



Genome-wide analysis of mRNA and microRNA expression in colorectal cancer and adjacent normal mucosa

Yuma Ito¹, Mitsumasa Osakabe^{1†}, Takeshi Niinuma², Noriyuki Uesugi¹, Ryo Sugimoto¹, Naoki Yanagawa¹, Koki Otsuka³, Akira Sasaki³, Takayuki Matsumoto⁴, Hiromu Suzuki²  and Tamotsu Sugai^{1†*} 

¹Department of Molecular Diagnostic Pathology, School of Medicine, Iwate Medical University, Yahaba, Japan

²Department of Molecular Biology, School of Medicine, Sapporo Medical University, Sapporo, Japan

³Department of Surgery, School of Medicine, Iwate Medical University, Yahaba, Japan

⁴Division of Gastroenterology, Department of Internal Medicine, Iwate Medical University, Yahaba, Japan

*Correspondence to: Tamotsu Sugai, Department of Molecular Diagnostic Pathology, School of Medicine, Iwate Medical University, 2-1-1, Shiwagun'yahabachou, Yahaba 028-3695, Japan. E-mail: tsugai@iwate-med.ac.jp

†These authors contributed equally to this study.

Abstract

mRNA expression varies in human cancers. Such altered mRNA expression is negatively regulated by the expression of microRNAs (miRNAs), which play an important role in human tumorigenesis. According to this theory, inverse mRNA/miRNA expression may be a direct driver of cancer development, and certain genetic events may occur prior to the development of any discernible histological abnormalities. We examined the inverse expression between mRNAs and their corresponding miRNAs in colorectal cancer (CRC) and adjacent normal mucosa and performed pathway analysis to identify mRNA/miRNA networks. The cancer samples were divided into first (20 cases) and second (24 cases) cohorts, and 48 samples were obtained from two sections of the normal mucosa adjacent to the tumors from the second cohort. We investigated mRNAs with commonly altered expression in CRC and adjacent normal mucosa using isolated cancer glands and normal crypts from the first cohort, compared with that of distal normal crypts, using an array-based method. As a result, significant inverse correlations between *CEACAM1* and miRNA-7114-5p and between *AK1* and miRNA-6780-5p were found in CRC and adjacent normal mucosa. We validated these correlations in the second cohort using RT-PCR. To confirm these findings, transfection and immunohistochemical assays were also performed, which verified the inverse correlation between *CEACAM1* and miRNA-7114-5p. Our findings suggest that the inverse correlations between the *CEACAM1*/miRNA-7114-5p and possibly *AK1*/miRNA-6780-5p pairs play an important role in early CRC development, and may help identify potential molecular targets for early detection of CRC.

Keywords: adjacent normal mucosa; array-based analysis; colorectal cancer; microRNA; mRNA

Received 22 September 2021; Revised 8 February 2022; Accepted 21 February 2022

No conflicts of interest were declared.

Introduction

Colorectal cancer (CRC) is one of the most common cancers, affecting 1.36 million people annually and the leading cause of cancer-related death worldwide and in developed countries [1,2]. Early detection of, and treatments for, CRC help decrease CRC-related death. Despite improvements in diagnosis, most CRC cases are not discovered until an advanced stage [1,2]. Identification of the molecular mechanisms of CRC is indispensable for resolving this problem. Cancers develop via accumulation of molecular events

including mutations in key genes (e.g. *APC*, *KRAS*, *DPC4*, and *TP53*), abnormal DNA methylation, and dysregulated microRNA (miRNA) expression [3–5]. Traditional theory proposes a three-step process of initiation, promotion, and progression [6], and multiple molecular alterations accumulate during this process [6,7]. A recent study showed that epigenetic changes in CRC occur long before the development of a clinically visible lesion harboring multiple genetic and epigenetic alterations, including DNA methylation and altered miRNA expression [8,9]. Many cancer cells acquire protumorigenic molecular alterations that do not induce

morphological changes but predispose to subsequent malignant transformation [10–12]. Such cells can form mucosa with increased risk of cancer development and can expand into surrounding tissues [10–13]. This process, described as ‘field cancerization’, can explain the molecular alterations seen in pre-neoplastic tissues, which appear as non-neoplastic mucosa histologically [10–13]. However, previous studies investigating colorectal carcinogenesis focused solely on cancer tissues or assumed that the mucosa adjacent to the neoplastic lesion is normal histologically and molecularly [10,13]. A recent study showed that normal mucosa surrounding CRC tumors with hallmark molecular events is considered ‘mucosa at risk’, which might be the earliest step in CRC formation [10–13].

According to this theory, changes in mRNA expression, a potential driver of tumorigenesis, occur in the surrounding mucosa before neoplastic formation. Additionally, such changes are regulated by miRNAs, a family of small non-coding RNAs that regulate biological processes including carcinogenesis [14,15]. miRNAs have been investigated in the field of oncological research, and evidence suggests that altered miRNA regulation may be involved in cancer pathogenesis by regulating the transcription of tumor suppressors and oncogenes [14,15]. Aberrant miRNA expression can also facilitate or disturb many cell signaling pathways [14,15]. Dysregulated expression of an mRNA and its corresponding miRNA in pre-neoplastic cells predisposes to cancer onset and progression to CRC, possibly affecting the expression patterns of transcripts in tumor cells [16,17]. Therefore, identification of paired mRNAs/miRNAs is essential for evaluating human tumorigenesis.

Here, we investigated dysregulated mRNA/miRNA pairs in isolated cancer and adjacent normal glands compared with expression levels in the distal mucosa using a crypt isolation method, which enables exclusion of interstitial cells. We also examined whether such dysregulated pairs are associated with cancer development risk in the mucosa.

Materials and methods

Patients

The workflow of this study is shown in Figure 1. Twenty cases of CRC were examined as the first cohort. The CRC stage was determined according to the guidelines of the Japanese Society for Cancer of the Colon and Rectum. The clinicopathological findings are shown in Table 1. The histological diagnosis was based on the World Health Organization

classification described previously [4]. Pathological factors were recorded in accordance with the Japanese Society for Cancer of the Colon and Rectum [18]. In addition, 24 CRCs were analyzed as the second (validation) cohort to confirm the validity of the results obtained from the first cohort (Table 1).

All patients and controls provided informed consent, and the study protocol was approved by the Ethics Committee of Iwate Medical University (reference number: MH2019-027).

Crypt isolation method

A crypt isolation method was used to obtain pure tumor and non-neoplastic glands from CRC patients. CRC tumor samples were obtained from the central area of the tumor [9]. Normal colonic mucosa was taken from the most distal portion of the colon and normal mucosa adjacent to the tumor (within 5 mm). The most distal mucosa was at least 5 cm away from the tumor. For the second cohort samples, we obtained isolated cancer glands and normal crypts from two different parts (distal and proximal) of normal mucosa located within 5 mm (median 3 mm) of the tumor. Crypt isolation was performed according to a previously reported method [9]. In brief, fresh mucosa and tumor tissue were cut with a sharp scalpel into minute pieces and incubated at 37 °C for 30 min in calcium- and magnesium-free Hanks’ balanced salt solution containing 30 mmol/l EDTA. The tissue was then stirred in the same calcium- and magnesium-free Hanks’ balanced salt solution for 30–40 min to accelerate isolation. Any remaining cancer tissue after this step was separated from the cancer stroma manually under a dissecting microscope (SZ60, Olympus, Tokyo, Japan). Using this process, all cancer tissue was clearly isolated from cancer stroma. Finally, although dis-cohesive cells should not be included in the isolated cancer samples, solid components might be. However, no solid components were evident in isolated samples.

Isolated glands were immediately fixed in 70% ethanol and stored at 4 °C until used for DNA extraction. The fixed isolated glands were observed under a dissection microscope. Isolated glands were processed routinely to confirm their histology using paraffin-embedded histological sections. Contamination, such as interstitial cells, was not observed in the samples. Morphological differences in isolated normal crypts could not distinguish adjacent normal crypts from distal normal crypts. In addition, differences in crypt morphology could not be identified among adjacent normal crypts. A representative example of the collection method used to obtain isolated tumor glands and normal isolated crypts is shown in Figure 2.

Workflow of this study

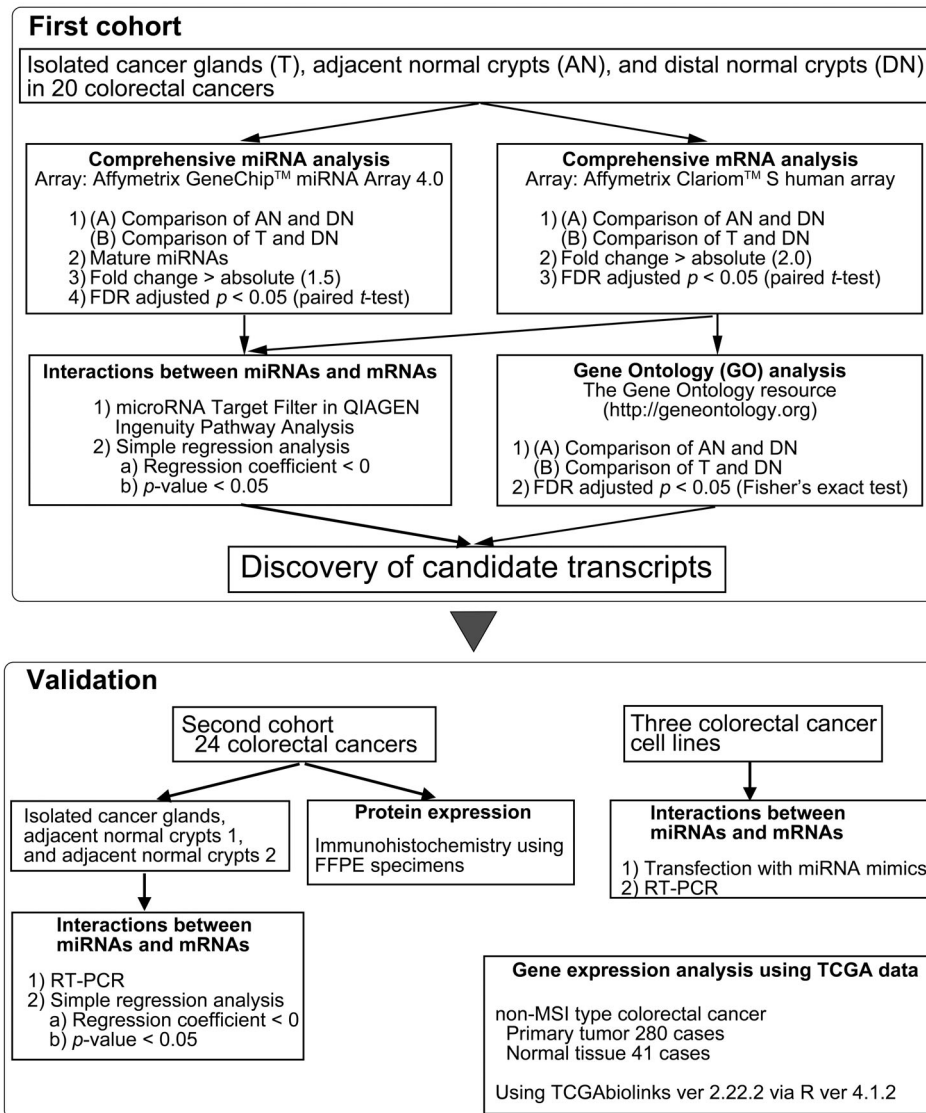


Figure 1. Workflow of this study. First, we examined the differences in mature miRNA expression levels between cancer or adjacent normal glands and distal normal glands using the following criteria: fold-change expression >1.5 and FDR-adjusted $p < 0.05$ for the GeneChip miRNA 4.0 microarray, and fold-change expression >2.0 and FDR-adjusted $p < 0.05$ for the Clariom S human array. Second, we assessed the interactions between miRNAs and mRNAs using the miRNA Target Filter in IPA to identify potential candidate mRNA/miRNA pairs. In addition, simple regression analysis was performed. Third, GO analysis of the mRNAs with altered expression identified from the arrays was used to determine their biological roles, and candidate mRNAs commonly found between adjacent normal crypts and CRC glands were subjected to GO analysis. In addition, miRNAs inversely correlated with mRNA expression were identified by IPA and simple regression analysis. Validation analyses including RT-PCR, immunohistochemistry, and transfection assays, were performed in cohort 2. Finally, the expression of candidate mRNAs and miRNAs in TCGA database was evaluated.

DNA extraction

DNA was extracted by standard SDS–proteinase K treatment and resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8.0]).

RNA isolation

RNAs were extracted using the mirVana™ miRNA Isolation kit (Thermo Fisher Scientific, Inc., Waltham, MA, USA) according to the manufacturer's

Table 1. Clinicopathological findings for the CRC cases

	Cohort 1 (%)	Cohort 2 (%)	P value
Total	20	24	
Sex			0.7600
Male	11 (55)	15 (62.5)	
Female	9 (45)	9 (37.5)	
Age, years, median (range)	72 (31–94)	72 (31–96)	0.8436
Location			0.9537
C/A/T/D/S/R	2/4/2/1/7/4	3/7/1/1/8/4	
Size, mm, median (range)	44.5 (21–115)	43 (21–115)	0.6266
Histological type			1.0000
Well differentiated	0 (0)	1 (4.2)	
Moderately differentiated	18 (90)	21 (87.5)	
Poorly differentiated	1 (5)	1 (4.2)	
Papillary type	1 (5)	1 (4.2)	
Stage			0.9150
I	1 (5)	0 (0)	
II	9 (45)	11 (45.8)	
III	7 (35)	9 (37.5)	
IV	3 (15)	4 (16.7)	

A, ascending colon; C, cecum; D, descending colon; R, rectum; S, sigmoid colon; T, transverse colon.

instructions. RNA quantity and quality were evaluated using the DU730 spectrophotometer (Beckman Coulter, Brea, CA, USA), and RNA integrity was determined by gel electrophoresis.

Clariom S human array and gene expression analysis

For each array experiment, 500 ng total RNA was used for labeling before hybridizing to the Clariom S human array (Thermo Fisher Scientific). Probe labeling, chip hybridization, and scanning were performed according to the manufacturer's instructions. Array data were generated using GeneSpring (version 14.9.1, Agilent Technologies, Santa Clara, CA, USA). Detailed information and methods were described previously [19]. mRNA expression was analyzed according to the following criteria: $p < 0.05$ with adjusted Benjamini–Hochberg/false discovery rate (FDR) correction and fold change expression >2.0 compared with expression in normal glands using the paired t -test.

miRNA microarray analysis

For microarray analysis, 200 ng RNA was polyadenylated and labeled using the FlashTag™ Biotin HSR RNA Labeling kit (Thermo Fisher Scientific, Aaltham, MA, USA) and then treated with DNA ligase. The labeled RNA was hybridized to GeneChip miRNA 4.0 microarrays (Thermo Fisher Scientific) at 48 °C for 16 h, followed by washing and staining with

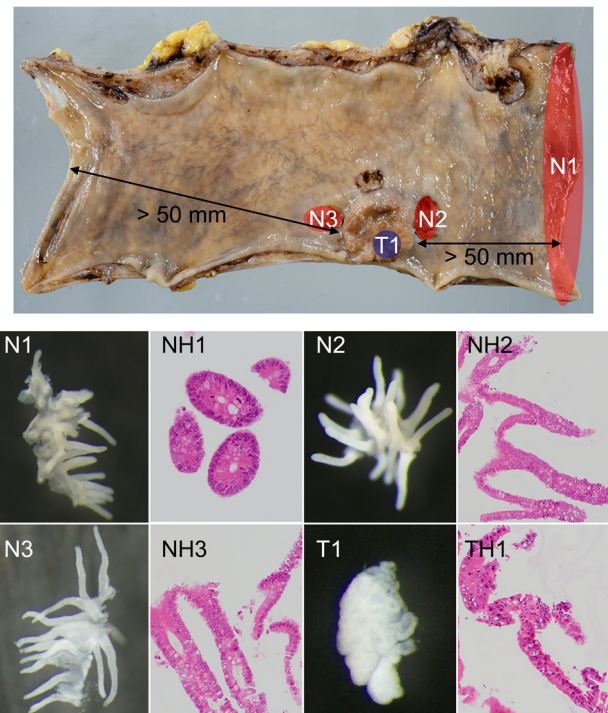


Figure 2. Representative sampling of an isolated CRC gland and normal mucosa. N1, distal normal mucosa; NH1, normal histological feature; N2, adjacent normal mucosa at distal site; NH2, normal histological feature; N3, adjacent normal mucosa at proximal site; NH3, normal histological feature; T1, isolated cancer gland; TH1, isolated cancer histology.

a streptavidin–PE solution. Stained arrays were scanned using the GeneChip™ Scanner 3000 7G System (Thermo Fisher Scientific). Array data were generated using GeneSpring (version 14.9.1, Agilent Technologies). miRNA expression was examined according to the following criteria: $p < 0.05$ with adjusted Benjamini–Hochberg/FDR correction and fold change in expression >1.5 compared with expression in normal glands using the paired t -test.

Functional and pathway analysis

Gene Ontology (GO) analysis is commonly used for functional analyses of large-scale transcriptomic and genomic data. The GO resource (<http://geneontology.org>) is a reference database used for systematic analysis of gene functions and of the biological significance of large lists of genes or transcripts.

Pathway crosstalk analysis

Enrichment map is a network-based method for gene-set enrichment visualization and interpretation. To

extract the interactions among significantly enriched signaling pathways, pathway crosstalk analysis was performed using the miRNA Target Filter in Ingenuity Pathway Analysis (IPA) (Qiagen, Valencia, CA, USA).

Simple linear regression analysis

Candidate mRNAs and their regulatory miRNAs, together with the corresponding expression profiles, were analyzed further. The detailed method is described in Supplementary materials and methods.

Immunohistochemistry, assessment of consensus molecular subtypes (CMSs), quantitative RT-PCR for validation analyses, and transfection with miRNA mimics are described in Supplementary materials and methods.

Statistical analysis

The correlations between miRNAs and mRNAs were examined by single regression analysis (regression coefficient < 0 and $p < 0.05$). We compared the CMS frequencies in CRC between the two cohorts using Fisher's exact test and compared immunohistochemical expression of candidate markers among distal and adjacent normal mucosae and CRC cancer tissue using the Friedman test. If significant differences ($p < 0.05$ with Bonferroni correction) were found among the three tissue groups, differences between two groups were analyzed using the Wilcoxon matched-pairs signed rank test. We used JMP Pro 13.0 software for Windows (SAS Institute Inc., Cary, NC, USA) for analyses, with a P value of < 0.05 indicating significance.

Results

Molecular characteristics of CRCs in cohorts 1 and 2

All isolated cancer glands examined were of the microsatellite stable (MSS) phenotype according to microsatellite analysis. We classified CRCs of cohorts 1 and 2 into CMSs by immunohistochemical analyses (<https://crcclassifier.shinyapps.io/appTesting/>) [20,21]. In cohort 1, 6 tumors were classified as CMS 2 and 14 as CMS 3; in cohort 2, 6 CRCs were classified as CMS 2 and 18 as CMS 3. There were no significant differences in CMS frequencies between cohorts 1 and 2 (Table 2). Representative immunohistochemical features are shown in supplementary material, Figure S1.

Next, *KRAS* and *BRAF* mutations were analyzed. The frequency of *KRAS* mutations did not differ between cohorts 1 and 2 (45.0% [9/20] versus 45.8%

Table 2. Comparison of the molecular characteristics between cohorts 1 and 2

	Total (%)	Cohort 1 (%)	Cohort 2 (%)	P value
Total	44	20	24	
CMS				0.7456
1	0 (0)	0 (0)	0 (0)	
2	12 (27.3)	6 (30.0)	6 (25.0)	
3	32 (72.7)	14 (70.0)	18 (75.0)	
4	0 (0)	0 (0)	0 (0)	
<i>KRAS</i> mutation	20 (45.5)	9 (45.0)	11 (45.8)	1.0000
<i>KRAS</i> mutation type				1.0000
Transition	13 (65.0)	6 (66.7)	7 (63.6)	
Transversion	7 (35.0)	3 (33.3)	4 (36.4)	
<i>BRAF</i> mutation	0 (0)	0 (0)	0 (0)	1.0000

[11/24], respectively). Transitions were the most common *KRAS* mutation (13 cases, 29.5%) in both cohorts. No *BRAF* mutations were detected in either cohort. These mutation data are summarized in Table 2. In addition, the frequency of *KRAS* mutations differed significantly between left- and right-sided CRCs (7/25 versus 13/19).

Comprehensive analysis of mRNA and miRNA expression patterns in isolated cancer, adjacent normal, and distal normal tissues

mRNA expression profiling in cancer and adjacent normal tissues

We examined global mRNA expression profiles, analyzing 20 isolated CRC glands and compared them with normal isolated glands. First, we investigated differences in mRNA expression between isolated cancer or adjacent normal glands and isolated distal normal glands: 182 mRNAs (0 upregulated and 182 downregulated mRNAs) were identified as differentially expressed in adjacent normal glands relative to distal normal glands. Next, 357 mRNAs (137 upregulated and 220 downregulated) were identified as differentially expressed in cancer glands relative to distal normal glands (Figure 3A,B and supplementary material, Tables S1 and S2).

miRNA expression profiling in cancer and adjacent normal tissues

We compared the miRNA expression profiles of the cancer and adjacent normal glands with those of distal normal glands: 40 miRNAs (29 upregulated and 11 downregulated, respectively) were differentially expressed between adjacent normal and distal glands. In addition, 109 miRNAs (79 upregulated and 30 downregulated miRNAs, respectively) were differentially

expressed in the cancer glands relative to the distal normal glands (Figure 3C,D and supplementary material, Tables S3 and S4).

Ingenuity analysis of mRNA and miRNA expression levels and simple regression analysis of microarray data

We examined inverse associations between miRNAs and their target mRNAs using Ingenuity pathway analysis (IPA). Of the 357 and 109 differentially expressed mRNAs and miRNAs, respectively, in the cancer glands, 914 mRNA/miRNA pairs with inverse expression patterns were identified (supplementary material, Figure S2). Of the 182 and 40 differentially expressed mRNAs and miRNAs, respectively, identified in the adjacent normal glands, 178 mRNA/miRNA pairs with inverse expression patterns were identified (supplementary material, Figure S3). Simple regression analysis of the microarray data was performed to select candidate mRNA/miRNA pairs (supplementary material, Table S5 and S6). Based on these criteria, 329 of the 914 mRNA/miRNA pairs in the isolated cancer tissue (supplementary material, Figure S4) and 13 of the

178 mRNA/miRNA pairs in the adjacent normal tissue were retained (supplementary material, Figure S5).

GO analysis of differentially expressed mRNAs in cancer and adjacent normal tissues

To comprehensively understand the biological roles of differentially expressed mRNAs in CRC, GO analysis was performed to determine associated functions and pathways. Of the three GO categories, biological process, molecular function, and cellular component, the differentially expressed mRNAs in both adjacent normal and cancer glands were significantly associated with the biological process 'neutrophil degranulation', the molecular function 'RNA binding', and the cellular component 'extracellular exosome' (Figure 4).

Candidate genes identified by GO analysis of cancer and adjacent normal tissues and their inverse correlations with miRNA expression

We explored the inverse correlations between mRNAs and their corresponding miRNAs in cancer, adjacent normal, and distal normal glands. The common

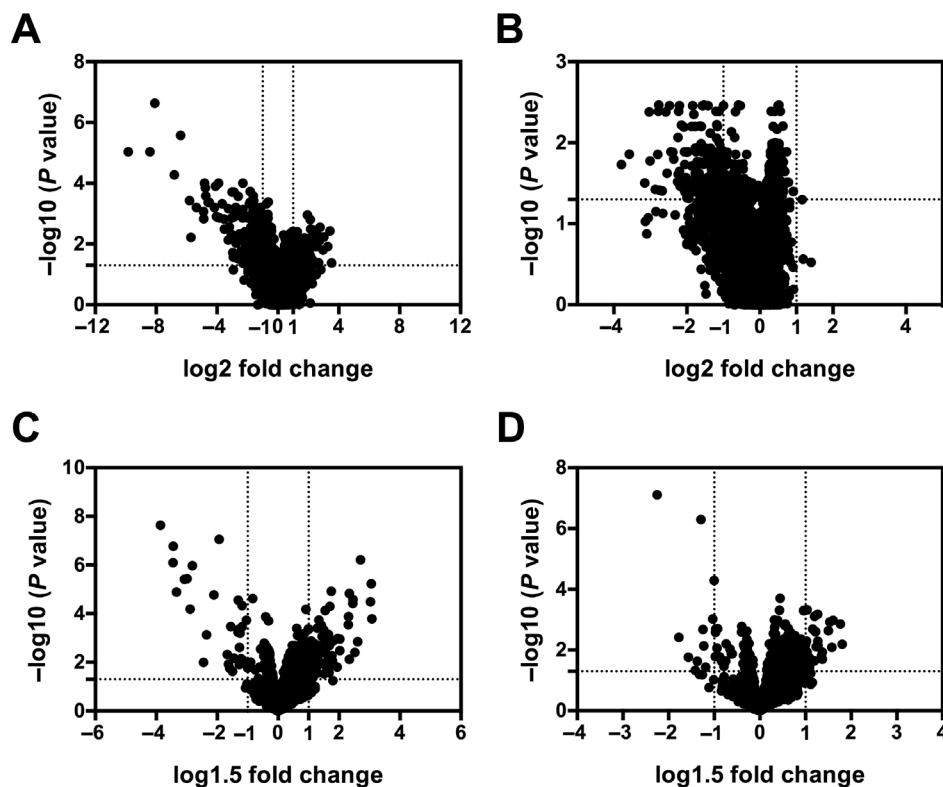


Figure 3. Volcano plots of differentially expressed genes. (A) Differential mRNA expression in cancer glands (357 total: 137 upregulated, 220 downregulated). (B) Differential mRNA expression in adjacent normal glands (182 total: 0 upregulated, 182 downregulated). (C) Differential miRNA expression in cancer glands (109 total: 79 upregulated, 30 downregulated). (D) Differential miRNA expression in adjacent normal glands (40 total: 29 upregulated, 11 downregulated).

candidate genes identified in the GO analysis of cancer and adjacent normal glands were analyzed. The inverse correlations between mRNA and miRNA expression in isolated cancer and adjacent normal glands are shown in supplementary material, Tables S5 and S6.

Candidate genes involved in neutrophil degranulation (GO:0043312) in both cancer and adjacent normal tissues and their inverse correlations with miRNA expression

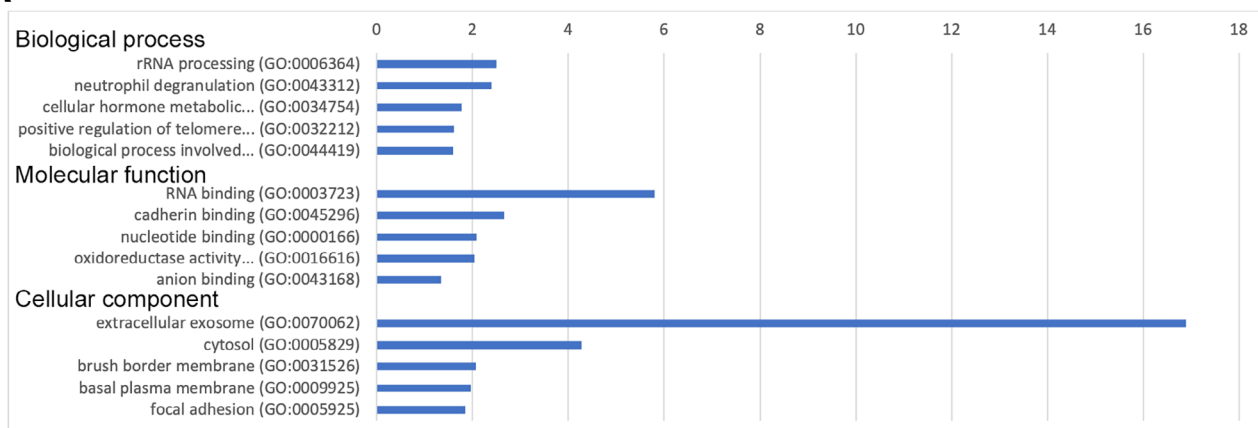
We compared the mRNAs associated with the GO term neutrophil degranulation differentially expressed in cancer glands with those differentially expressed in adjacent normal glands. Eight mRNAs (*CD59*, *CD63*, *CEACAM1* [carcinoembryonic antigen-related cell adhesion molecule 1], *DGAT10* [diacylglycerol O-acyltransferase 1], *HMGB1* [high mobility group

box 1], *MVP* [major vault protein], *PRDX6* [periredoxin 6], and *VAPA* [vacuole membrane protein associated protein]) were commonly expressed between cancer and adjacent normal glands (supplementary material, Figure S6). Six mRNA/miRNA pairs showed an inverse correlation between mRNA and miRNA expression in cancer glands, whereas two pairs, *CEACAM1*/miRNA-29a-3p and *CEACAM1*/miRNA-7114-5p, revealed an inverse correlation in adjacent normal glands (supplementary material, Table S7). Among these, only the *CEACAM1*/miRNA-7114-5p pair was common to the cancer and adjacent normal tissues (Figure 5A,B).

Candidate genes involved in RNA binding (GO:0003723) in both cancer and adjacent normal tissues

Next, we examined the candidate mRNAs associated with the GO term RNA binding that were commonly

A



B

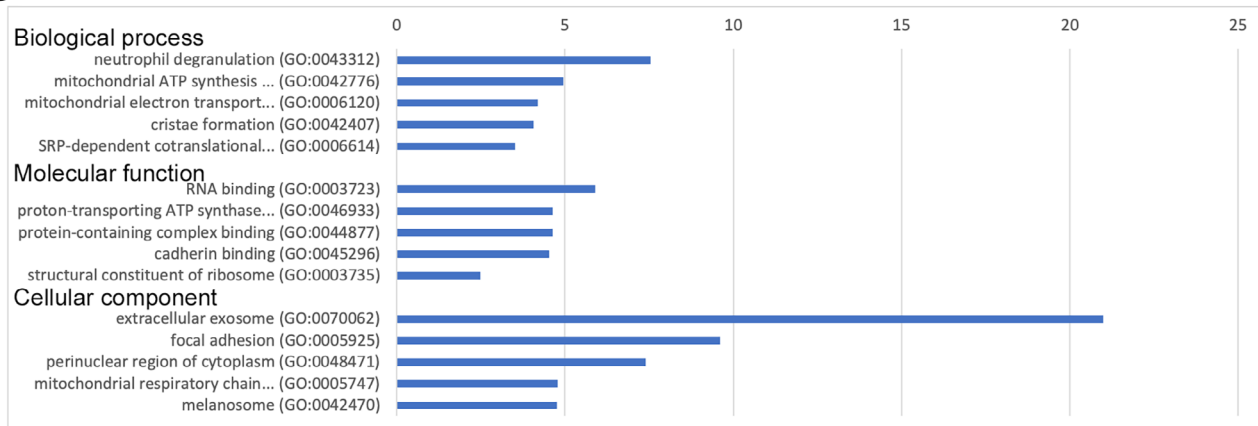


Figure 4. GO analysis of the significantly differentially expressed mRNAs in (A) the cancer and (B) adjacent normal tissues.

differentially expressed in cancer and adjacent normal glands (supplementary material, Figure S7). However, none of the common mRNAs formed an inverse relationship with any miRNA (supplementary material, Table S8).

Candidate genes involved in extracellular exosomes (GO:0070062) in both cancer and adjacent normal tissues

Among the differentially expressed candidate mRNAs associated with the GO term extracellular exosome (supplementary material, Figure S8), 20 were common to isolated cancer and adjacent normal tissues. These mRNAs formed 58 and 21 mRNA/miRNA pairs in cancer and adjacent normal samples, respectively. Among these, there was a statistical inverse correlation between 18 mRNA/miRNA pairs in the cancer tissue (supplementary material, Table S9). In the adjacent normal glands, three mRNA/miRNA pairs had inverse correlations: *AK1* (adenylate kinase 1)/miRNA-6780b-5p, *CEACAM1*/miRNA-29a-3p, and *CEACAM1*/miRNA-7114-5p. Two of these mRNA/miRNA pairs, *AK1*/miRNA-6780b-5p and *CEACAM1*/miRNA-7114-5p,

were common to cancer and adjacent normal glands (Figure 5).

Candidate mRNA/miRNA pairs validated by RT-PCR in cohort 2 (validation cohort)

We validated the inverse relationships between candidate mRNAs and miRNAs, including *AK1* versus miRNA-6780-5p and *CEACAM1* versus miRNA-7114-5p, in cancer and adjacent normal tissues, compared with distal normal tissues, using RT-PCR. The adjacent normal tissues used in the validation analysis were divided into two components. Expression of *CEACAM1* was inversely correlated with that of miRNA-7114-5p in cancer tissues and both components of adjacent normal tissues compared with distal normal tissues. We also found that *AK1* expression was negatively correlated with miRNA-6780-5p expression in isolated cancer tissues and both adjacent normal tissue components, compared with distal normal tissues. Correlation details are shown in Figure 6A. Finally, we examined differences in inverse correlations of the *CEACAM1*/miRNA-7114-5p and *AK1*/miRNA-6780b-5p pairs between left- and right-

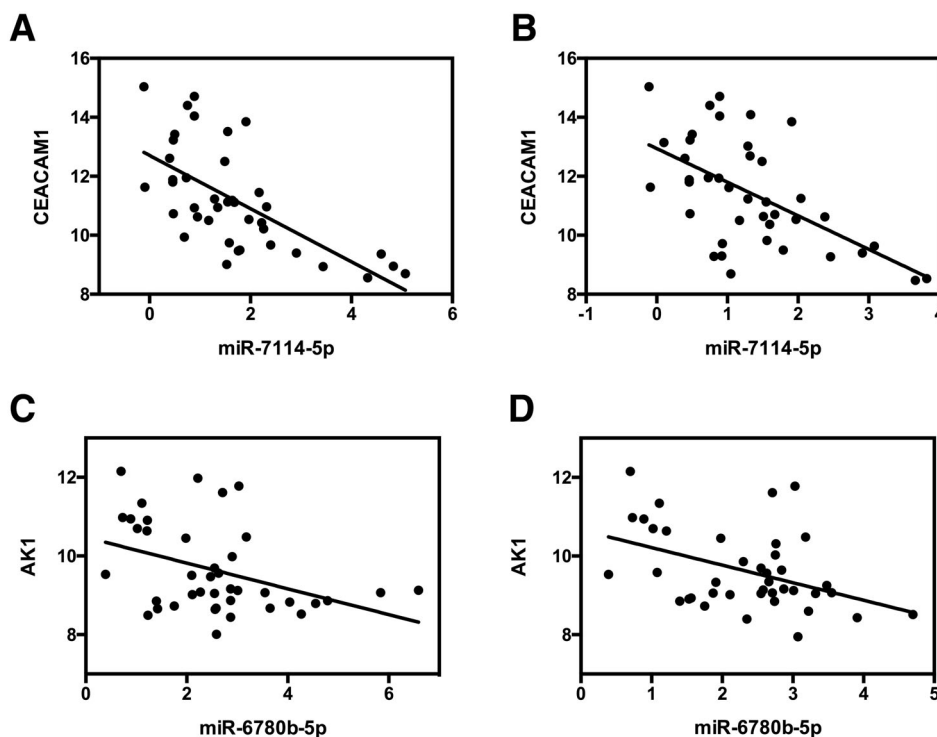


Figure 5. Simple linear regression analysis of *CEACAM1* versus miR7114-5p expression and *AK1* versus miRNA-6780b-5p. (A) *CEACAM1* and miR7114-5p expression in isolated cancer glands. (B) *CEACAM1* and miRNA-7114-5p expression in isolated adjacent normal crypts. (C) *AK1* and miRNA-6780b-5p expression in cancer glands. (D) *AK1* and miRNA-6780b-5p expression in isolated normal glands.

sided CRCs. A strict correlation was almost identified between *CEACAM1* and miRNA-7114-5p. However, no difference in the *AK1*/miRNA-6780b-5p correlation was found between left- and right-sided CRCs (Table 3).

Effects of transfected miRNA-7114-5p and miRNA-6780-5p mimics on *CEACAM1* and *AK1* expression

We transfected CRC cells with miRNA-7114-5p and miRNA-6780-5p mimics and assessed their effects on the expression of candidate target genes. Ectopic expression of miRNA-7114-5p suppressed *CEACAM1* expression, suggesting that *CEACAM1* is a potential target gene of miRNA-7114-5p in CRC cells

(Figure 6B: a–c). However, we were unable to demonstrate that ectopic expression of miRNA-6780-5p inhibits *AK1* expression (Figure 6B: d–f).

Immunohistochemical examination of *CEACAM1* and *AK1* in distal normal mucosa, adjacent normal mucosa, and cancer tissue

The immunostaining intensity in normal epithelial and cancer cells was classified as negative, weak, moderate, or strong, and the immunostaining area was semi-quantified and categorized as 0, 1–25, 26–50, or 51–100%, respectively. An immunohistochemical score (sum of the immunostaining intensity and area scores) ≥ 4 was considered positive [22]. Although there was a significant difference in the *CEACAM1*

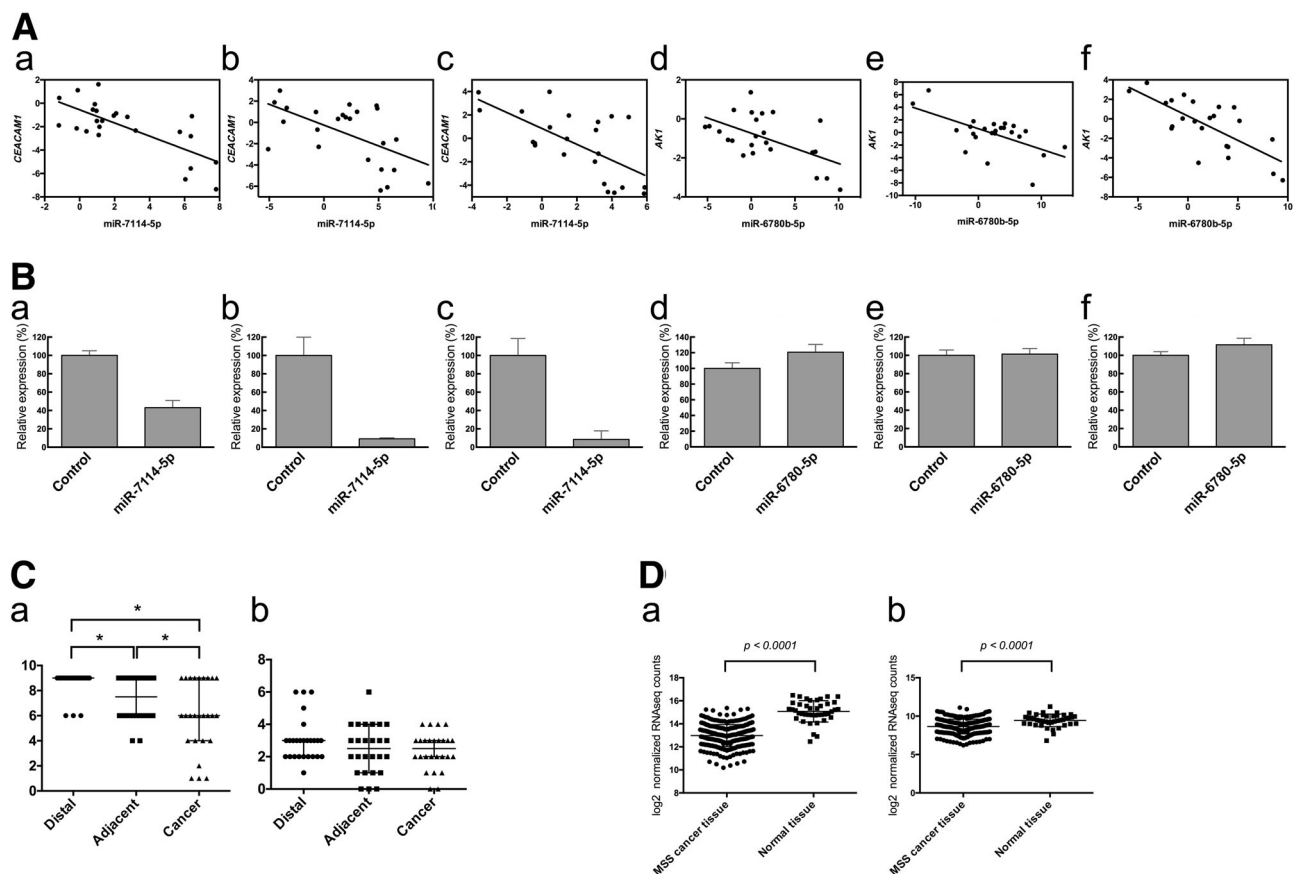


Figure 6. Validation of miRNA/mRNA associations by regression analysis, transfection assay, immunohistochemical analysis, and TCGA analysis. (A) Regression analyses of *CEACAM1* versus miRNA-7114-5p and *AK1* versus miRNA-6780b-5p expression (cohort 2). (a–c) *CEACAM1*/miRNA-7114-5p correlation in cancer glands (a), adjacent normal glands at site 1 (b), and adjacent normal glands at site 2 (c). (d–f) *AK1*/miRNA-6780b-5p correlation in cancer glands (d), adjacent normal glands at site 1 (e), and adjacent normal glands at site 2 (f). (B) Transfection assay. (a–c) Expression of *CEACAM1* in DLD1 (a), HCT116 (b), and RKO (c) cells. (d–f) Expression of *AK1* in DLD1 (d) HCT116 (e), and RKO cells (f). (C) Immunohistochemical analyses of *CEACAM1* (a) and *AK1* (b) in cohort 2. * $p < 0.05$. (D) Expression of *CEACAM1* (a) and *AK1* (b) in cancer versus normal tissues in TCGA database.

Table 3. Simple regression analysis between mRNAs and miRNAs in the validation analysis

	Regression coefficient	Adjusted coefficient of determination	95% CI	P value
<i>CEACAM1</i> versus miRNA-7114-5p				
Isolated cancer glands	-0.56	0.5	-0.8, -0.33	< 0.0001
Left side	-0.73	0.67	-1.05, -0.41	0.0004
Right side	-0.45	0.34	-0.86, -0.04	0.0338
Isolated adjacent normal glands at site 1	-0.39	0.29	-0.64, -0.14	0.0038
Left side	-0.32	0.26	-0.63, -0.01	0.0444
Right side	-0.52	0.3	-1.04, 0	0.0484
Isolated adjacent normal glands at site 2	-0.68	0.4	-1.04, -0.32	0.0008
Left side	-0.79	0.77	-1.07, -0.51	0.0001
Right side	-0.38	-0.04	-1.44, 0.68	0.4392
<i>AK1</i> versus miRNA-6780b-5p				
Isolated cancer glands	-0.16	0.28	-0.26, -0.05	0.0043
Left side	-0.15	0.29	-0.28, -0.02	0.0321
Right side	-0.17	0.18	-0.39, 0.05	0.1086
Isolated adjacent normal glands at site 1	-0.33	0.33	-0.52, -0.13	0.0019
Left side	-0.18	-0.01	-0.61, 0.24	0.3586
Right side	-0.36	0.5	-0.6, -0.12	0.0086
Isolated adjacent normal glands at site 2	-0.5	0.5	-0.71, -0.28	<0.0001
Left side	-0.59	0.71	-0.84, -0.34	0.0003
Right side	-0.2	-0.05	-0.8, 0.41	0.4795

CI, confidence interval.

immunohistochemical score between distal normal and adjacent normal mucosae and between adjacent and cancer cells (Figure 6C: a), no such differences in the AK1 immunohistochemical score were found (Figure 6C: b). Representative immunohistochemical features are shown in supplementary material, Figure S9. Antibodies used for immunohistochemical staining are shown in supplementary material, Table S10.

Associations between *CEACAM1* and miRNA-7114-5p and between *AK1* and miRNA-6780-5p expression in TCGA

We attempted to explore miRNA and mRNA expression data from The Cancer Genome Atlas (TCGA) colon adenocarcinoma RNA-Seq, HTSeq-Count, and miRNA-Seq datasets miRNA profiling data from NCI Genomic Data Commons (<https://gdc.cancer.gov>). The expression of *CEACAM1* and *AK1* was downregulated in this database (Figure 6D: a and b). However, we could not find public data for the association between *CEACAM1* and miRNA-7114-5p or between *AK1* and miRNA-6780-5p expression.

Discussion

This study investigated the dysregulated mRNA/miRNA pairs found in CRC tissues and those in

adjacent normal tissues and their associations with specific pathways. The background levels of two mRNA/miRNA pairs (*AK1*/miRNA-6780-5p and *CEACAM1*/miRNA-7114-5p) in the adjacent normal mucosa were closely associated with development of CRC. In addition, we examined differences in the mRNA/miRNA pairs identified between left- and right-sided CRC. This may be important considering the molecular difference between the two sides. This is the first study to find that inversely correlated mRNA/miRNA pairs in the adjacent normal mucosa independently predict pathway-specific predisposition to CRC. The results suggest that field cancerization of adjacent normal mucosa due to retention of inverse correlations between specific mRNA and miRNA pairs plays an important role in CRC development, and that CRC pathogenesis is predetermined by early molecular events defined by interactions between mRNAs and miRNAs expressed in adjacent mucosa.

We examined the molecular characteristics of CRCs in cohorts 1 and 2 to identify differences between cohorts [20,21]. There were no significant differences in the frequencies of molecular characteristics between the cohorts, suggesting common alterations in cohorts 1 and 2.

Two main mechanisms contribute to gene expression inhibition: DNA methylation and high miRNA expression [8,9,14]. This study focused on miRNA regulation of mRNAs and found inverse correlations between *CEACAM1* and miRNA-7114-5p and between *AK1* and

miRNA-6780-5p. This suggests that *CEACAM1* and *AK1* are already downregulated in adjacent normal mucosa, prior to CRC development. In addition, this finding is very interesting in that normal crypts adjacent to CRC with dysregulated expression of *CEACAM1* and *AK1* are not morphologically different from crypts obtained from distal mucosa. Another important finding of this study is that *CEACAM1* and *AK1* are regulated by specific miRNAs (miRNA-7114-5p and miRNA-6780-5p, respectively), which were not associated with the expression of other mRNAs that contribute to cancer development. Our results suggest that inverse correlations of *CEACAM1*/miRNA-7114-5p and possibly *AK1*/miRNA-6780-5p predispose to CRC development. In addition, we found at least inverse correlations of the *CEACAM1*/miRNA-7114-5p pair in left-sided and right-sided colons, suggesting common alteration in both segments.

Carcinoembryonic antigen-related cell adhesion (CEACAM) molecules are members of the glycosylphosphatidylinositol-linked immunoglobulin superfamily [23]. Previous studies have demonstrated that expression of *CEACAM-1*, also known as CD66a, is dysregulated (down- or upregulated) in several human cancers, including melanoma and lung, gastric, and colon cancers, compared with corresponding normal tissues [23–26]. This suggests that CEACAM-1 functions in both a tumor-suppressive manner and an oncogenic manner in carcinogenesis [24–27]. In addition, CEACAM-1 is the major antigen of the CD66 cluster of granulocyte differentiation antigens [26,28] and was implicated in the neutrophil degranulation pathway, which is associated with cancer development [26,28]. Functionally, it was shown that CEACAM1 is closely associated with reduced cell adhesion and apoptosis in cancer cells, suggesting enhanced cancer mobility and consequently cancer metastasis [26]. Reduced expression of CEACAM1 in CRC is not well reported, and its mechanism remains unknown. One study showed that *CEACAM1* expression was regulated by miRNA-342 and induced lumen formation in a three-dimensional model of mammary glands [29]. However, miRNA-regulated expression of *CEACAM1* has not been reported in CRC. In the present study, the inverse relationship between *CEACAM1* and miRNA-7114-5p expression was confirmed in two different adjacent mucosal components in the validation cohort. This association was supported by transfection and immunohistochemical assays. *CEACAM1* expression was also downregulated in the public genome database, suggesting *CEACAM1* to be downregulated by high miRNA-7114-5p expression. Thus, downregulated *CEACAM1* expression by miRNA-7114-5p may play an important role in early colorectal carcinogenesis.

There is growing evidence of the significance of metabolic signaling in human diseases [30,31]. New molecules crucial for cell homeostasis and function are being investigated. High-energy phosphoryl transfer systems are required to mediate intracellular communication between ATP-consuming and ATP-producing cellular compartments and consequently can maintain normal cellular growth and development [30,31]. The major component of the cellular phosphotransfer system is AK [30,31]. In recent studies, the significance of organized phosphotransfer was demonstrated in human diseases [30,31]. In contrast, extracellular exosomes are key structures enabling communication between cells and contributing to tumor development [32,33]. A recent study showed that AK1 plays an important role in extracellular exosomes [33]. These findings suggest that the metabolic reprogramming associated with AK1 may be critical for CRC pathogenesis. In the present study, we demonstrated that reduced *AK1* expression was present in both CRC and also the adjacent normal mucosa and was associated with high expression of miRNA-6780-5p. This association was confirmed in the validation cohort samples. Additionally, downregulated expression of *AK1* was verified in the public genome database. However, such association was not found in transfection and immunohistochemical assays. This discrepancy may be due to several reasons. First, *AK1* expression may be regulated by multiple factors, such as DNA methylation and different miRNAs. Second, *in vivo* results cannot always be replicated *in vitro*. Third, the discrepancy between the mRNA and immunohistochemical expression of AK1 may be due to non-specific immunohistochemical reactions. Further studies will be required to identify an association between AK1 and miRNA-6780-5p.

This study has some limitations. First, heterogeneous miRNA/mRNA expression within the same tumor was not well evaluated. Sampling was conducted at the invasive front of the CRC tumor, which is thought to be an appropriate sampling site. Second, non-invasive lesions, such as colorectal adenoma, were not examined; it would be interesting to identify normal mucosal risk in early-stage colorectal neoplastic lesions. In particular, identification of the morphological characteristics of normal crypts in adjacent normal mucosa would have a significant clinical impact. Third, the sample size was small. However, isolated glands from which pure target cells can be obtained only allow the evaluation of commonly dysregulated miRNA/mRNA pairs between CRC and adjacent normal mucosa. We believe that this finding is important for elucidating early-stage colorectal carcinogenesis, in which morphological changes are not yet observed. Finally, a prospective study may be required to

identify the pathological and clinical significance of our findings.

In conclusion, we examined correlations between mRNA and miRNA pairs in adjacent normal mucosa and CRC tissue using an array-based analysis. As a result, we found inverse correlations between the expression of *CEACAM1* and miRNA-7114-5p and between that of *AKI* and miRNA-6780-5p in both CRC and also the adjacent normal mucosa, compared with the distal mucosa. These associations were validated in a second cohort by RT-PCR. Using pathway analysis, *CEACAM1* and *AKI* were found to be closely associated with neutrophil degranulation and extracellular exosomes, respectively. These pathways were implicated as cancer-associated pathways in recent studies [26,30]. The current results suggest that the inverse correlation between *CEACAM1* and miRNA-7114-5p and the possible inverse correlation between *AKI* and miRNA-6780-5p contribute to early development of CRC. Finally, these two mRNA/miRNA pairs may help predict the risk of tumor growth in the normal mucosa. However, further study is required to determine whether they affect colorectal carcinogenesis.

Acknowledgements

We gratefully acknowledge the technical assistance from Ms. E. Sugawara and Mrs. Ishikawa. We also thank the members of the Department of Molecular Diagnostic Pathology, Iwate Medical University, for the additional support.

Author contributions statement

YI contributed to the preparation of the manuscript, including the data collection and sampling. MO, NU and RS performed all data collection and statistical analyses. NY contributed to the pathological analyses. KO, AS and TM provided clinical support during the preparation of the manuscript. TN and HS assisted with the molecular analyses. TS contributed to the preparation of the manuscript, including all aspects of the data collection and analysis.

Data availability statement

The data supporting the findings of our study are available from the corresponding author upon reasonable request.

References

- Haggard FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clin Colon Rectal Surg* 2009; **22**: 191–197.
- Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Prz Gastroenterol* 2019; **14**: 89–103.
- Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. *Gastroenterology* 2010; **138**: 2059–2072.
- Hamilton SR, Sekine S. *Tumors of the Colon and Rectum, Malignant Epithelial Tumors. WHO Classification of Tumours of the Digestive System*. International Agency for Research on Cancer: Lyon, 2019; 208–212.
- Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012; **487**: 330–337.
- Amaro A, Chiara S, Pfeffer U. Molecular evolution of colorectal cancer: from multistep carcinogenesis to the big bang. *Cancer Metastasis Rev* 2016; **35**: 63–74.
- Paterson C, Clevers H, Bozic I. Mathematical model of colorectal cancer initiation. *Proc Natl Acad Sci U S A* 2020; **117**: 20681–20688.
- Worthley DL, Whitehall VL, Buttenshaw RL, et al. DNA methylation within the normal colorectal mucosa is associated with pathway-specific predisposition to cancer. *Oncogene* 2010; **29**: 1653–1662.
- Sugai T, Yoshida M, Eizuka M, et al. Analysis of the DNA methylation level of cancer-related genes in colorectal cancer and the surrounding normal mucosa. *Clin Epigenetics* 2017; **9**: 55.
- Patel A, Tripathi G, Gopalakrishnan K, et al. Field cancerization in colorectal cancer: a new frontier or pastures past? *World J Gastroenterol* 2015; **21**: 3763–3772.
- Lochhead P, Chan AT, Nishihara R, et al. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Mod Pathol* 2015; **28**: 14–29.
- Hawthorn L, Lan L, Mojica W. Evidence for field effect cancerization in colorectal cancer. *Genomics* 2014; **103**: 211–221.
- Chai H, Brown RE. Field effect in cancer—an update. *Ann Clin Lab Sci* 2009; **39**: 331–337.
- Zhou X, Xu X, Wang J, et al. Identifying miRNA/mRNA negative regulation pairs in colorectal cancer. *Sci Rep* 2015; **5**: 12995.
- Peng Y, Croce CM. The role of microRNAs in human cancer. *Signal Transduct Target Ther* 2016; **1**: 15004.
- Slattery ML, Herrick JS, Pellatt DF, et al. MicroRNA profiles in colorectal carcinomas, adenomas and normal colonic mucosa: variations in miRNA expression and disease progression. *Carcinogenesis* 2016; **37**: 245–261.
- Hamfjord J, Stangeland AM, Hughes T, et al. Differential expression of miRNAs in colorectal cancer: comparison of paired tumor tissue and adjacent normal mucosa using high-throughput sequencing. *PLoS One* 2012; **7**: e34150.
- Japanese Society for Cancer of the Colon and Rectum. *Japanese Classification of Colorectal Carcinoma* (Second English Edition). Kanehara Co: Tokyo, 2009; 30–63.
- Sugai T, Osakabe M, Niinuma T, et al. Comprehensive analyses of microRNA and mRNA expression in colorectal serrated lesions

- and colorectal cancer with a microsatellite instability phenotype. *Genes Chromosomes Cancer* 2022; **61**: 161–171.
20. Guinney J, Dienstmann R, Wang X, *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015; **21**: 1350–1356.
 21. Li X, Larsson P, Ljuslinder I, *et al.* A modified protein marker panel to identify four consensus molecular subtypes in colorectal cancer using immunohistochemistry. *Pathol Res Pract* 2021; **220**: 153379.
 22. Hashimoto M, Uesugi N, Osakabe M, *et al.* Expression patterns of microenvironmental factors and tenascin-C at the invasive front of stage II and III colorectal cancer: novel tumor prognostic markers. *Front Oncol* 2021; **11**: 690816.
 23. Takeuchi A, Yokoyama S, Nakamori M, *et al.* Loss of CEACAM1 is associated with poor prognosis and peritoneal dissemination of patients with gastric cancer. *Sci Rep* 2019; **9**: 12702.
 24. Thies A, Moll I, Berger J, *et al.* CEACAM1 expression in cutaneous malignant melanoma predicts the development of metastatic disease. *J Clin Oncol* 2002; **20**: 2530–2536.
 25. Laack E, Nikbakht H, Peters A, *et al.* Expression of CEACAM1 in adenocarcinoma of the lung: a factor of independent prognostic significance. *J Clin Oncol* 2002; **20**: 4279–4284.
 26. Han ZM, Huang HM, Sun YW. Effect of CEACAM-1 knockdown in human colorectal cancer cells. *Oncol Lett* 2018; **16**: 1622–1626.
 27. Mollinedo F. Neutrophil degranulation, plasticity, and cancer metastasis. *Trends Immunol* 2019; **40**: 228–242.
 28. Uribe-Querol E, Rosales C. Neutrophils in cancer: two sides of the same coin. *J Immunol Res* 2015; **2015**: 983698.
 29. Weng C, Nguyen T, Shively JE. miRNA-342 regulates CEACAM1-induced lumen formation in a three-dimensional model of mammary gland morphogenesis. *J Biol Chem* 2016; **291**: 16777–16786.
 30. Klepinin A, Zhang S, Klepinina L, *et al.* Adenylate kinase and metabolic signaling in cancer cells. *Front Oncol* 2020; **10**: 660.
 31. Qin T, Zhao H, Shao Y, *et al.* High expression of AK1 predicts inferior prognosis in acute myeloid leukemia patients undergoing chemotherapy. *Biosci Rep* 2020; **40**: BSR20200097.
 32. Su C, Zhang J, Yarden Y, *et al.* The key roles of cancer stem cell-derived extracellular vesicles. *Signal Transduct Target Ther* 2021; **6**: 109.
 33. Dai J, Su Y, Zhong S, *et al.* Exosomes: key players in cancer and potential therapeutic strategy. *Signal Transduct Target Ther* 2020; **5**: 145.
 34. Boland CR, Thibodeau SN, Hamilton SR, *et al.* A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* 1998; **58**: 5248–5257.
 35. Sugai T, Eizuka M, Takahashi Y, *et al.* Molecular subtypes of colorectal cancers determined by PCR-based analysis. *Cancer Sci* 2017; **108**: 427–434.
 36. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta C(T)) method. *Methods* 2001; **25**: 402–408.
- References 34–36 are cited only in the supplementary material.

SUPPLEMENTARY MATERIAL ONLINE

Supplementary materials and methods

Figure S1. Representative staining images for CMS classification

Figure S2. Network of the inversely correlated mRNAs and miRNAs in cancer tissues

Figure S3. Network of the inversely correlated mRNAs and miRNAs in adjacent normal tissues

Figure S4. Network of the mRNAs and miRNAs with significant inverse correlations in cancer tissues

Figure S5. Network of the mRNAs and miRNAs with significant inverse correlations in the adjacent normal tissues

Figure S6. Venn diagram of mRNAs associated with neutrophil degranulation (GO:0043312) that were differentially expressed in CRC with the MSS phenotype and in the adjacent normal mucosa compared with the distal mucosa

Figure S7. Venn diagram of mRNAs associated with RNA binding (GO:0003723) that were differentially expressed in CRC with the MSS phenotype and in adjacent normal mucosa compared with the distal mucosa

Figure S8. Venn diagram of mRNAs associated with the extracellular exosome (GO:0070062) that was differentially expressed in CRC with the MSS phenotype and adjacent normal mucosa compared with the distal mucosa

Figure S9. Representative immunohistochemical staining images for CEACAM1 and AK-1

Table S1. Differentially expressed mRNAs in cancer glands (cohort 1)

Table S2. Differentially expressed mRNAs in adjacent normal glands (cohort 1)

Table S3. Differentially expressed miRNAs in cancer glands (cohort 1)

Table S4. Differentially expressed miRNAs in adjacent normal glands (cohort 1)

Table S5. Simple regression analysis between mRNAs and miRNAs in cancer glands (cohort 1)

Table S6. Simple regression analysis between mRNAs and miRNAs in adjacent normal glands (cohort 1)

Table S7. Inversely correlated miRNAs and mRNAs common between cancer glands and adjacent normal glands involved in 'neutrophil degranulation (GO:0043312)'

Table S8. Inversely correlated miRNAs and mRNAs common between cancer glands and adjacent normal glands involved in 'RNA binding (GO:0003723)'

Table S9. Inversely correlated miRNAs and mRNAs common between cancer glands and adjacent normal glands involved in 'extracellular exosomes (GO:0070062)'

Table S10. Antibodies used for immunohistochemical staining