply deep learning to digitized pathology images to explore compartment level expression and prognostic impact of IRS-1, IRS-2, RUNX3, and SMAD4 in resected tumors from 452 colon cancer patients. Furthermore, we elucidate their correlations with the expression of miRs 126, 17-5p, and 20a-5p as well as the density of CD3+ and CD8+ tumor-infiltrating lymphocytes.

2. Materials and Methods

2.1. Study Population

Patients undergoing radical surgery for colon cancer in various hospitals in Northern Norway in the time period of 1998–2007 were eligible for inclusion. From an initial 861 identified patients, 452 patients were finally included in the study. The main exclusion criteria were metastatic disease/prior malignancy within the last 5 years before diagnosis, missing tissue blocks/inadequate tissue for TMA construction, and faulty coding (rectal cancer, mainly). Follow-up was completed on 1 December 2017. The study population was previously published in detail [2].

2.2. Tissue Microarray Construction

All colon cancer cases were reviewed by two pathologists. The most representative areas of tumor without necrosis were selected. Using a 0.6 mm diameter stylet, a total of 4 cores were sampled for each included patient, securing both tumor and stromal tissue. The TMAs were assembled using a tissue-arraying instrument (Beecher Instruments, Silver Springs, MD, USA). The detailed methodology has been previously reported [33]. Sections were cut on a MICROM HM 335 S microtome, transferred to Super Frost Plus slides, and dried at room temperature before staining.

2.3. Immunohistochemistry and In Situ Hybridization

The Discovery Ultra Research instrument Roche 05987750001 was used to examine the protein expression of the six biomarkers in colon cancer TMAs. The antibodies used for this study were IRS1 ab40077, IRS2 ab134101, SMAD4 ab40759, and RUNX3 ab135248, sourced from Abcam. In addition, we used CD3 (2GV6) Roche 05278422001 and CD8 (SP57) Roche 05937248001 for in vitro diagnostic (IVD) use. All antibodies were validated for IHC-P (formalin fixed and paraffin embedded tissue) by the supplier. Optimization of dilutions, incubation times and temperatures were done in-house. Advised positive tissue controls from supplier were tested for each antibody. Staining and antibody specificity was verified by an internal tissue control (TMA multi control) containing several normal and cancer tissues. Negative controls were conducted by omitting the primary antibody, The negative controls were mainly clean, but weak brown non-specific staining in RUNX3 DAB stain was observed. Details of the optimized IHC protocols are given in Table S1. Table S2 shows product information of antibodies and reagents. The methodology for in situ hybridization of miRs 126, 17-5p, and 20a-5p was previously described [2,3].

2.4. Digitization/Immunohistochemistry Scoring

TMA slides were digitized using a Pannoarmic Flash III digital slide scanner (3DHistech, Budapest, Hungary) and processed in QuPath vs. 0.3.2 according to Bankhead et al. [34]. Only cores containing tumor tissue were used for analyses. Cells were identified and classified using a StarDist deep learning (DL) model trained on the hematoxylin channel of the Lizard dataset [35–37]. Briefly, we extracted the hematoxylin optical density image channel, normalized it, and applied the DL model. As an output, we obtained segmented nuclei classified into six classes (neutrophil, epithelial, lymphocyte, plasma, eosinophil, and connective). Due to limitations with the DL method, lymphocytes, plasma cells, eosinophils and connective tissue cells were combined into a single stromal class. The final analyses were conducted on the tumor epithelial cells and the combined stromal class cells. In addition, minor filtering based on size and circularity was applied. The mean marker intensity and the estimated cytoplasm (an arbitrary expansion from the nucleus) were then calculated separately for each nucleus. The final score for each compartment is the median score of all its nuclei. QuPath scripts to run and generate the final scores for each marker are available upon request to the corresponding author. An optimal cutoff strategy was applied to dichotomize markers for survival analyses. For non-significant markers, the median cut-off was chosen.

2.5. Statistics

Statistical tests were performed in Rstudio 2021.09.0 build 351 (RStudio PBC) using R version 4.0.4. Disease-specific survival (DSS) was defined as the interval from surgery to the time of colon cancer death. Before analyses, expressions of the investigated markers were rescaled to a range between 0 and 1 using max–min scaling. Kaplan–Meier plots visualized the dichotomized molecular marker's impact on patient survival. The differences between the survival curves were tested using the log-rank test. Multivariable analyses were conducted using Cox regression. All significant variables from the univariate analyses available to a majority of the patients were entered into the initial models. A sequential backward conditional approach was adapted, where variables with p-values above 0.1 were dropped at each step. Chi-squared and Fischer's exact tests were used to examine the association between molecular marker expression and clinicopathological variables. Pearson correlation was used to examine the associations between marker expressions. *r* values of ±0, 0–0.2, 0.2–0.3, 0.3–0.5, 0.5–0.7 and >0.7 were considered negative, weak, weak/moderate, moderate, moderate/strong, and strong, respectively. Hierarchical clustering with distance calculated based on the *r* values of their correlations was applied to visualize patterns in the correlation data.

3. Results

3.1. Patient Characteristics

The patients characteristics were previously reported [2]. A total of 452 patients were included in the study. All patients were treated with curative intent, and most patients were diagnosed with pTNM stage II (48.5%) and III (35.6%). Survivors were followed up on for a median of 173 months. At the end of follow-up period, 119 patients experienced a recurrence and 108 had succumbed to their disease.

3.2. Expression of SMAD4, RUNX3, IRS-1, and IRS-2 and Their Correlations with Clinicopathological Variables

Expression of the investigated markers are illustrated in Figure 1. As can be seen, SMAD4 and RUNX3 were expressed both in nucleus and cytoplasm whereas IRS-1 and IRS-2 expression was restricted to the cytoplasm. IRS-1 and 2 and SMAD4 expression was evenly distributed in tumor epithelial cells, spindle shaped cells/stromal cells, and immune cells. RUNX3 was predominantly expressed in tumor infiltrating lymphocytes (TILs) and to a lesser extent in other cells. IRS-1 and 2 were in some cases highly expressed in collagen-like sheets.

Correlations between investigated biomarkers and clinicopathological variables are presented in Table S3. Expression of SMAD4 in tumor cytoplasm was correlated with weight loss and pathological stage, whereas SMAD4 in stromal cytoplasm was correlated with pathological stage only. SMAD4 in stromal nucleus was correlated with site and histological grade. Expression of RUNX3 in tumor nucleus and cytoplasm were both correlated with site. Expression of RUNX3 in stromal nucleus and cytoplasm were both correlated with site, and RUNX3 in stromal nucleus was correlated with histological grade, whereas RUNX3 in stromal cytoplasm was correlated with pathological stage. Expression of IRS-1 in stromal cytoplasm was correlated with ECOG status, and IRS-2 in stromal cytoplasm was correlated with weight loss.

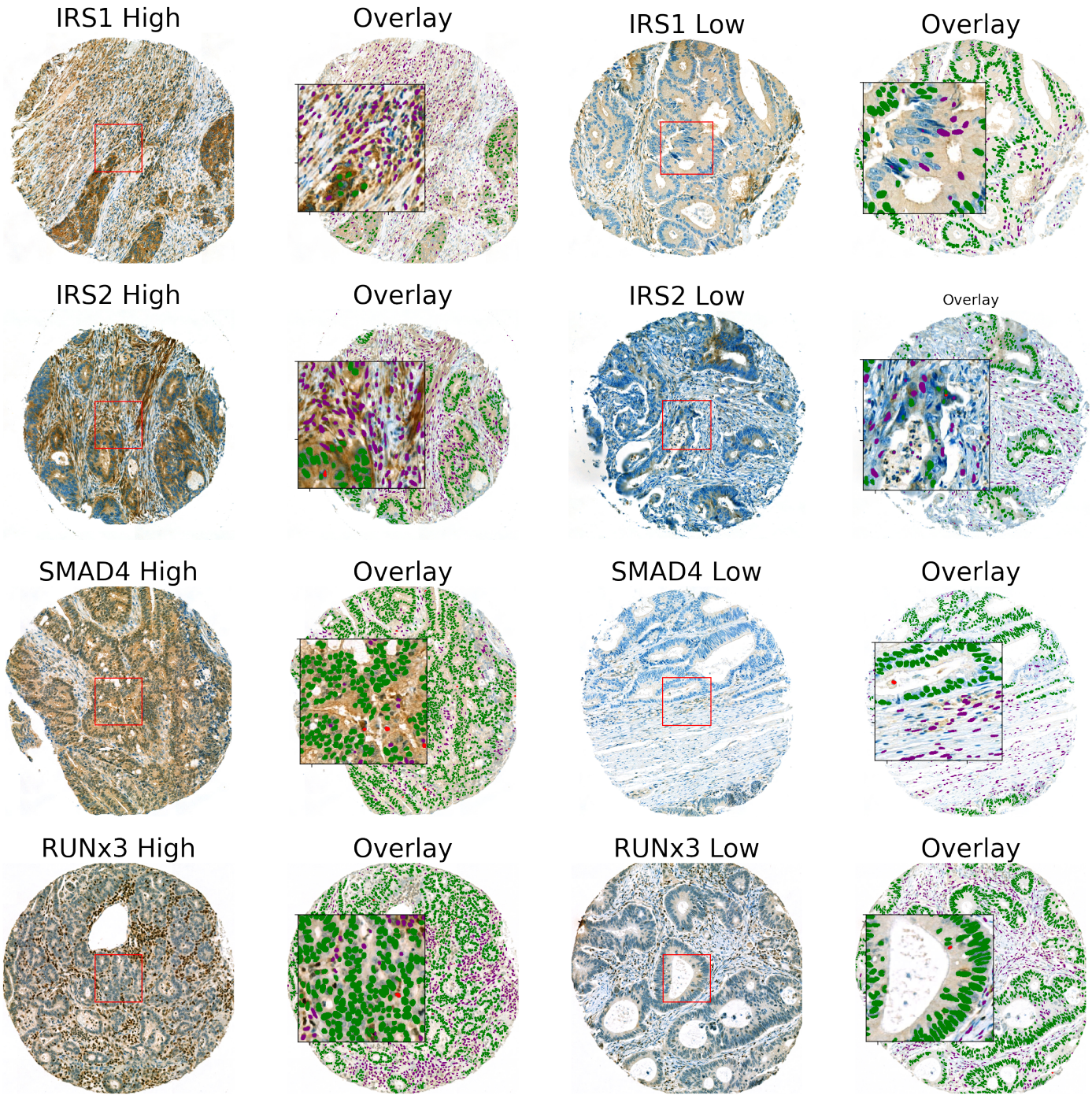

Figure 1. High and low expression of IRS1, IRS2, SMAD4, and RUNX3 in tumor and stromal cells. For SMAD4 and RUNX3, representative cores with high expression in both nucleus and cytoplasm were chosen.

3.3. Correlations between Investigated Biomarkers and CD3, CD8, miR-17-5p, miR-20a-5p and miR-126

Correlations between investigated markers, miRs, and lymphocyte markers are presented in Figure 2. As expected, moderate to strong correlations were observed between epithelial and stromal expression for each marker. Furthermore, extensive correlations between IRS-1 and IRS-2, RUNX3 and SMAD4 in tumor and stroma, were observed ($0.15 < r < 0.60$). S-miR-126 showed a weak/moderate correlation with SMAD4 ($0.20 < r < 0.30$) and S-miR-126 and S-miR-17-5p showed weak/moderate correlations with CD3+ and CD8+ TIL density ($0.15 < r < 0.30$). No other relevant correlations between miRs and other investigated

markers were observed. Of note, CD3+ and CD8+ TILs density showed moderate/strong correlations with S-RUNX3 ($0.35 < r < 0.60$) and weak/moderate correlations with S- and T-SMAD4 ($0.15 < r < 0.30$).

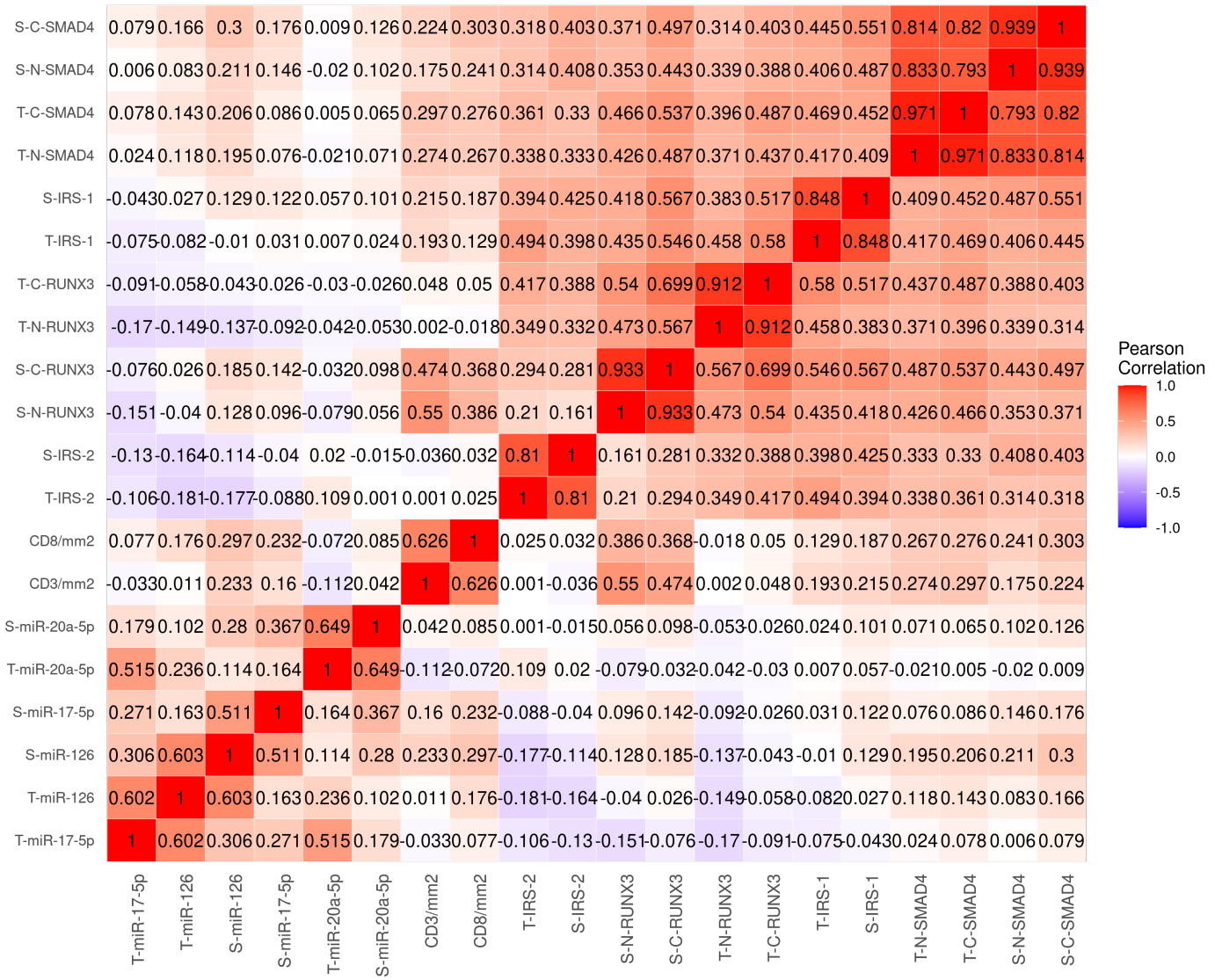

Figure 2. Correlations between IRS1, IRS2, SMAD4, and RUNX3 in tumor and stromal cell nucleus and/or cytoplasm, miR17, miR20a, and miR126 in tumor and stromal compartments and CD3+ and CD8+ cell density. Abbreviations: S, stroma; T, tumor; C, cytoplasm; N, nucleus.

3.4. Univariate Analyses

Univariate survival analyses of clinicopathological variables in this cohort have been reported earlier, showing that age, weight loss, pathological stage, histological grade, vascular infiltration, and resection margins were significant indicators of DSS [2]. Univariate analyses of investigated biomarkers are presented in Table 1 and Figure 3. Increased expression of SMAD4 in nucleus and cytoplasm, the tumor epithelial compartment, and cytoplasm in the stromal compartment; RUNX3 in the nucleus or cytoplasm in both the tumor epithelial and stromal compartments; and IRS-1 in the stromal cytoplasm were all significant predictors of a favorable DSS.

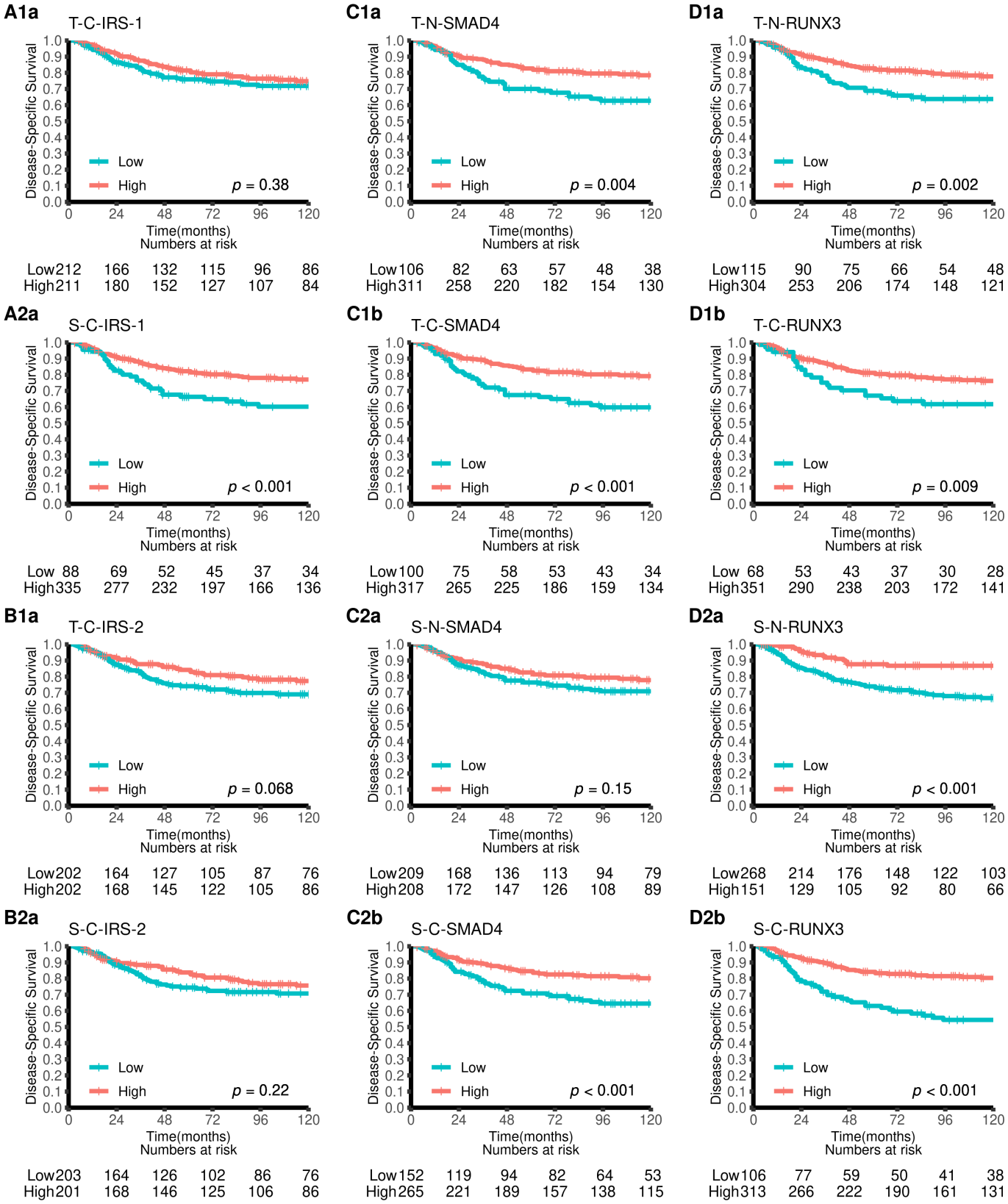

Figure 3. Disease-specific survival curves for the expression of IRS1 and IRS in tumor cell cytoplasm (**A1**–**A2**,**B1**–**B2**), SMAD4 (**C1**–**C2**), and RUNX3 (**D1**–**D2**) in tumor cell nucleus and cytoplasm using the optimal cutoffs for each marker. Abbreviations: S, stroma; T, tumor; C, cytoplasm; N, nucleus.

Table 1. Univariate analyses of tumor IRS1, IRS2, SMAD4, and RUNx3 in tumor and stromal cell nucleus and/or cytoplasm (log-rank test, n = 452).

	Tumor					**Stromal**				
	N(%)	**5 Year**	**Median**	**HR(95% CI)**	***p***	**N(%)**	**5 Year**	**Median**	**HR(95% CI)**	***p***
C-IRS1					0.380					<0.001
Low	212(47)	76	NA	1		88(19)	66	182	1	
High	211(47)	81	NA	0.84(0.57–1.24)		335(74)	82	NA	0.5(0.31–0.82)	
Missing	29(6)					29(6)				
C-IRS2					0.068					0.220
Low	202(45)	74	NA	1		203(45)	74	NA	1	
High	202(45)	84	NA	0.69(0.46–1.03)		201(44)	84	NA	0.78(0.52–1.16)	
Missing	48(11)					48(11)				
N-SMAD4					0.004					0.150
Low	106(23)	70	NA	1		209(46)	76	NA	1	
High	311(69)	83	NA	0.55(0.35–0.88)		208(46)	83	NA	0.74(0.5–1.11)	
Missing	35(8)					35(8)				
C-SMAD4					<0.001					<0.001
Low	100(22)	67	NA	1		152(34)	71	NA	1	
High	317(70)	83	NA	0.48(0.29–0.77)		265(59)	85	NA	0.49(0.32–0.74)	
Missing	35(8)					35(8)				
N-RUNX3					0.002					<0.001
Low	115(25)	69	NA	1		268(59)	74	NA	1	
High	304(67)	83	NA	0.53(0.34–0.83)		151(33)	88	NA	0.37(0.25–0.56)	
Missing	33(7)					33(7)				
C-RUNX3					0.009					<0.001
Low	68(15)	67	NA	1		106(23)	63	182	1	
High	351(78)	81	NA	0.55(0.32–0.95)		313(69)	84	NA	0.36(0.23–0.58)	
Missing	33(7)					33(7)				

Abbreviations: C, cytoplasm; N, nucleus; NA, not applicable.

3.5. Multivariate Analyses

Multivariable analyses are summarized in Table 2. Increasing age at diagnosis, pathological stage III, and a resection margin of <1 mm were unfavorable predictors of DSS. Furthermore, tumor epithelial expression of SMAD4 in cytoplasm (HR 0.58, 95% CI 0.43–0.80, $p < 0.001$) and RUNX3 in nucleus (HR 0.62, 95% CI 0.45–0.84, $p = 0.002$) were independent positive predictors of DSS. In the stromal compartment, IRS-1 in cytoplasm (HR 0.64, 95% CI 0.47–0.87, $p = 0.005$), SMAD4 in cytoplasm (HR 0.67, 95% CI 0.5–0.91, $p = 0.009$), and RUNX3 in cytoplasm (HR = 0.62, 95% CI 0.44–0.87, $p = 0.006$) were independent predictors of a favorable DSS.

3.6. Co-Expressions

Co-expression analyses between SMAD4 in tumor epithelial and stromal cytoplasm and RUNX3 in tumor epithelial nuclei and stromal cytoplasm are presented in Table S4 and Figure S1. Patients with increased/preserved SMAD4 and RUNX3 expression in either the tumor epithelium (HR = 0.34, 95% CI 0.18–0.66, $p = 0.001$) or the stromal compartment (HR = 0.34, 95% CI 0.2–0.57, $p < 0.001$) had significantly better prognoses compared to those with decreased/lost expression.

Table 2. Multivariable models including statistically significant clinicopathological variables and investigated biomarkers from univariable analyses (Cox proportional hazards test, n = 452). Separate models for each marker in each compartment (tumor T1 and T2 and stromal S1–S3).

	Tumor	**Stromal**								
	T1	**T2**	**S1**	**S2**	**S3**					
	HR(95% CI)	***p***	**HR(95% CI)**	***p***	**HR(95% CI)**	***p***	**HR(95% CI)**	***p***	**HR(95% CI)**	***p***
Age	1.03(1.01–1.05)	0.005	1.02(1–1.05)	0.0141	1.02(1.01–1.04)	0.013	1.02(1–1.04)	0.020	1.02(1–1.05)	0.018
pTNM										
pTNM I	1		1		1		1		1	
pTNM II	1.7(0.66–4.42)	0.274	2.32(0.89–6.02)	0.083	1.89(0.73–4.9)	0.188	1.75(0.67–4.54)	0.252	2.17(0.83–5.63)	0.113
pTNM III	5.24(2.07–13.28)	<0.001	6.71(2.65–16.99)	<0.001	6.16(2.44–15.56)	<0.001	5.39(2.13–13.67)	<0.001	5.65(2.23–14.32)	<0.001
Margins										
0 mm	1		1		1		1		1	
<1 mm	0.58(0.26–1.27)	0.174	0.55(0.25–1.21)	0.135	0.62(0.28–1.38)	0.238	0.7(0.32–1.54)	0.375	0.49(0.22–1.07)	0.075
1–2 mm	0.17(0.05–0.54)	0.003	0.16(0.05–0.58)	0.005	0.27(0.1–0.78)	0.015	0.24(0.08–0.75)	0.014	0.16(0.05–0.57)	0.005
2–10 mm	0.33(0.16–0.69)	0.003	0.45(0.22–0.92)	0.028	0.45(0.22–0.92)	0.028	0.43(0.21–0.89)	0.024	0.39(0.19–0.79)	0.009
10–50 mm	0.44(0.22–0.86)	0.017	0.64(0.33–1.24)	0.185	0.55(0.28–1.07)	0.080	0.58(0.3–1.14)	0.112	0.53(0.28–1.02)	0.056
>50 mm	0.35(0.14–0.89)	0.028	0.41(0.17–0.99)	0.049	0.41(0.17–1.01)	0.054	0.42(0.17–1.09)	0.074	0.35(0.15–0.86)	0.022
C-IRS1										
Low					1					
High					0.64(0.47–0.87)	0.005				
N-SMAD4										
Low										
High		NS						NS		
C-SMAD4										
Low	1						1			
High	0.58(0.43–0.8)	<0.001					0.67(0.5–0.91)	0.009		
N-RUNX3										
Low			1						1	
High			0.62(0.45–0.84)	0.002					0.68(0.44–1.03)	0.068
C-RUNX3										
Low									1	
High				NS					0.62(0.44–0.87)	0.006

Abbreviations: S, stroma; T, tumor; C, cytoplasm; N, nucleus; NS, not significant.

4. Discussion

To our knowledge, this is the first study to separately investigate tumor epithelial and stromal cell expression of IRS-1 and 2, SMAD4, and RUNX3 in colon cancer patients. We demonstrate that increased expression of RUNX3 and SMAD4 in tumor epithelial and stromal compartments and IRS-1 in stroma are independent predictors of a favorable prognosis in this patient group (Tables 1 and 2 and Figure 3). The largest difference in five-year DSS was observed for high vs. low RUNX3 expression in stromal cell cytoplasm (84% vs. 63%). Moreover, we have previously assessed the prognostic impact of miRs 126, 17-5p, and 20a-5p expression in tumor epithelial and stromal tissue from early stage colon cancer patients and observed that the investigated miRs were positive prognostic factors of DSS [2,3]. According to several cell line studies, the investigated markers are key players in pathways regulated by these miRs [7,38,39]. Surprisingly, with the exception of weak correlations ($0.2 < r < 0.25$) between miR-126 and SMAD4, no relevant correlations between the miRs and the investigated protein markers were observed. However, moderate to strong correlations were observed between the CD3+ and CD8+ TIL density and stromal cell RUNX3 expression.

Our results corroborate previous observations that RUNX3 is an important tumor repressor in colorectal cancer. In a large trial comprising 849 stage I–IV colorectal cancer patients, Soong et al. demonstrated that RUNX3 expression in the nucleus and not in the cytoplasm was a positive prognostic factor [39]. Shin et al. demonstrated that hypermethylation of RUNX3 was associated with an unfavorable prognosis in a small CRC cohort comprising 62 patients, indicating an inactivated form because of transcriptional silencing [40]. Ogino et al. found that patients with metastatic, microsatellite stable colorectal cancer with hypermethylated RUNX3 had an impaired prognosis when receiving combination chemotherapy compared to patients with unmethylated RUNX3 [41]. Furthermore, hypermethylation of RUNX3 is part of a panel of markers defining the CpG island methylator phenotype (CIMP) of colon cancer [42]. Berg et al. documented an OR of 3.4 for local recurrence for patients with loss of RUNX3 expression in a small series comprising 64 stage II–III colon cancer patients evaluated for MSI, CIMP, and copy number variation [43]. Soong et al. was the only other study showing differentiated expression of RUNX3 between tumor cell nuclei and cytoplasm [39]. In addition to prognostic significance, they demonstrated a reciprocal relationship between nuclear and cytoplasmic expression of RUNX3 and stage. Intriguingly, patients with neither nuclear nor cytoplasmic expression showed similar survival to patients with only nuclear expression. In our material we observed a trend towards impaired prognosis for patients with low nuclear and high cytoplasmic expression of RUNX3 in the tumor epithelial compartment. However, using our method and cut-offs, it is not possible to conclude that this staining represents exclusive cytoplasmic expression in tumor epithelial cells and not confounding immune cell expression. Kim et al. used a combination of preclinical models and retrospective data to investigate the connection between RUNX3 and headgehog (Hh) signaling in CRC [44]. Similar to us and others, they found that loss of RUNX3 indicated a poor prognosis. Furthermore, they observed that RUNX3 was a negative regulator of GLI1, the main activator of genes regulated by the Hh pathway. These results suggest that RUNX3 is strongly involved in regulating the Hh pathway in CRC [44]. Moreover, and of particular interest, we found that high stromal cell expression of RUNX3 was associated with favorable survival and was moderately to strongly correlated with CD3+ and CD8+ TILs density. We were not able to identify previous studies investigating compartment specific expression of RUNX3 in CRC. RUNX3 is an important mediator during lymphocyte differentiation into CD4+ and CD8+ cytotoxic T -lymphocytes (CTL) and natural killer cell progenitors into natural killer (NK) cells, and is thus likely pivotal in the development of anti-tumor immunity [24]. Furthermore, data from inflammatory bowel disease and preclinical models suggest that RUNX3 induced polarization of CD4+ cells towards a CD4+ CTL phenotype may be especially important in the gut epithelium [25,45,46]. Although our experimental data suggest a link between

lymphocyte expression of RUNX3 and prognosis, we were not able to assert the stromal RUNX3 signal to specific cell types due to limitations with our method. Nevertheless, our results strongly suggest RUNX3 in conjunction with TILs and/or other immune cells as drivers of the favorable survival we observe in the stromal RUNX3 high group.

SMAD4 is a known positive prognostic biomarker in CRC [11]. In our univariate analyses, patients with high expression of SMAD4 in either the tumor epithelial (HR 0.5 for both nucleus and cytoplasm in univariate analyses, and 0.62 for cytoplasm in multivariate analyses) or the stromal compartment (HR 0.5 and 0.67 for cytoplasm in univariate and multivariable analyses, respectively) exhibited beneficial prognosis vs. patients with low expression. These results are in line with those presented by Voorneveld et al. in a large meta-analysis of SMAD4 in CRC in 2015 [47]. Furthermore, supporting the robustness of our findings, they reported the percentage of preserved SMAD4 across the studies to be 50–90%, which corresponds favorably to our chosen cutoff, where approximately 75% of patients are in the group with high SMAD4 expression. Interestingly, in a recent study on patients resected for liver metastases, Kawaguchi et al. reported that the detrimental effect of RAS alterations was abrogated if both TP53 and SMAD4 was preserved [48]. Furthermore, the small group with concurrent RAS, TP53, and SMAD4 alterations exhibited the worst prognosis in both resectable and unresectable patients. These findings suggest that SMAD4 expression may be used to determine whether patients with limited liver metastases are likely to benefit from metastasectomy. We found SMAD4 expression to be a positive prognostic factor in the stromal compartment. Surprisingly, we were only able to identify one study that investigated the expression of SMAD4 in stroma. Contrary to our findings, Bacman et al. did not find any correlations between loss of SMAD4 in neither the tumor, nor the stromal compartment, in a study of 310 stage II–III CRC patients [49]. However, similar to our results, loss of SMAD4 in tumor and stroma was correlated with high grade tumors. Interestingly, Mesker et al. found a positive correlation between preserved SMAD4 and the total amount of tumor associated stroma [50]. Their results indicate that tumor SMAD4 signaling interacts with the tumor micro-environment. Furthermore, at least one preclinical study suggests that preserved SMAD4 signaling is necessary for T-cell suppression of CRC development [51]. Based on weak to moderate correlations with tumor epithelial and stromal RUNX3, IRS-1, CD8+ TIL density, and stromal miR126 (Figure 2), stromal SMAD4 expression in colon cancer is likely involved in complex interactions between tumor epithelial cells and different types of stromal cells. However, further studies are needed to draw conclusions on the role of stromal SMAD4 expression in colon cancer. SMAD4 and RUNX3 are both integral players in the TGF-β pathway [7,8]. In our material, SMAD4 and RUNX3 were weakly to moderately correlated both between and within the tumor epithelial and stromal compartments ($0.3 < r < 0.6$, Figure 2). Hence, it was pertinent to test their compartment specific co-expressions (Table S4, Figure S1). Not surprisingly, patients with a SMAD4-/RUNX3- pattern exhibited the worst prognosis in both compartments (5-year DSS 60%). Interestingly, preserved SMAD4 or RUNX3 resulted in similar 5-year DSSs in the tumor epithelial compartment (70%), but not in the stromal compartment (SMAD4-/RUNX3+ 80%, SMAD4+/RUNX3- 60%). The latter indicates a more important role for RUNX3 in the stromal compartment. These co-expression analyses, combined with our observation of RUNX3 being predominantly expressed in TILs, indicate that RUNX3 plays an important role in TIL-mediated colon cancer suppression.

Numerous studies in other cancer types have shown the oncogenic capacity of IRS-1 and 2, but little is known of their prognostic impact in colorectal cancer [16–19]. Interestingly, data from epidemiological studies suggest that specific genetic polymorphisms in the IRS-1 and 2 genes are associated with decreased or increased risk of developing colorectal cancer [52,53]. However, we did not observe any association between tumor expression of IRS-1 and 2 and survival in our cohort. In the stromal compartment, IRS-1 expression was an independent positive prognosticator. Stromal IRS-1 was moderately ($0.4 < r < 0.6$) correlated with tumor epithelial and stromal SMAD4, and tumor epithelial and stromal

RUNX3, indicating a possible link between stromal IRS-1 and pathways involving activated SMAD4 and RUNX3.

5. Future Works

As a notion for future works, we would like to outline the following. (1): S-RUNX3 was predominantly expressed in lymphocyte-like cells. However, morphological classification is prone to error, especially when it is largely based on the hematoxylin staining alone. Further studies should explore S-RUNX3 expression using multiplex immunohistochemistry or fluorescent labeling differentiating specific cell types. We would start with the pan T-lymphocyte marker CD3 and expand the panel of markers if this initial co-expression does not fully explain our current findings. Furthermore, as different types of immune cell activation must be considered as a pan-cancer event, future studies may also include other cancer types. (2): The interactions between T-SMAD4/T-RUNX3 and S-SMAD4/S-RUNX3 warrants further scrutiny in a clinical setting. Even though pre-clinical studies suggest that their silencing results in a deregulation of the TGF-β pathway, clinical confirmation is needed, as other pathways may be involved in this complex interplay.

6. Conclusions

We present the first study that differentiates between expression of IRS-1 and 2, RUNX3 and SMAD4 in the tumor epithelial and stromal compartments of colon cancer patients. We confirm previous studies reporting preserved RUNX3 and SMAD4 in the tumor epithelial compartment as positive prognosticators in colon cancer. Furthermore, we present novel data on the positive prognostic value of stromal RUNX3, SMAD4, and IRS-1 in colon cancer. In addition, we show that stromal RUNX3 is correlated with TIL density in colon cancer. Not surprisingly, co-expression analyses indicate that RUNX3 plays an important role in TIL-mediated colon cancer suppression. Clinical implications of loss of RUNX3 and/or SMAD4 expression in the tumor epithelial compartment may be used to identify stage II patients in need of adjuvant chemotherapy, or patients with T1-3N1 tumors eligible for extended chemotherapy compared to the three month standard for this risk group. However, a comprehensive prospective validation is warranted before these potential biomarkers are implemented in a clinical setting.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers15051448/s1, Table S1. HC Procedure Discovery Ultra; Table S2. Product information of antibodies and reagents; Table S3. Dichotomized IRS-1, IRS-2, SMAD4 and RUNX3 in tumor and stromal cell nucleus and/or cytoplasm and their distribution over and correlation with clinicopathological variables (chi-square and Fisher's exact tests); Table S4. (A) Univariate analyses of co-expression analyses between SMAD4 in tumor cytoplasm and RUNX3 in tumor nucleus (A1) and between SMAD4 in stromal cytoplasm and RUNX3 in and stromal cytoplasm (A2, log-rank test test, n = 452). (B) Multivariable models including co-expressions of SMAD4 in tumor cytoplasm and RUNX3 in tumor nucleus (B1) and SMAD4 in stromal cytoplasm and RUNX3 in stromal cytoplasm (B2) and relevant clinicopathological variables (cox proportional hazards test, n = 452); Figure S1. Disease-specific survival curves for the co-expressions between (A) SMAD4 in tumor cytoplasm and RUNX3 in tumor nucleus and (B) SMAD4 in stromal cytoplasm and RUNX3 in stromal cytoplasm.

Author Contributions: All authors contributed to study design. H.S. collected the clinical data. L.-T.R.B. collected and revised the pathological specimens. M.I.P. and A.P.G.L. conducted the IHC experiments. H.S., T.K.K. and S.A. analyzed and interpreted the data. H.S. and T.K.K. wrote the manuscript. All authors participated during revisions and approved the final manuscript.

Funding: The works was funded by UiT The Arctic University of Norway. North Norwegian Health Authority grant no. HNF1521-20. The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki. It was approved by the Regional Committee for Medical and Health Research Ethics North (REK Nord, protocol ID: 2011/2151) and Institutional Review Boards of the participating institutions (PVO project number: 0312). The reporting of clinicopathological variables, survival data, and biomarker expression was conducted in accordance with the REMARK guidelines 51.

Informed Consent Statement: Patient consent was waived due to the retrospective nature of the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy regulations.

Acknowledgments: We would like to thank Magnus L. Persson for preparing the TMA-slides.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics, 2022. *CA Cancer J. Clin.* **2022**, *72*, 7–33. [CrossRef] [PubMed]
2. Selven, H.; Busund, L.T.R.; Andersen, S.; Bremnes, R.M.; Kilvær, T.K. High expression of microRNA-126 relates to favorable prognosis for colon cancer patients. *Sci. Rep.* **2021**, *11*, 9592. [CrossRef] [PubMed]
3. Selven, H.; Andersen, S.; Pedersen, M.I.; Lombardi, A.P.G.; Busund, L.T.R.; Kilvær, T.K. High expression of miR-17-5p and miR-20a-5p predicts favorable disease-specific survival in stage I-III colon cancer. *Sci. Rep.* **2022**, *12*, 7080. [CrossRef]
4. Shaw, L.M. The insulin receptor substrate (IRS) proteins. *Cell Cycle* **2011**, *10*, 1750–1756. [CrossRef] [PubMed]
5. Levanon, D.; Groner, Y. Structure and regulated expression of mammalian RUNX genes. *Oncogene* **2004**, *23*, 4211–4219. [CrossRef]
6. McCarthy, A.J.; Chetty, R. Smad4/DPC4. *J. Clin. Pathol.* **2018**, *71*, 661–664. [CrossRef] [PubMed]
7. Zhou, Y.; Feng, X.; Liu, Y.l.; Ye, S.c.; Wang, H.; Tan, W.k.; Tian, T.; Qiu, Y.m.; Luo, H.s. Down-Regulation of miR-126 Is Associated with Colorectal Cancer Cells Proliferation, Migration and Invasion by Targeting IRS-1 via the AKT and ERK1/2 Signaling Pathways. *PLoS ONE* **2013**, *8*, e81203. [CrossRef]
8. Ito, Y.; Miyazono, K. RUNX transcription factors as key targets of TGF-β superfamily signaling. *Curr. Opin. Genet. Dev.* **2003**, *13*, 43–47. [CrossRef]
9. Zhao, M.; Mishra, L.; Deng, C.X. The role of TGF-β/SMAD4 signaling in cancer. *Int. J. Biol. Sci.* **2018**, *14*, 111–123. [CrossRef] [PubMed]
10. Hennige, A.; Lammers, R.; Arlt, D.; Höppner, W.; Strack, V.; Niederfellner, G.; Seif, F.; Häring, H.U.; Kellerer, M. Ret oncogene signal transduction via a IRS-2/PI 3-kinase/PKB and a SHC/Grb-2 dependent pathway: Possible implication for transforming activity in NIH3T3 cells. *Mol. Cell. Endocrinol.* **2000**, *167*, 69–76. [CrossRef]
11. Isaksson-Mettävainio, M.; Palmqvist, R.; Forssell, J.; Stenling, R.; Oberg, A. SMAD4/DPC4 Expression and Prognosis in Human Colorectal Cancer. *Anticancer Res.* **2006**, *26*, 507–510. [PubMed]
12. Mu, W.P.; Wang, J.; Niu, Q.; Shi, N.; Lian, H.F. Clinical significance and association of RUNX3 hypermethylation frequency with colorectal cancer: A meta-analysis. *OncoTargets Ther.* **2014**, *7*, 1237. [CrossRef]
13. Wang, H.; Wang, J.; Li, D.; Zhu, Z.; Pei, D. A functional polymorphism within the distal promoter of *RUNX3* confers risk of colorectal cancer. *Int. J. Biol. Markers* **2022**, *37*, 40–46. [CrossRef]
14. Sesti, G.; Federici, M.; Hribal, M.L.; Lauro, D.; Sbraccia, P.; Lauro, R. Defects of the insulin receptor substrate (IRS) system in human metabolic disorders. *FASEB J.* **2001**, *15*, 2099–2111. [CrossRef] [PubMed]
15. Dearth, R.K.; Cui, X.; Kim, H.J.; Hadsell, D.L.; Lee, A.V. Oncogenic Transformation by the Signaling Adaptor Proteins Insulin Receptor Substrate (IRS)-1 and IRS-2. *Cell Cycle* **2007**, *6*, 705–713. [CrossRef]
16. Nehrbass, D.; Klimek, F.; Bannasch, P. Overexpression of insulin receptor substrate-1 emerges early in hepatocarcinogenesis and elicits preneoplastic hepatic glycogenosis. *Am. J. Pathol.* **1998**, *152*, 341–345.
17. Boissan, M.; Beurel, E.; Wendum, D.; Rey, C.; Lécluse, Y.; Housset, C.; Lacombe, M.L.; Desbois-Mouthon, C. Overexpression of Insulin Receptor Substrate-2 in Human and Murine Hepatocellular Carcinoma. *Am. J. Pathol.* **2005**, *167*, 869–877. [CrossRef] [PubMed]
18. Bergmann, U.; Funatomi, H.; Kornmann, M.; Beger, H.G.; Korc, M. Increased Expression of Insulin Receptor Substrate-1 in Human Pancreatic Cancer. *Biochem. Biophys. Res. Commun.* **1996**, *220*, 886–890. [CrossRef]
19. Kornmann, M.; Maruyama, H.; Bergmann, U.; Tangvoranuntakul, P.; Beger, H.G.; White, M.F.; Korc, M. Enhanced expression of the insulin receptor substrate-2 docking protein in human pancreatic cancer. *Cancer Res.* **1998**, *58*, 4250–4254.
20. Chang, Q.; Li, Y.; White, M.F.; Fletcher, J.A.; Xiao, S. Constitutive activation of insulin receptor substrate 1 is a frequent event in human tumors: Therapeutic implications. *Cancer Res.* **2002**, *62*, 6035–6038.
21. Ito, Y. RUNX genes in development and cancer: Regulation of viral gene expression and the discovery of RUNX family genes. *Adv. Cancer Res.* **2008**, *99*, 33–76.
22. Ragnarsson, G.; Eiriksdottir, G.; Johannsdottir, J.T.; Jonasson, J.G.; Egilsson, V.; Ingvarsson, S. Loss of heterozygosity at chromosome 1p in different solid human tumours: Association with survival. *Brit. J. Cancer* **1999**, *79*, 1468–1474. [CrossRef]
23. Soong, R.; Shah, N.; Peh, B.K.; Chong, P.Y.; Ng, S.S.; Zeps, N.; Joseph, D.; Salto-Tellez, M.; Iacopetta, B.; Ito, Y. The expression of RUNX3 in colorectal cancer is associated with disease stage and patient outcome. *Brit. J. Cancer* **2009**, *100*, 676–679. [CrossRef]

24. Seo, W.; Nomura, A.; Taniuchi, I. The Roles of RUNX Proteins in Lymphocyte Function and Anti-Tumor Immunity. *Cells* **2022**, *11*, 3116. [CrossRef] [PubMed]
25. Reis, B.S.; Rogoz, A.; Costa-Pinto, F.A.; Taniuchi, I.; Mucida, D. Mutual expression of the transcription factors Runx3 and ThPOK regulates intestinal CD4+ T cell immunity. *Nat. Immunol.* **2013**, *14*, 271–280. [CrossRef] [PubMed]
26. Shi, Y.; Massagué, J. Mechanisms of TGF-β Signaling from Cell Membrane to the Nucleus. *Cell* **2003**, *113*, 685–700. [CrossRef]
27. Javelaud, D.; Mauviel, A. Crosstalk mechanisms between the mitogen-activated protein kinase pathways and Smad signaling downstream of TGF-β: Implications for carcinogenesis. *Oncogene* **2005**, *24*, 5742–5750. [CrossRef] [PubMed]
28. Lee, M.K.; Pardoux, C.; Hall, M.C.; Lee, P.S.; Warburton, D.; Qing, J.; Smith, S.M.; Derynck, R. TGF-β activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *EMBO J.* **2007**, *26*, 3957–3967. [CrossRef]
29. Bakin, A.V.; Tomlinson, A.K.; Bhowmick, N.A.; Moses, H.L.; Arteaga, C.L. Phosphatidylinositol 3-Kinase Function Is Required for Transforming Growth Factor β-mediated Epithelial to Mesenchymal Transition and Cell Migration. *J. Biol. Chem.* **2000**, *275*, 36803–36810. [CrossRef]
30. Hahn, S.A.; Schutte, M.; Shamsul Hoque, A.T.M.; Moskaluk, C.A.; da Costa, L.T.; Rozenblum, E.; Weinstein, C.L.; Fischer, A.; Yeo, C.J.; Hruban, R.H.; et al. *DPC4*, A Candidate Tumor Suppressor Gene at Human Chromosome 18q21.1. *Science* **1996**, *271*, 350–353. [CrossRef]
31. Xu, X.; Kobayashi, S.; Qiao, W.; Li, C.; Xiao, C.; Radaeva, S.; Stiles, B.; Wang, R.H.; Ohara, N.; Yoshino, T.; et al. Induction of intrahepatic cholangiocellular carcinoma by liver-specific disruption of Smad4 and Pten in mice. *J. Clin. Investig.* **2006**, *116*, 1843–1852. [CrossRef] [PubMed]
32. Miyaki, M.; Kuroki, T. Role of Smad4 (DPC4) inactivation in human cancer. *Biochem. Biophys. Res. Commun.* **2003**, *306*, 799–804. [CrossRef]
33. Bremnes, R.M.; Veve, R.; Gabrielson, E.; Hirsch, F.R.; Baron, A.; Bemis, L.; Gemmill, R.M.; Drabkin, H.A.; Franklin, W.A. High-Throughput Tissue Microarray Analysis Used to Evaluate Biology and Prognostic Significance of the E-Cadherin Pathway in Non–Small-Cell Lung Cancer. *J. Clin. Oncol.* **2002**, *20*, 2417–2428. [CrossRef]
34. Bankhead, P.; Loughrey, M.B.; Fernández, J.A.; Dombrowski, Y.; McArt, D.G.; Dunne, P.D.; McQuaid, S.; Gray, R.T.; Murray, L.J.; Coleman, H.G.; et al. QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* **2017**, *7*, 16878. [CrossRef] [PubMed]
35. Graham, S.; Jahanifar, M.; Azam, A.; Nimir, M.; Tsang, Y.W.; Dodd, K.; Hero, E.; Sahota, H.; Tank, A.; Benes, K.; et al. Lizard: A Large-Scale Dataset for Colonic Nuclear Instance Segmentation and Classification. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 11–17 October 2021; pp. 684–693. [CrossRef]
36. Weigert, M.; Schmidt, U. Nuclei Instance Segmentation and Classification in Histopathology Images with Stardist. In Proceedings of the 2022 IEEE International Symposium on Biomedical Imaging Challenges (ISBIC), Kolkata, India, 28–31 March 2022; pp. 1–4. [CrossRef]
37. Schmidt, U.; Weigert, M.; Broaddus, C.; Myers, G. Cell Detection with Star-Convex Polygons. In Proceedings of the Medical Image Computing and Computer Assisted Intervention — MICCAI 2018, Granada, Spain, 16–20 September 2018; Frangi, A.F., Schnabel, J.A., Davatzikos, C., Alberola-López, C., Fichtinger, G., Eds.; Springer: Cham, Switzerland, 2018; pp. 265–273.
38. Cheng, D.; Zhao, S.; Tang, H.; Zhang, D.; Sun, H.; Yu, F.; Jiang, W.; Yue, B.; Wang, J.; Zhang, M.; et al. MicroRNA-20a-5p promotes colorectal cancer invasion and metastasis by downregulating Smad4. *Oncotarget* **2016**, *7*, 45199–45213. [CrossRef] [PubMed]
39. Song, J.; Liu, Y.; Wang, T.; Li, B.; Zhang, S. MiR-17-5p promotes cellular proliferation and invasiveness by targeting RUNX3 in gastric cancer. *Biomed. Pharmacother.* **2020**, *128*, 110246. [CrossRef]
40. Shin, E.J.; Kim, H.J.; Son, M.W.; Ahn, T.S.; Lee, H.Y.; Lim, D.R.; Bae, S.B.; Jeon, S.; Kim, H.; Jeong, D.; et al. Epigenetic inactivation of RUNX3 in colorectal cancer. *Ann. Surg. Treat. Res.* **2018**, *94*, 19–25. [CrossRef] [PubMed]
41. Ogino, S.; Meyerhardt, J.A.; Kawasaki, T.; Clark, J.W.; Ryan, D.P.; Kulke, M.H.; Enzinger, P.C.; Wolpin, B.M.; Loda, M.; Fuchs, C.S. CpG island methylation, response to combination chemotherapy, and patient survival in advanced microsatellite stable colorectal carcinoma. *Virchows Arch.* **2007**, *450*, 529–537. [CrossRef]
42. Weisenberger, D.J.; Siegmund, K.D.; Campan, M.; Young, J.; Long, T.I.; Faasse, M.A.; Kang, G.H.; Widschwendter, M.; Weener, D.; Buchanan, D.; et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat. Genet.* **2006**, *38*, 787–793. [CrossRef] [PubMed]
43. Berg, M.; Nordgaard, O.; Kørner, H.; Oltedal, S.; Smaaland, R.; Søreide, J.A.; Søreide, K. Molecular Subtypes in Stage II-III Colon Cancer Defined by Genomic Instability: Early Recurrence-Risk Associated with a High Copy-Number Variation and Loss of RUNX3 and CDKN2A. *PLoS ONE* **2015**, *10*, e0122391. [CrossRef]
44. Kim, B.R.; Na, Y.J.; Kim, J.L.; Jeong, Y.A.; Park, S.H.; Jo, M.J.; Jeong, S.; Kang, S.; Oh, S.C.; Lee, D.H. RUNX3 suppresses metastasis and stemness by inhibiting Hedgehog signaling in colorectal cancer. *Cell Death Differ.* **2019**, *27*, 676–694. [CrossRef]
45. Sugai, M.; Aoki, K.; Osato, M.; Nambu, Y.; Ito, K.; Taketo, M.M.; Shimizu, A. Runx3 Is Required for Full Activation of Regulatory T Cells To Prevent Colitis-Associated Tumor Formation. *J. Immunol.* **2011**, *186*, 6515–6520. [CrossRef] [PubMed]
46. Garrity-Park, M.M.; Loftus, E.V.; Bryant, S.C.; Smyrk, T.C. A Biomarker Panel to Detect Synchronous Neoplasm in Non-neoplastic Surveillance Biopsies from Patients with Ulcerative Colitis. *Inflamm. Bowel Dis.* **2016**, *22*, 1568–1574. [CrossRef]
47. Voorneveld, P.W.; Jacobs, R.J.; Kodach, L.L.; Hardwick, J.C. A Meta-Analysis of SMAD4 Immunohistochemistry as a Prognostic Marker in Colorectal Cancer. *Transl. Oncol.* **2015**, *8*, 18–24. [CrossRef]

48. Kawaguchi, Y.; Kopetz, S.; Panettieri, E.; Hwang, H.; Wang, X.; Cao, H.S.T.; Tzeng, C.W.D.; Chun, Y.S.; Aloia, T.A.; Vauthey, J.N. Improved Survival over Time After Resection of Colorectal Liver Metastases and Clinical Impact of Multigene Alteration Testing in Patients with Metastatic Colorectal Cancer. *J. Gastrointest. Surg.* **2021**, *26*, 583–593. [CrossRef]
49. Bacman, D.; Merkel, S.; Croner, R.; Papadopoulos, T.; Brueckl, W.; Dimmler, A. TGF-beta receptor 2 downregulation in tumour-associated stroma worsens prognosis and high-grade tumours show more tumour-associated macrophages and lower TGF-beta1 expression in colon carcinoma: A retrospective study. *BMC Cancer* **2007**, *7*, 156. [CrossRef]
50. Mesker, W.E.; Liefers, G.J.; Junggeburt, J.M.C.; van Pelt, G.W.; Alberici, P.; Kuppen, P.J.K.; Miranda, N.F.; van Leeuwen, K.A.M.; Morreau, H.; Szuhai, K.; et al. Presence of a High Amount of Stroma and Downregulation of SMAD4 Predict for Worse Survival for Stage I–II Colon Cancer Patients. *Anal. Cell. Pathol.* **2009**, *31*, 169–178. [CrossRef]
51. Kim, B.G.; Li, C.; Qiao, W.; Mamura, M.; Kasperczak, B.; Anver, M.; Wolfraim, L.; Hong, S.; Mushinski, E.; Potter, M.; et al. Smad4 signalling in T cells is required for suppression of gastrointestinal cancer. *Nature* **2006**, *441*, 1015–1019. [CrossRef] [PubMed]
52. Slattery, M.L.; Samowitz, W.; Curtin, K.; Ma, K.N.; Hoffman, M.; Caan, B.; Neuhausen, S. Associations among IRS1, IRS2, IGF1, and IGFBP3 genetic polymorphisms and colorectal cancer. *Cancer Epidemiol. Biomarkers Prev.* **2004**, *13*, 1206–1214. [CrossRef]
53. Sauerbrei, W.; Taube, S.E.; McShane, L.M.; Cavenagh, M.M.; Altman, D.G. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): An Abridged Explanation and Elaboration. *JNCI J. Natl. Cancer Inst.* **2018**, *110*, 803–811. [CrossRef] [PubMed]