

Comprehensive genomic analysis contrasting primary colorectal cancer and matched liver metastases

AKIO SHIOMI^{1,2}, MASATOSHI KUSUHARA³, TAKASHI SUGINO⁴, TEIICHI SUGIURA⁵,
KEIICHI OHSHIMA⁶, TAKESHI NAGASHIMA^{7,8}, KENICHI URAKAMI⁷,
MASAKUNI SERIZAWA⁹, HIDEYUKI SAYA² and KEN YAMAGUCHI¹⁰

¹Division of Colon and Rectal Surgery, Shizuoka Cancer Center, Shizuoka 411-8777; ²Division of Gene Regulation Institute for Advanced Medical Research, Keio University School of Medicine, Tokyo 160-858;

³Regional Resources Division, Shizuoka Cancer Center Research Institute;

Divisions of ⁴Pathology and ⁵Hepato-Biliary-Pancreatic Surgery, Shizuoka Cancer Center; ⁶Medical Genetics Division;

⁷Cancer Diagnostics Research Division, Shizuoka Cancer Center Research Institute, Shizuoka 411-8777;

⁸SRL, Inc., Tokyo 163-0409; ⁹Drug Discovery and Development Division,

Shizuoka Cancer Center Research Institute; ¹⁰Shizuoka Cancer Center, Shizuoka 411-8777, Japan

Received November 25, 2020; Accepted March 9, 2021

DOI: 10.3892/ol.2021.12727

Abstract. Recent studies have revealed that colorectal cancer (CRC) displays intratumor genetic heterogeneity, and that the cancer microenvironment plays an important role in the proliferation, invasion and metastasis of CRC. The present study performed genomic analysis on paired primary CRC and synchronous colorectal liver metastasis (CRLM) tissues collected from 22 patients using whole-exome sequencing, cancer gene panels and microarray gene expression profiling. In addition, immunohistochemical analysis was used to confirm the protein expression levels of genes identified as highly expressed in CRLM by DNA microarray analysis. The present study identified 10 genes that were highly expressed in CRLM compared with in CRC, from 36,022 probes obtained

from primary CRC, CRLM and normal liver tissues by gene expression analysis with DNA microarrays. Of the 10 genes identified, five were classified as encoding 'matricellular proteins' [(osteopontin, periostin, thrombospondin-2, matrix Gla protein (MGP) and glycoprotein nonmetastatic melanoma protein B (GPNMB)] and were selected for immunohistochemical analysis. Osteopontin was strongly expressed in CRLM (6 of 22 cases: 27.3%), but not in CRC (0 of 22: 0%; $P=0.02$). Periostin also exhibited strong immunoreactivity in CRLM (17 of 22: 68.2%) compared with in CRC (7 of 22: 31.8%; $P=0.006$). Thrombospondin-2 exhibited strong immunoreactivity in both CRC and CRLM (54.5% in CRC, 45.5% in CRLM; $P=0.55$). GPNMB and MGP were rarely positive for both CRC and CRLM. A comparison of immunoreactive positive factors for these five genes revealed the complexities of gene expression in CRLM. Of the cases examined, 16 (72.7%) cases of CRC showed zero or only one positive immunoreactive factor. By contrast, CRLM showed more frequent and multiple immunoreactive factors; for example, 16 cases (72.7%) shared two or more factors, which was statistically more frequent than in CRC ($P=0.007$). The present study revealed the genomic heterogeneity between paired primary CRC and CRLM, in terms of cancer cell microenvironment. This finding may lead to novel diagnostic and therapeutic targets in the era of genome-guided personalized cancer treatment.

Correspondence to: Dr Akio Shiomi, Division of Colon and Rectal Surgery, Shizuoka Cancer Center, 1007 Shimonagakubo, Nagaizumi, Sunto, Shizuoka 411-8777, Japan
E-mail: a.shiomi@scchr.jp; koolina2001@hotmail.co.jp

Abbreviations: CRC, colorectal cancer; CRLM, colorectal liver metastasis; Project HOPE, High-tech Omics-based Patient Evaluation; WES, whole-exome sequencing; GEP, gene expression profiling; SNV, single nucleotide variants; GPNMB, glycoprotein nonmetastatic melanoma protein B; MGP, matrix Gla protein; THBS2, thrombospondin-2; TMB, tumor mutation burden; SMAP, Shizuoka Multi-omics Analysis Protocol; ECM, extracellular matrix; CMS, consensus molecular subtypes; TCGA, The Cancer Genome Atlas; ICGC, International Cancer Genome Consortium; SIBLING, small integrin-binding ligand N-linked glycoprotein; MMP, matrix metalloproteinases

Key words: colorectal cancer, liver metastasis, gene expression, heterogeneity, microenvironment

Introduction

Colorectal cancer (CRC) is a common malignancy; it has the third highest incidence rate among all cancers and is the second leading cause of cancer-related deaths worldwide (1). Despite advances in our understanding of the clinicopathological features of CRC and the improvements in diagnosis and treatment thereof, mortality from CRC is expected to increase, with approximately 860,000 deaths reported each year (2). Unfortunately, at the time of diagnosis, 19.6% of

patients with primary CRC have concurrent distant metastases, with the highest frequency observed in liver (10.9%), peritoneum (4.5%), and lung (2.4%) (3). The prognosis of patients with liver metastasis from colorectal cancer (CRLM) has improved dramatically with the availability of new and effective cytotoxic and targeted agents, as well as aggressive surgical resection. Moreover, recent improvements in innovative agents have led to an increased response rate in unresectable CRLM. The surgical resection rate after downstaging in initially unresectable CRLM has been reported to be up to 40% (so called 'conversion therapy') (4). Despite these developments, the oncological outcomes for patients with CRLM remains unsatisfactory; for example, CRLM recurrence after hepatectomy is common, with a 50% rate reported in remnant liver (5-7).

Advances in molecular biology over the past decade have facilitated a better understanding of the development of several kinds of cancers, and a more precise use of innovative targeted therapies. Indeed, recent studies have shown that several cancers display intratumor genetic heterogeneity. Given the intratumor genetic heterogeneity described in several cancers, the metastatic process itself may result in clonal selection in the progression from primary to metastatic disease. The proliferation, invasion, and metastasis of cancer cells are not defined by the nature of the cancer cell itself, but rather by the adaptation of the microenvironment via interactions between the cancer cell and its surrounding tissues (8,9).

It is considered that elucidating intratumor genetic heterogeneity will lead to new diagnostic and therapeutic methods for CRC. Herein, we performed molecular analysis on paired patients with primary CRC and synchronous CRLM, resected at the same institute. We investigated the molecular characteristics using whole-exome sequencing, cancer gene panels, fusion gene panels and microarray gene expression profiling. A notable feature of this study is that it fully matches all of the clinicopathological data, as the cases are from the same medical institution. Most previous studies used metachronous or unpaired patient samples. Consequently, the analyses presented in those studies could be biased by intertumoral heterogeneity and the administration of chemotherapy between the time of resection of primary CRC and metastatic sites, which might alter the genetic characteristics of the clones. Moreover, previous studies used a limited set of biomarkers.

To our knowledge, no other study has analyzed such a completely paired sample of primary CRC and synchronous CRLM using a major complement of exhaustive genetic analyses with next-generation sequencing.

Materials and methods

Clinical data. Surgically resected tumor specimens, normal liver tissues and corresponding peripheral blood samples were obtained from 22 consecutive patients who underwent both colectomy and hepatectomy for CRC and synchronous CRLM between January 2014 and March 2015. All patients were enrolled in Project HOPE (High-tech Omics-based Patient Evaluation), a study launched at our institute with the aim of evaluating the biological characteristics of cancer by multiomics-based analyses (10). The clinicopathological data of patients were reviewed retrospectively. A prospective

colorectal database, containing information regarding patient characteristics, preoperative assessment, operative characteristics, postoperative complications, pathological characteristics, and oncological outcomes, maintained at the hospital was used for this retrospective analysis.

Project HOPE (High-tech Omics-based Patient Evaluation). In the present study, we evaluated fresh frozen tumor tissues obtained from both primary CRC and CRLM using whole-exome sequencing (WES), cancer gene panel sequencing, fusion gene panel sequencing and microarray-based gene expression profiling (GEP). We also acquired peripheral blood samples from patients and paired them with the corresponding resected tissue samples from the HOPE study.

Ethical considerations. The research plan of Project HOPE was designed according to the revised Ethical Guidelines for Human Genome/Gene Analysis Research in Japan (11), and the study protocol was approved by the Institutional Review Board at the Shizuoka Cancer Center. Written informed consent was obtained from all participants. This retrospective study was also approved by the same board (Authorization no. 30-5).

Whole-exome sequencing. DNA was extracted from blood and flash-frozen tissues using a QIAamp DNA Blood Kit (cat. no. 51185; Qiagen), except that the tissues were treated with Proteinase (cat. no. K19133; Qiagen). DNA was quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Inc.) and a Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Inc.).

DNA sample with A 260/280 ratio >1.8 was used for DNA sequencing. The exome library used for WES was prepared using an Ion Torrent AmpliSeq RDY Exome Kit (cat. no. A27193; Thermo Fisher Scientific, Inc.) in accordance with the manufacturer's instructions. A total of 100 ng of DNA was used for target amplification under the following conditions: 99°C for 2 min, followed by 10 cycles at 95°C for 15 sec and 60°C for 16 min, and a final hold at 10°C. Incorporated primer sequences were partially digested using FuPa reagent (Ion Torrent AmpliSeq RDY Exome Kit).

Proton adapters were ligated to the amplicons at 22°C for 30 min, followed by 72°C for 10 min, and the library was purified with Agencourt Ampure XT beads (cat. no. A63881, Thermo Fisher Scientific, Inc.). Libraries were quantified using a quantitative polymerase chain reaction (qPCR), and DNA (8 pM) was sequenced using a semiconductor DNA sequencer (Ion Torrent Proton Sequencer; Thermo Fisher Scientific, Inc.) by 200 cycles single-end sequencing according to the manufacturer's instructions. The average values of coverage of WES were about 100. Matched tumor-normal pair somatic variants were identified using Ion Reporter ver. 4.4 software (Thermo Fisher Scientific Inc.) (12) after base calling, quality trimming, and mapping to the hg19/GRCh37 reference genome using Torrent Suite software ver. 4.4 (Thermo Fisher Scientific, Inc.) (13). In this step, sequence data derived from tumor and blood samples were analyzed separately, and the latter were used as matched controls. In this process, only somatic variants remain after the subtraction of variants from blood data from those acquired from tumor data. In this variant-call workflow, we identified somatic mutations that satisfied the thresholds quality score ≥ 60 or depth of

coverage ≥ 20 . Somatic variants were inspected manually using the Integrative Genomics Viewer (14). Annotation of detected single nucleotide variants (SNVs) was performed using the following databases that included germline and somatic variants: COSMIC (15), ClinVar (16), dbSNP (17), UniProt (18) and DrugBank (19). The details are described by Nagashima *et al* (10,20).

Cancer cell gene panel sequencing. The DNA library comprising 409 genes implicated in cancer was prepared using the Ion AmpliSeq Comprehensive Cancer Panel Kit (cat. no. 4477685; Thermo Fisher Scientific, Inc.). A total of 10 ng of tumor DNA was used for target amplification under the following conditions: 99°C for 2 min, followed by 12 cycles at 99°C for 15 sec and 60°C for 16 min, and a final hold at 10°C. After adapter ligation and library purification, libraries were quantified using a quantitative polymerase chain reaction (qPCR), and DNA (8 pM) was sequenced using Ion Torrent Proton Sequencer according to the manufacturer's instructions. The average values of coverage of WES analysis were $\sim 1,200$. Data processing and annotation were the same as described above for WES, except for variant calls identified by subtracting 409 gene variants from WES blood data. Ion Torrent The details are described by Shimoda *et al* (21).

Comprehensive gene expression analysis using DNA microarray. Fresh tumor and adjacent normal tissues were soaked in RNAlater reagent (Thermo Fisher Scientific, Inc.), and total RNA was isolated and purified using an miRNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. RNA quality was evaluated using an RNA integrity number, which was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies). RNA samples with an RNA integrity number ≥ 6.0 were used for gene expression analysis. Gene expression analysis was performed using a SurePrint G3 Human Gene Expression 8x60K v2 Microarray (Agilent Technologies) kit using a One-color Low Input Quick Amp Labeling kit (Agilent Technologies) according to the manufacturer's instructions. Data processing to generate raw signal intensity data was performed with GeneSpring version 13.1.1 software (Agilent Technologies). Data analysis was performed using GeneSpring GX (Agilent Technologies) and Microsoft Excel. Raw signal intensity values were normalized to the 75th percentile and translated into \log_2 ratio against average. Among all 50,599 probe sets, 36,022 probes which excluded too low signal, too stable signal or abnormal data were used for the following analysis. For the extraction of genes with higher expression in CRLM than in CRC, the normalized signal values for each probe were averaged in the two groups and in the group of normal liver tissue to use further screening.

Fusion gene panel sequencing. The preparation of total RNA was described above. Total RNAs (10 ng) were used as templates to prepare cDNAs using the SuperScript VILO cDNA Synthesis Kit (cat. no. 11754050; Thermo Fisher Scientific, Inc.). The Ion AmpliSeq Library Kit 2.0 (cat. no. 4480442; Thermo Fisher Scientific, Inc.) was used to construct an Ion Torrent adapter-ligated library in accordance with the manufacturer's instructions, and the Ion Proton Sequencing 200 Kit (cat. no. A26433; Thermo Fisher Scientific, Inc.) was used for

nucleotide sequencing in accordance with the manufacturer's protocol. All data were analyzed using the Ion Reporter server. Pre-installed software, Ion AmpliSeq RNA Fusion workflow was used to detect fusion transcripts targeted by the fusion panel.

Fusion gene panel used in this study was custom-made-panel, "HOPE fusion panel" which we designed to amplify 491 fusion transcripts (22). Fusion gene data were obtained from the website of the Sanger Institute COSMIC (15). The COSMIC database v71 includes more than 600 fusion genes. We selected as much of the sequence of the target fusion gene as possible and excluded multiple fusion genes with complex structures (e.g., inversions) with the same breakpoint, ultimately resulting in a selection of 491 fusion genes. The list of panel genes and details are described by Urakami *et al* (22). Primers for detecting fusion transcripts and their 5' and 3' partners were designed using the Ion Ampliseq Designer website (<https://www.ampliseq.com/>; Thermo Fisher Scientific, Inc.). The reference file for this workflow was constructed from all fusion variants in the fusion gene panel. The workflow used a related BED file that describes the breakpoints between the two genes that are associated with each fusion. The reference and BED files are included in the IonReporter Software.

Immunohistochemical analysis. Routine pathological diagnosis was achieved using surgically resected tumors fixed in 10% formalin and embedded in paraffin. Paraffin sections (3 μm thickness) containing representative histology of the tumor were used for immunohistochemical analysis. Immunohistochemical staining was performed using the Bond III automated stainer and BOND Polymer Refine Detection kit (Leica Biosystems). The sections were pretreated with Bond Epitope Retrieval Solution for 20 min at 100°C and then reacted with the primary antibodies (Bond Epitope Retrieval Solution 1 for osteopontin and MGP; Bond Epitope Retrieval Solution 2 for periostin, thrombospondin 2, and GPNMB). After reaction with diaminobenzidine chromogen, the sections were counterstained with hematoxylin. The stained sections were evaluated independently by two investigators (A.S. and T.S.) blinded to patient data. Primary antibodies for osteopontin, periostin, thrombospondin-2, and glycoprotein nonmetastatic melanoma protein B (GPNMB) were obtained from Abcam. Primary antibodies for matrix Gla protein (MGP) were obtained from Thermo Fisher Scientific.

Statistical analysis. Non-parametric variables are reported as medians (range). For comparisons between two groups, χ^2 and Wilcoxon signed rank test were used. When the expected count was under five, Fisher's exact test was used. To compare the equality of variances, one-way analysis of variance was performed to calculate the F-value. The Bonferroni method was implemented to adjust for multiple comparisons. $P < 0.05$ was considered to be significant. All statistical analyses were performed using IBM SPSS Statistics software for Windows, version 24 (IBM; SPSS Inc.).

Data availability. The dataset presented in the current study has been submitted to the National Bioscience Database Center (NBDC) under the accession number hum0127.

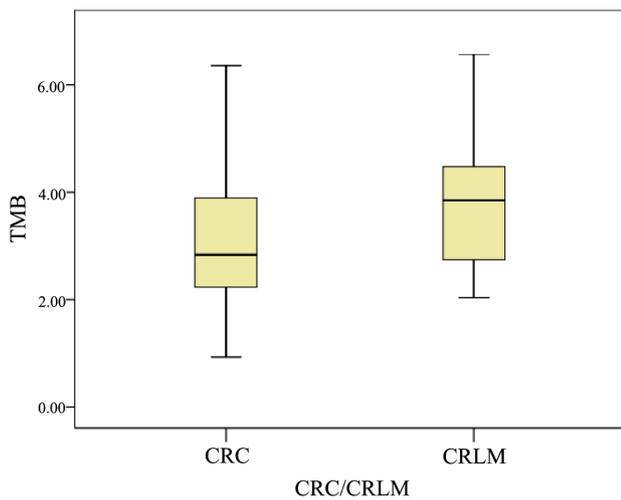


Figure 1. Comparison of tumor mutation burden between primary CRC and CRLM. The TMB counts in CRLM tended to be more frequent than those of primary CRC, although there were no statistically significant differences ($P=0.07$). TMB, tumor mutation burden; CRC, colorectal cancer; CRLM, colorectal liver metastasis.

Results

Patients' characteristics. The data set consisted of 22 pairs of primary CRC and CRLM. The patients' characteristics are shown in Table I. A total of 22 paired patients were enrolled in this study, including 13 (59.1%) colon and 9 (40.9%) rectal cancer patients. Right-sided colon was defined as caecum, ascending, and transverse colon. Left-sided colon was defined as descending and sigmoid colon. All patients had synchronous liver metastasis and did not have any other extrahepatic metastasis. Patients with multiple cancers such as synchronous or metachronous malignancy (within 5 years) other than carcinoma in situ, familial adenomatous polyposis, and appendiceal cancer were excluded. The median number of CRLM of patients was 2 (1-14). The median size of CRLM was 30 (17-110) mm. Of the cases, 21 of 22 (95.5%) received two-stage hepatectomy. The median interval between resection of primary CRC and CRLM was 6 (0-15) weeks. No patients received adjuvant/neoadjuvant chemo/chemoradiotherapy before and after the resection for primary CRC or hepatectomies for CRLM.

Comparison of tumor mutation burden between primary CRC and CRLM. The tumor mutation burden (TMB), also referred to as 'mutation load', represents the number of single nucleotide variations (SNV) per mega base, and has received increasing attention owing to its potential responses to immune checkpoint inhibitors. TMB was determined for the 22 paired samples (Fig. 1). The median TMB in primary CRC was 2.8 (0.9-6.4), whereas it was 3.8 (2.0-6.6) in CRLM. The TMB counts in CRLM tended to be more frequent than those of primary CRC, although there were no statistically significant differences ($P=0.07$).

Gene mutation profiling. Genomic alterations contributing to tumorigenesis in primary CRC and CRLM were analyzed using an analysis pipeline called 'Shizuoka Multi-omics Analysis

Table I. Patient characteristics (n=22).

Variable	Value
Sex	
Male	12
Female	10
Age, years	64.5 (27-86) ^a
<50	3
≥50	19
Primary tumor site	
Colon	13
Right side	5
Left side	8
Rectum	9
CEA (ng/ml)	24.7 (1.8-1421) ^a
≤5.0	7
>5.0	15
Histological type of primary site	
Differentiated	21
Undifferentiated	1
pT stage	
pT1	0
pT2	1
pT3	9
pT4	12
pN stage	
pN0	3
pN1	9
pN2	10
Synchronous presentation of CRLM	
Absent	22
Present	0
Lymphatic invasion	
Absent	12
Present	10
Venous invasion	
Absent	4
Present	18
Median number of CRLM	2 (1-14)
Largest size of CRLM, cm	30 (17-110) ^a
Multiple CRLM	
Absent	9
Present	13
Extrahepatic disease	
Absent	22
Present	0
Neoadjuvant chemo(radio)therapy	
Absent	22
Present	0

^aValues for age, CEA, and largest size of CRLM are provided as the median (range). CEA, carcinoembryonic antigen; CRLM, colorectal liver metastasis.

Table II. Genes whose expression are higher in CRLM through DNA microarray analysis.

Gene symbol	Gene name	Location	Variance		Bonferroni
			F-value	P-value	<0.05/20=0.0025
SPP1	Osteopontin	ECM	558.088	<0.0001	0.0025
TIMP1	Tissue inhibitor of metalloproteinase 1	Secreted	489.451	<0.0001	0.0025
THBS2	Thrombospondin-2	ECM	258.758	<0.0001	0.0025
POSTN	periostin	ECM	248.512	<0.0001	0.0025
GPNMB	Glycoprotein nonmetastatic melanoma protein B	Membrane (with ECM domain)	245.809	<0.0001	0.0025
MGP_2	Matrix Gla protein	ECM	198.335	<0.0001	0.0025
MGP	Matrix Gla protein	ECM	191.511	<0.0001	0.0025
COL10A1	Collagen type X α 1 chain	ECM	130.898	0.0008	0.0025
CCDC80	Coiled-coil domain containing 80	Cytosol	125.409	0.001	0.0025
CCDC146	Coiled-coil domain containing 146	Nucleus	113.416	0.0016	0.0025
CCL18	Chemokine (C-C motif) ligand 18	Secreted	104.601	0.0024	0.0025

ECM, extracellular matrix.

Protocol (SMAP) (10)'. Detected genomic alterations and number of cases are shown in Fig. 2A. In primary CRC, the alteration of TP53, APC, and KRAS was detected in 16 (72.7%), 14 (63.6%), and 9 (40.9%) cases, respectively. In CRLM, the alteration of TP53, APC, and KRAS was detected in 16 (72.7%), 14 (63.6%), 9 (40.9%) cases, respectively.

Concordance between mutations in matched pairs of primary CRC and CRLM. The concordance between genomic alterations in paired primary CRC and CRLM varied from 16.8 to 86.8% in each paired case (Fig. 2B). The concordance between the sequence variation in KRAS, which is known to be an important factor related to tumor metastasis was 100%. On the other hand, the concordance in APC was only 66.7% (Fig. 2C).

Gene expression analysis using DNA microarrays. To select candidate genes that showed significantly higher expression in CRLM than primary CRC, gene expression was compared using the average expression value of 36,022 probes in primary CRC, CRLM, and normal liver tissues (i.e., non-cancerous tissues of liver metastasis cases). First, we selected 1,794 probes whose expression was up to 1.5 times higher in CRLM than in primary CRC. Second, to remove contamination of normal liver tissues in liver metastasis samples, we compared expression levels in primary CRC with those in non-cancerous liver tissues and selected 198 probes whose expression was higher in primary CRC than in normal liver. Third, we selected 20 probes (19 genes) as highly expressed probes in CRLM with a normalized ratio >1.0. Finally, we examined the difference in expression level of these 20 probes in paired samples of CRC and CRLM, and selected 11 probes (10 genes) that showed significantly higher expression in CRLM compared to CRC (Table II; Fig. 3). Fig. 4A shows the characteristic expression profiles of the 11 probes in the 22 paired cases. Fig. 4B shows a comparison in gene expression between primary CRC and CRLM; all 11 probes showed significantly higher expression in CRLM (Wilcoxon signed rank test, $P < 0.001$).

Immunohistochemical analysis. Immunohistochemical analysis was performed to confirm the protein expression of genes identified by DNA microarray analysis. Among the 10 genes that were highly expressed in CRLM, the five that were classified as encoding 'matricellular proteins', i.e., osteopontin, periostin, thrombospondin-2, MGP, and GPNMB, were selected for this analysis. GPNMB is a type I transmembrane protein with three unique domains: an extracellular domain, a transmembrane domain, and a cytoplasmic domain. The extracellular domain is composed of two regions with distinct properties: the integrin-binding motif and the polycystic kidney disease domain. Thus, we included GPNMB as a matricellular protein in this study. Representative images of immunohistochemical staining are shown in Fig. 5. First, immunoreactive scores were defined independently according to the intensity of staining and the proportion of stained structures. Staining intensity was scored as: 0, no staining; 1, weak staining; and 2, strong staining. Proportions of stained tumor cells/extracellular matrix (ECM) were classified as: 0, $\leq 5\%$ positive cells; 1, 6-25% positive cells; 2, 26-50% positive cells; and 3, $\geq 51\%$ positive cells. Total scores for intensity and proportion were used to signify the levels of protein expression. In this study, a score of ≤ 3 was considered to represent negative expression, and a score of ≥ 4 was considered to represent positive expression. Table III shows the positive immunoreactivity scores for these five proteins.

Osteopontin showed strong immunoreactivity in tumor cells and ECM of CRLM (6 of 22 cases: 27.3%) compared to no immunoreactivity in primary CRC (0 of 22 cases: 0.0%). The immunoreactivity score was significantly higher in CRLM than in primary CRC ($P = 0.02$).

Periostin also showed strong immunoreactivity in the ECM of CRLM (17 of 22 cases: 68.2%) compared to that in primary CRC (7 of 22 cases: 31.8%). Periostin showed no immunoreactivity in the tumor cells of primary CRC and CRLM. Periostin was not expressed with high frequency in primary CRC. However, when it was expressed, it showed

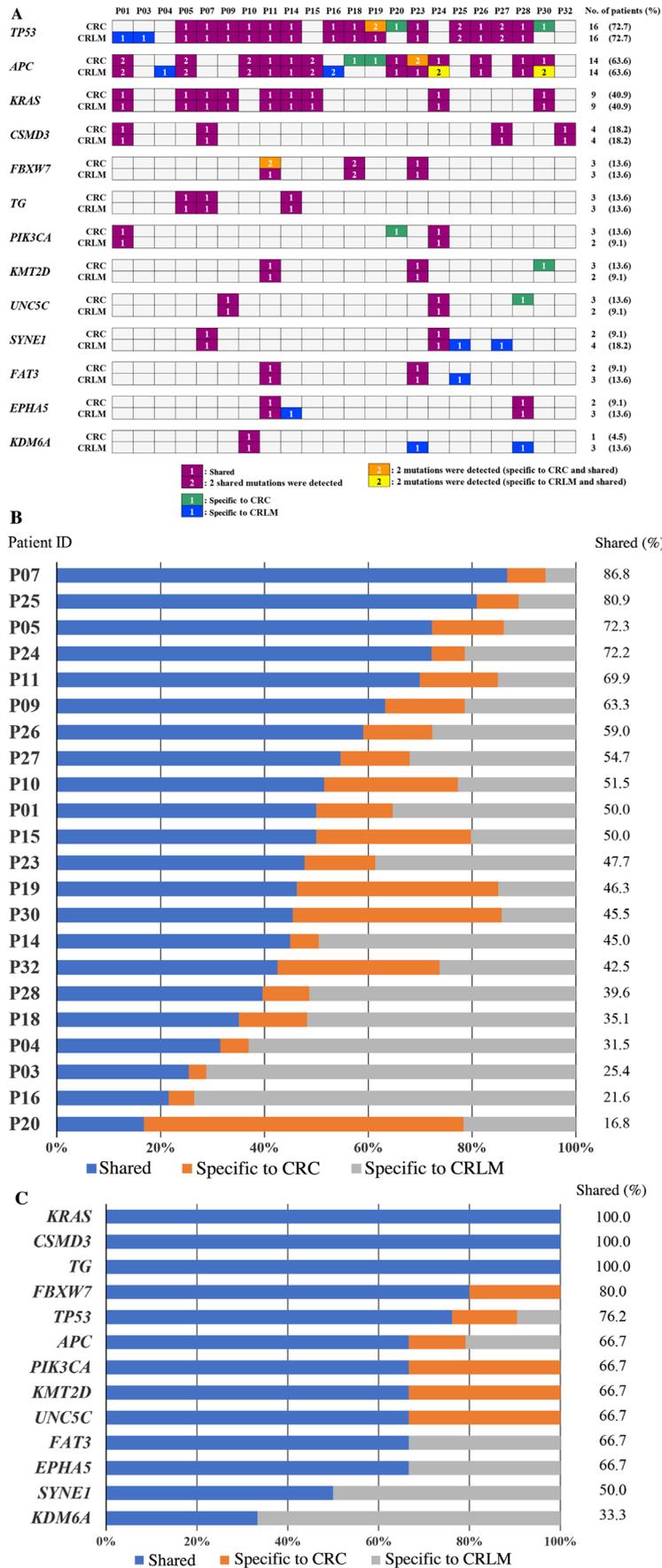


Figure 2. Concordance between genomic mutations in matched pairs of primary CRC and CRLM. (A) Gene mutation profiling of primary CRC and CRLM. (B) Concordance between genomic alterations in paired primary CRC and CRLM. The concordance between genomic alterations in primary CRC and CRLM ranged from 16.8 to 86.8% in each paired case. (C) Concordance between sequence variations in representative genes in paired primary CRC and CRLM. The concordance between the sequence variation in *KRAS* was 100%. On the other hand, the concordance between the sequence variation in *APC* was only 66.7%. CRC, colorectal cancer; CRLM, colorectal liver metastasis.

Table III. Expression status of each factor by immunostaining.

Protein	Tumor type	Positive		Negative		P-value
		n	%	n	%	
Osteopontin	Primary CRC	0	0	22	100	0.02
	CRLM	6	27.3	16	72.7	
Periostin	Primary CRC	7	31.8	15	68.2	0.006
	CRLM	17	68.2	5	31.8	
Thrombospondin-2	Primary CRC	12	54.5	10	45.5	0.55
	CRLM	10	45.5	12	54.5	
GPNMB	Primary CRC	4	18.2	18	81.8	1
	CRLM	5	22.7	17	77.3	
MGP	Primary CRC	3	13.6	19	86.4	1
	CRLM	4	18.2	18	81.8	

CRC, colorectal cancer; CRLM, colorectal liver metastasis.

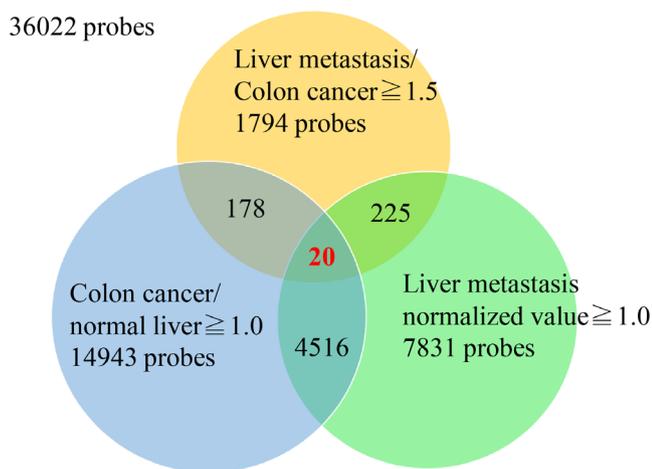


Figure 3. Selection of candidate genes showing significantly higher expression in CRLM than CRC. From 36,022 probes, 20 probes were chosen as highly expressed probes in CRLM. CRC, colorectal cancer; CRLM, colorectal liver metastasis.

strong immunoreactivity in the peripheral invasive part of primary lesions.

Thrombospondin-2 showed strong immunoreactivity in tumor cells and ECM of both primary CRC and CRLM, and there were no statistically significant differences (P=0.55).

GPNMB showed weak immunoreactivity in tumor cells, but strong immunoreactivity in ECM of both primary CRC and CRLM. These expression patterns were also observed in macrophages, but no statistically significant differences were observed.

MGP showed strong immunoreactivity in ECM of primary CRC (3 of 22 cases: 13.6% and CRLM (4 of 22 cases: 18.2%), but no statistically significant differences were observed.

Comparison of number of immunoreactive factors. To determine whether differences in the complexity of expression patterns between primary CRC and CRLM could be verified, and to determine genomic heterogeneity, we compared the

number of positive immunoreactive factors for 22 pairs of primary CRC and CRLM (Table IV). Sixteen cases (72.7%) of primary CRC showed either no or one positive immunoreactive factor, and six cases (27.3%) were positive for two or more immunoreactive factors. On the other hand, immunoreactivity was more frequently observed in CRLM, with 16 cases (72.7%) sharing two or more factors, which was significantly more than that observed in primary CRC (P=0.007).

Discussion

In the present study, we demonstrated the existence of genomic heterogeneity between paired primary CRC and CRLM by using a large complement of exhaustive genetic analyses with next-generation sequencing. To the best of our knowledge, no other study has analyzed such a completely paired sample of primary CRC and synchronous CRLM. Elucidation of the heterogeneity of microenvironment-related factors on the proliferation, invasion, and metastasis of cancer cells will lead to novel diagnostic and therapeutic targets for CRC in the era of genome-guided personalized cancer treatment.

Recent advances in next-generation sequencing technologies have made it possible to analyze large numbers of sequences, leading to international cancer genome analysis projects such as the Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC). CRC can be classified into four gene expression-based subtypes with distinguishing features, the consensus molecular subtypes (CMSs): CMS1 (microsatellite instability immune, 14%), CMS2 (canonical, 37%), CMS3 (metabolic, 13%), and CMS4 (mesenchymal, 23%) (23). This intertumoral heterogeneity has led to the finding that different CRC subtypes have a different genetic makeup, clinical behavior, pathological features, and responses to treatment (24-26).

In addition to the intertumoral heterogeneity mentioned above, intratumoral heterogeneity relates to the genetic heterogeneity between cancer cells within a single tumor. During carcinogenesis, genetic abnormalities accumulate continually,

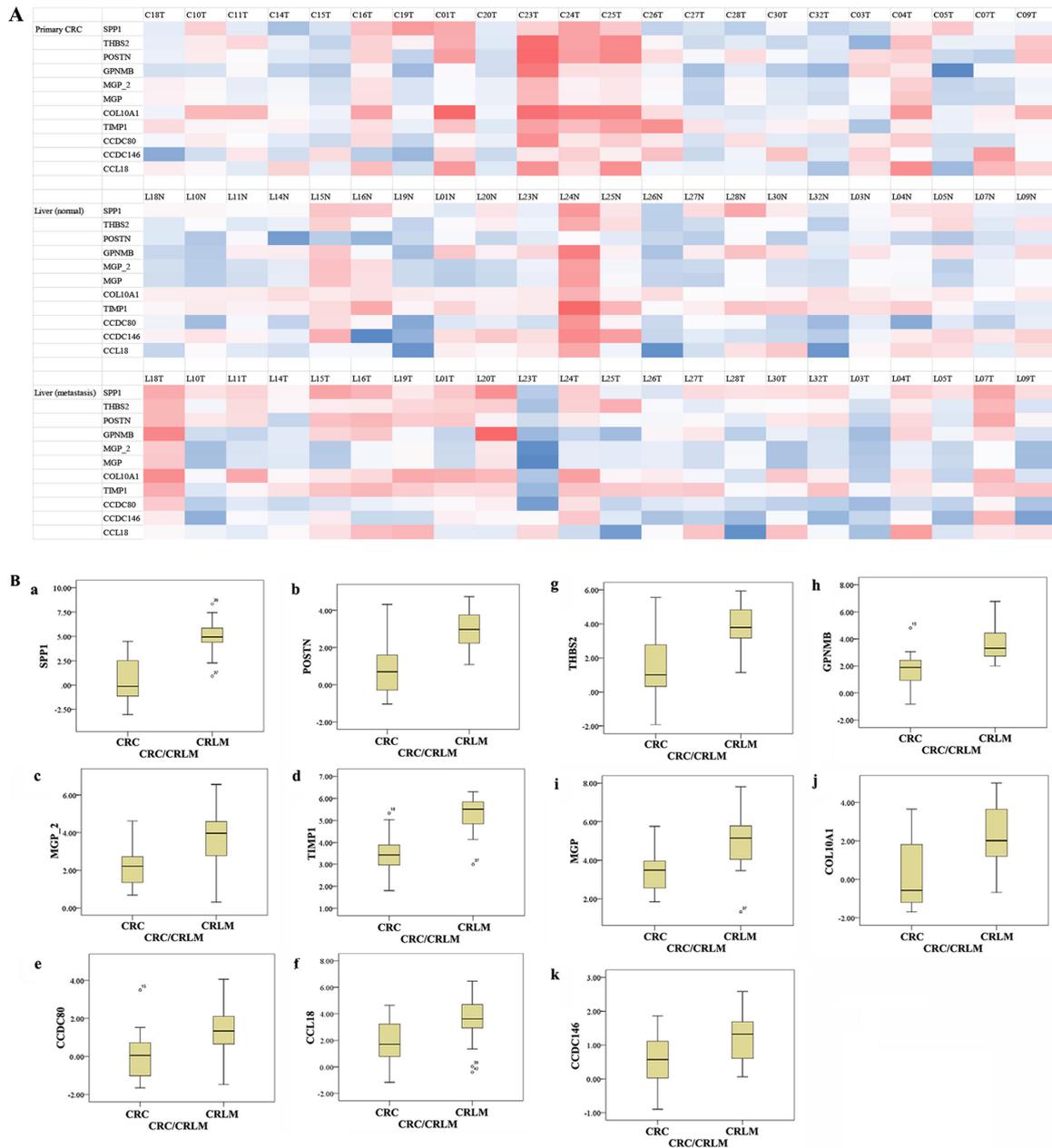


Figure 4. (A) Heat map of the expression profiles from the microarray analysis. Three heat maps representing gene expression levels (z-scores) in CRC (upper), CRLM (middle) and normal liver tissue (lower), respectively. Raw signal intensity values of DNA microarray were log transformed and normalized to the 75th percentile. We transformed those normalized intensity value into z-scores per probe according to the formula: $Z\text{-score}=(x-\alpha)/\beta$. (x : normalized intensity of the selected probe in each sample, α : mean of the normalized values of 22 samples in the selected probe, β : standard deviation of the normalized values of 22 samples in the selected probe). Resultant z-score in each sample which are higher or lower than the mean of normalized intensity in each probe are displayed as positive value (red) or negative value (blue), respectively. (B) Comparison of gene expression between primary CRC and CRLM. (B-a) SPP1, (B-b) POSTN, (B-c) MGP2, (B-d) TIMP1, (B-e) CCDC80, (B-f) CCL18, (B-g) THBS2, (B-h) GPNMB, (B-i) MGP, (B-j) COL10A1 and (B-k) CCDC146. All 11 probes showed significantly higher expression in CRLM than primary CRC ($P < 0.001$). CRC, colorectal cancer; CRLM, colorectal liver metastasis.

Table IV. Comparison of positive immunoreactivity between primary CRC and CRLM.

Number of factors	Primary CRC (n=22)	CRLM (n=22)	P-value
0 or 1 factor	16 (72.7%)	6 (27.3%)	0.007
2 or more factors	6 (27.3%)	16 (72.7%)	

CRC, colorectal cancer; CRLM, colorectal liver metastasis.

allowing the cells an increased ability to expand and invade. As a result of this continuous process, cancers become genetically heterogeneous, with an indeterminate number of coexisting genomic clones. These clones have different functional characteristics such as the ability to form metastases or respond to chemotherapy.

Cancer cells survive and proliferate in a microenvironment created by the cells themselves, various stromal cells, and the stromal tissue. The stromal cells that form cancer tissues include fibroblasts, vascular and lymphangial endothelial cells, lymphocytes, and macrophages. Both cells interact with cancer

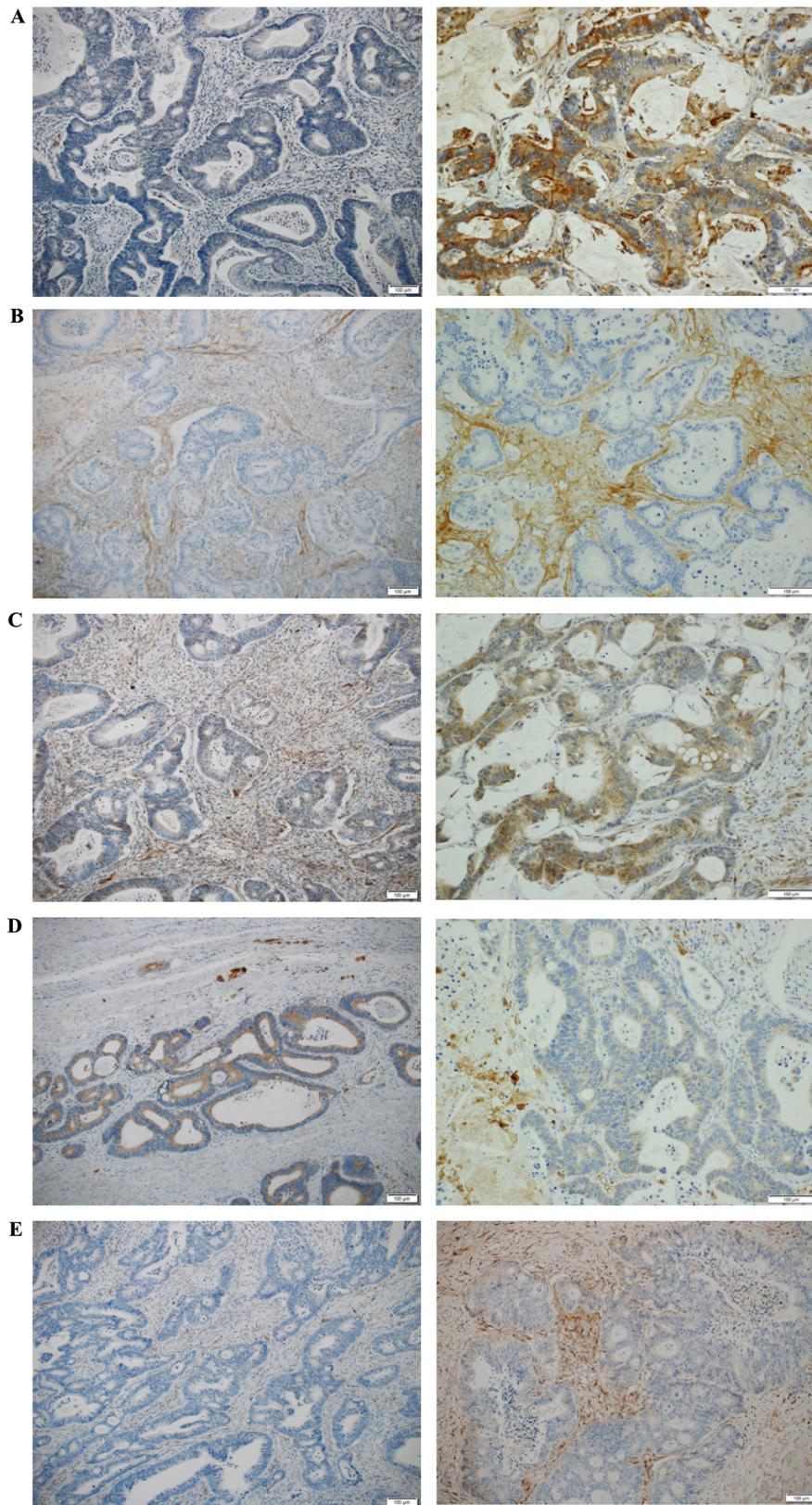


Figure 5. Representative images of immunohistochemical staining in paired CRC and CRLM. (A) Osteopontin, (B) periostin, (C) thrombospondin-2, (D) GPNMB and (E) MGP. Pictures lined in the left side are representative images of immunohistochemical staining captured from CRC. Representative images of immunohistochemical staining captured from paired CRLM are lined in the right side of this figure. Osteopontin and periostin showed statistically higher levels of immunoreactivity in CRLM compared to CRC. CRC, colorectal cancer; CRLM, colorectal liver metastasis; GPNMB, glycoprotein nonmetastatic melanoma protein B; MGP, matrix Gla protein.

cells, imparting their characteristic biological features on the cancer (27,28). Distant metastasis of cancer has also been

implicated in interactions between cancer cells and stromal cells in the cancer microenvironment (8,9).

In this study, using the DNA microarray analysis, we determined that many highly expressed genes were classified as encoding 'matricellular proteins', which interact with the ECM. Periostin, a secreted adhesion-related protein that is expressed in the periosteum and periodontal ligaments, also acts as a critical regulator in the formation and maintenance of bone and teeth, as well as playing an important role in tumorigenesis (29). Recent studies have shown that periostin is highly expressed in various human cancers, and it has been suggested that periostin promotes tumor growth and metastasis (30-32). Moreover, periostin has been reported to enhance the metastatic growth of colon cancer by both preventing stress-induced apoptosis in cancer cells and augmenting endothelial cell survival to promote angiogenesis. The expression level of periostin in metastatic tumors is reported to be noticeably higher than that in the matched primary colon cancer (33).

Osteopontin is a multifunctional ECM phosphorylated glycoprotein (glycol-phosphoprotein) that belongs to the Small Integrin-Binding Ligand N-linked Glycoprotein (SIBLING) family, and is reported to play an important role in the tumorigenesis, progression and prognosis of various cancers by regulating cell-matrix interactions and cell signaling through binding with integrins and CD44 receptors (34-39). A pooled data analysis showed that high osteopontin expression was significantly associated with high tumor grade, invasion, lymph node metastasis, tumor distant metastasis and poor survival in CRC (40).

Thrombospondin-2 (THBS2) is a member of the ECM glycoproteins that mediate ECM assembly, cell-matrix interactions, degradation of matrix metalloproteinases (MMP)-2 and MMP-9, and interact with multiple cell receptors and growth factors. The implication of THBS2 expression in CRC has been controversial. Several studies reported an inverse correlation between THBS2 expression level and malignancy grade (41,42). In contrast, recent studies reported that THBS2 expression in CRC was positively correlated with TNM stage and is a strong prognostic indicator (43,44).

The GPNMB gene is reported to be overexpressed in numerous cancers and is often associated with the metastatic phenotype (45-49). The extracellular domain of GPNMB interacts with integrins to facilitate the recruitment of immune-suppressive and proangiogenic cells to the tumor microenvironment, thereby enhancing tumor migration and invasion (50). GPNMB expressed in immune cells such as macrophages and dendritic cells (14,29) may impair T-cell activation to down-modulate anti-tumor immune responses (51-53). However, the role of GPNMB is complex; it appears to have an inhibitory role in some cancers, but may promote metastasis in others.

MGP is an ECM protein containing post-translationally modified γ -carboxylglutamate residues due to vitamin K-dependent carboxylation. MGP was initially thought to be involved in the inhibition of calcification of arteries and cartilage. Further investigation demonstrated that MGP had a wider range of activities, which were dependent upon phosphorylation-carboxylation status, protein expression and variants. Recent studies showed that MGP plays a role in tumor angiogenesis by increasing vascular endothelial growth factor gene expression (54-56). Recently, MGP was reported to be upregulated in a variety of tumors, including ovarian,

breast, urogenital and skin cancer. However, in colon and lung cancers, an inverse correlation between MGP expression and survival was observed (57).

In the present study, exhaustive genetic analysis using next-generation sequencing and a comparison of immunoreactive factors revealed the complexities of gene expression in CRLM. It is especially notable that compared to primary CRC, CRLM has greater genomic heterogeneity associated with the ECM. This result corroborates our hypotheses. To proliferate in the liver, which differs environmentally from the original colorectal tissue in which they exist naturally, cancer cells must modify their microenvironment to make it more amenable for survival. A suitable microenvironment cannot be regulated by a single factor; instead, complex factors are involved, especially in metastatic sites. The complexity of intratumor heterogeneity, which we revealed in this study, may be an underlying cause of the resistance to treatment for metastatic disease.

Although differences in genomic mutational profiling between primary and metastatic sites have been reported in some studies (58-60), most studies have used metachronous or unpaired patient samples. Moreover, a limited set of biomarkers was employed to show that metastatic sites have more inherited mutations than primary CRC. These studies have potential biases related to both intra- and intertumoral heterogeneity.

It has been reported that the several kinds of systemic chemotherapy could alter the genomic landscape in some cancers (61,62). Thus, adjuvant chemotherapy has the potential to alter the genomic mutational profiles of recurrent cancers, changing them from those of the original primary tumors (63). Previous analyses evaluating the heterogeneity of metachronous tumors could be biased by the administration of chemotherapy between the resection of primary CRC and metastatic sites, potentially producing alterations in genomic clones (58,64). Our study eliminates this bias since we selected paired samples of synchronous tumors.

This study has two important limitations. First, the study included a relatively small number of patients. We plan to continue our analysis using an increased number of cases in the future. Next, although the gene expression of thrombospondin-2, GPNMB, and MGP in CRLM was more frequent than in primary CRC according to the DNA microarray analysis, the immunohistochemical analysis revealed no differences in expression. We consider that the discrepancy between the tissue regions analyzed by DNA microarray and those subjected to immunohistochemical analysis led to this result. The relatively low immunoreactivity scores of GPNMB and MGP may be attributed to the timing of degradation and wash-out of these proteins from the tissue (56).

There are three future perspectives from this study. First, we also obtained data of surgically resected tumor specimens and corresponding peripheral blood samples not only for CRLM, but also pulmonary metastasis, peritoneal metastasis, and ovarian metastasis through WES, cancer gene panel sequencing, fusion gene panel sequencing and microarray-based GEP under the framework of 'project HOPE'. We will therefore be able to perform further investigations using these samples. Second, studies of circulating tumor cells (liquid biopsy) may be useful for establishing

early CRLM diagnosis. It may also lead to better predictive biomarkers to identify patients who might benefit from adjuvant chemotherapy. Third, the genes that were overexpressed in this study have the potential to be new therapeutic targets. For example, the administration of anti-periostin antibody significantly inhibited the growth of primary tumors as well as metastatic tumors in a murine model of breast cancer (29). Osteopontin-inhibition is also reported to be a favorable therapeutic approach to metastatic disease (65-68). An antibody-drug conjugate targeting GPNMB, called glembatumumab vedotin (CDX-011), is currently being assessed in clinical studies for various cancers (69).

In conclusion, we examined the genomic heterogeneity between paired primary CRC and CRLM in terms of the microenvironment on the proliferation, invasion, and metastasis of cancer cells. These findings will lead to new diagnostic and therapeutic targets for CRC in the era of genome-guided personalized cancer treatment.

Acknowledgements

The authors would like to thank Mr. Koji Muramatsu (Division of Pathology, Shizuoka Cancer Center) for his technical pathological assistance.

Funding

This research was funded by the prefectural government of Shizuoka Prefecture, Japan.

Availability of data and materials

The dataset presented in the current study has been submitted to the National Bioscience Database Center (NBDC) under the accession number hum0127 (<https://humandbs.biosciencedbc.jp/en/hum0127-v2>).

Authors' contributions

AS, MK, TaS, TeS, KO, TN, KU, MS, HS and KY designed this retrospective study. AS and TeS enrolled patients to Project HOPE. AS, MK, TaS, MS and TN analyzed and interpreted the data. AS and MK confirmed the authenticity of all the raw data. All authors drafted the manuscript, revised the paper, and read and approved the final manuscript.

Ethics approval and consent to participate

The research plan of Project HOPE was designed according to the revised Ethical Guidelines for Human Genome/Gene Analysis Research in Japan (http://www.lifescience.next.go.jp/files/pdf/n1115_01.pdf), and the study protocol was approved by the Institutional Review Board at the Shizuoka Cancer Center (authorization no. 30-5). This retrospective study was also approved by the same board. Written informed consent was obtained from all participants.

Patient consent for publication

Consent for publication was obtained from all participants.

Competing interests

The authors declare that they have no competing interest.

References

1. DeSantis CE, Lin CC, Mariotto AB, Siegel RL, Stein KD, Kramer JL, Alteri R, Robbins AS and Jemal A: Cancer treatment and survivorship statistics, 2014. *CA Cancer J Clin* 64: 252-271, 2014.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A: 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. Hashiguchi Y, Muro K, Saito Y, Ito Y, Ajioka Y, Hamaguchi T, Hasegawa K, Hotta K, Ishida H, Ishiguro M, *et al*; Japanese Society for Cancer of the Colon and Rectum: Japanese Society for Cancer of the Colon and Rectum (JSCCR) guidelines 2019 for the treatment of colorectal cancer. *Int J Clin Oncol* 25: 1-42, 2020.
4. Folprecht G, Gruenberger T, Bechstein WO, Raab HR, Lordick F, Hartmann JT, Lang H, Frilling A, Stoecklacher J, Weitz J, *et al*: Tumour response and secondary resectability of colorectal liver metastases following neoadjuvant chemotherapy with cetuximab: The CELIM randomised phase 2 trial. *Lancet Oncol* 11: 38-47, 2010.
5. Majeed AW: Surgery for colorectal liver metastases with hepatic lymph node involvement: A systematic review. *Br J Surg* 87: 1737, 2000.
6. Martin LW and Warren RS: Current management of colorectal liver metastases. *Surg Oncol Clin N Am* 9: 853-878, 2000.
7. Penna C and Nordlinger B: Colorectal metastasis (liver and lung). *Surg Clin North Am* 82: 1075-1090, x-xi, 2002.
8. Fidler IJ and Kripke ML: Metastasis results from preexisting variant cells within a malignant tumor. *Science* 197: 893-895, 1977.
9. Futakuchi M, Nannuru KC, Varney ML, Sadanandam A, Nakao K, Asai K, Shirai T, Sato SY and Singh RK: Transforming growth factor-beta signaling at the tumor-bone interface promotes mammary tumor growth and osteoclast activation. *Cancer Sci* 100: 71-81, 2009.
10. Nagashima T, Yamaguchi K, Urakami K, Shimoda Y, Ohnami S, Ohshima K, Tanabe T, Naruoka A, Kamada F, Serizawa M, *et al*: Japanese version of The Cancer Genome Atlas, JCGA, established using fresh frozen tumors obtained from 5143 cancer patients. *Cancer Sci* 111: 687-699, 2020.
11. Japanese ethical guidelines for human genome/gene analysis research. <https://www.mhlw.go.jp/general/seido/kousei/i-kenkyu/genome/0504sisin.html>. Accessed: March 25, 2021.
12. Ion Reporter Software User Guide: Tumor-Normal pair workflow. https://tools.thermofisher.com/content/sfs/manuals/IonReporter_v50_Help.pdf. Accessed: March 25, 2021.
13. Torrent Suite v4.4.3 User and Admin Guide. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0019144_TorrentSuite_5_14_UG.pdf. Accessed: March 25, 2021.
14. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G and Mesirov JP: Integrative genomics viewer. *Nat Biotechnol* 29: 24-26, 2011.
15. Bamford S, Dawson E, Forbes S, Clements J, Pettett R, Dogan A, Flanagan A, Teague J, Futreal PA, Stratton MR, *et al*: The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer* 91: 355-358, 2004.
16. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM and Maglott DR: ClinVar: Public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* 42: D980-D985, 2014.
17. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM and Sirotkin K: dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res* : 308-311, 2001.
18. UniProt Consortium: UniProt: A hub for protein information. *Nucleic Acids Res* 43: D204-D212, 2015.
19. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, Chang Z and Woolsey J: DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* 34: D668-D672, 2006.
20. Nagashima T, Shimoda Y, Tanabe T, Naruoka A, Saito J, Serizawa M, Ohshima K, Urakami K, Ohnami S, Ohnami S, *et al*: Optimizing an ion semiconductor sequencing data analysis method to identify somatic mutations in the genomes of cancer cells in clinical tissue samples. *Biomed Res* 37: 359-366, 2016.

21. Shimoda Y, Nagashima T, Urakami K, Tanabe T, Saito J, Naruoka A, Serizawa M, Mochizuki T, Ohshima K, Ohnami S, *et al*: Integrated next-generation sequencing analysis of whole exome and 409 cancer-related genes. *Biomed Res* 37: 367-379, 2016.
22. Urakami K, Shimoda Y, Ohshima K, Nagashima T, Serizawa M, Tanabe T, Saito J, Usui T, Watanabe Y, Naruoka A, *et al*: Next generation sequencing approach for detecting 491 fusion genes from human cancer. *Biomed Res* 37: 51-62, 2016.
23. Guinney J, Dienstmann R, Wang X, de Reyniès A, Schlicker A, Sonesson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, *et al*: The consensus molecular subtypes of colorectal cancer. *Nat Med* 21: 1350-1356, 2015.
24. Sadanandam A, Lyssiotis CA, Homicsko K, Collisson EA, Gibb WJ, Wullschlegler S, Ostos LC, Lannon WA, Grotzinger C, Del Rio M, *et al*: A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat Med* 19: 619-625, 2013.
25. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, *et al*: PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med* 372: 2509-2520, 2015.
26. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, *et al*: Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357: 409-413, 2017.
27. Polanska UM and Orimo A: Carcinoma-associated fibroblasts: Non-neoplastic tumour-promoting mesenchymal cells. *J Cell Physiol* 228: 1651-1657, 2013.
28. Ishii G, Ochiai A and Neri S: Phenotypic and functional heterogeneity of cancer-associated fibroblast within the tumor microenvironment. *Adv Drug Deliv Rev* 99: 186-196, 2016.
29. Kyutoku M, Taniyama Y, Katsuragi N, Shimizu H, Kunugiza Y, Iekushi K, Koibuchi N, Sanada F, Oshita Y and Morishita R: Role of periostin in cancer progression and metastasis: Inhibition of breast cancer progression and metastasis by anti-periostin antibody in a murine model. *Int J Mol Med* 28: 181-186, 2011.
30. Erkan M, Kleeff J, Gorbachevski A, Reiser C, Mitkus T, Esposito I, Giese T, Büchler MW, Giese NA and Friess H: Periostin creates a tumor-supportive microenvironment in the pancreas by sustaining fibrogenic stellate cell activity. *Gastroenterology* 132: 1447-1464, 2007.
31. Kudo Y, Ogawa I, Kitajima S, Kitagawa M, Kawai H, Gaffney PM, Miyachi M and Takata T: Periostin promotes invasion and anchorage-independent growth in the metastatic process of head and neck cancer. *Cancer Res* 66: 6928-6935, 2006.
32. Puglisi F, Puppini C, Pegolo E, Andreatta C, Pascoletti G, D'Aurizio F, Pandolfi M, Fasola G, Piga A, Damante G, *et al*: Expression of periostin in human breast cancer. *J Clin Pathol* 61: 494-498, 2008.
33. Moniuszko T, Wincewicz A, Koda M, Domysławska I and Sulkowski S: Role of periostin in esophageal, gastric and colon cancer. *Oncol Lett* 12: 783-787, 2016.
34. Sodek J, Ganss B and McKee MD: Osteopontin. *Crit Rev Oral Biol Med* 11: 279-303, 2000.
35. Zhang J, Takahashi K, Takahashi F, Shimizu K, Ohshita F, Kameda Y, Maeda K, Nishio K and Fukuchi Y: Differential osteopontin expression in lung cancer. *Cancer Lett* 171: 215-222, 2001.
36. Song G, Cai QF, Mao YB, Ming YL, Bao SD and Ouyang GL: Osteopontin promotes ovarian cancer progression and cell survival and increases HIF-1 α expression through the PI3-K/Akt pathway. *Cancer Sci* 99: 1901-1907, 2008.
37. Kim JY, Bae BN, Kim KS, Shin E and Park K: Osteopontin, CD44, and NF κ B expression in gastric adenocarcinoma. *Cancer Res Treat* 41: 29-35, 2009.
38. Likui W, Hong W and Shuwen Z: Clinical significance of the upregulated osteopontin mRNA expression in human colorectal cancer. *J Gastrointest Surg* 14: 74-81, 2010.
39. Bramwell VH, Tuck AB, Chapman JA, Anborgh PH, Postenka CO, Al-Katib W, Shepherd LE, Han L, Wilson CF, Pritchard KI, *et al*: Assessment of osteopontin in early breast cancer: Correlative study in a randomised clinical trial. *Breast Cancer Res* 16: R8, 2014.
40. Zhao M, Liang F, Zhang B, Yan W and Zhang J: The impact of osteopontin on prognosis and clinicopathology of colorectal cancer patients: A systematic meta-analysis. *Sci Rep* 5: 12713, 2015.
41. Tokunaga T, Nakamura M, Oshika Y, Abe Y, Ozeki Y, Fukushima Y, Hatanaka H, Sadahiro S, Kijima H, Tsuchida T, *et al*: Thrombospondin 2 expression is correlated with inhibition of angiogenesis and metastasis of colon cancer. *Br J Cancer* 79: 354-359, 1999.
42. de Fraipont F, Nicholson AC, Feige JJ and Van Meir EG: Thrombospondins and tumor angiogenesis. *Trends Mol Med* 7: 401-407, 2001.
43. Kim H, Watkinson J, Varadan V and Anastassiou D: Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving INHBA, THBS2 and COL11A1. *BMC Med Genomics* 3: 51, 2010.
44. Tian Q, Liu Y, Zhang Y, Song Z, Yang J, Zhang J, Guo T, Gao W, Dai F and He C: THBS2 is a biomarker for AJCC stages and a strong prognostic indicator in colorectal cancer. *J BUON* 23: 1331-1336, 2018.
45. Rich JN, Shi Q, Hjelmeland M, Cummings TJ, Kuan CT, Bigner DD, Counter CM and Wang XF: Bone-related genes expressed in advanced malignancies induce invasion and metastasis in a genetically defined human cancer model. *J Biol Chem* 278: 15951-15957, 2003.
46. Kuan CT, Wakiya K, Dowell JM, Herndon JE II, Reardon DA, Graner MW, Riggins GJ, Wikstrand CJ and Bigner DD: Glycoprotein nonmetastatic melanoma protein B, a potential molecular therapeutic target in patients with glioblastoma multiforme. *Clin Cancer Res* 12: 1970-1982, 2006.
47. Rose AA, Pepin F, Russo C, Abou Khalil JE, Hallett M and Siegel PM: Osteoactivin promotes breast cancer metastasis to bone. *Mol Cancer Res* 5: 1001-1014, 2007.
48. Rose AA, Annis MG, Dong Z, Pepin F, Hallett M, Park M and Siegel PM: ADAM10 releases a soluble form of the GPNMB/Osteoactivin extracellular domain with angiogenic properties. *PLoS One* 5: e12093, 2010.
49. Rose AA, Grosset AA, Dong Z, Russo C, Macdonald PA, Bertos NR, St-Pierre Y, Simantov R, Hallett M, Park M, *et al*: Glycoprotein nonmetastatic B is an independent prognostic indicator of recurrence and a novel therapeutic target in breast cancer. *Clin Cancer Res* 16: 2147-2156, 2010.
50. Taya M and Hammes SR: Glycoprotein non-metastatic melanoma protein B (GPNMB) and cancer: A novel potential therapeutic target. *Steroids* 133: 102-107, 2018.
51. Zhou LT, Liu FY, Li Y, Peng YM, Liu YH and Li J: Gpnmb/osteoactivin, an attractive target in cancer immunotherapy. *Neoplasia* 59: 1-5, 2012.
52. Selim AA: Osteoactivin bioinformatic analysis: Prediction of novel functions, structural features, and modes of action. *Med Sci Monit* 15: MT19-MT33, 2009.
53. Singh M, Del Carpio-Cano F, Belcher JY, Crawford K, Frara N, Owen TA, Popoff SN and Safadi FF: Functional roles of osteoactivin in normal and disease processes. *Crit Rev Eukaryot Gene Expr* 20: 341-357, 2010.
54. Mertsch S, Schurgers LJ, Weber K, Paulus W and Senner V: Matrix gla protein (MGP): An overexpressed and migration-promoting mesenchymal component in glioblastoma. *BMC Cancer* 9: 302, 2009.
55. Kuzontkoski PM, Mulligan-Kehoe MJ, Harris BT and Israel MA: Inhibitor of DNA binding-4 promotes angiogenesis and growth of glioblastoma multiforme by elevating matrix GLA levels. *Oncogene* 29: 3793-3802, 2010.
56. Gheorghe SR and Crăciun AM: Matrix Gla protein in tumoral pathology. *Clujul Med* 89: 319-321, 2016.
57. Caiado H, Conceição N, Tiago D, Marreiros A, Vicente S, Enriquez JL, Vaz AM, Antunes A, Guerreiro H, Caldeira P, *et al*: Evaluation of MGP gene expression in colorectal cancer. *Gene* 723: 144120, 2020.
58. Hope NR and Murray GI: The expression profile of RNA-binding proteins in primary and metastatic colorectal cancer: Relationship of heterogeneous nuclear ribonucleoproteins with prognosis. *Hum Pathol* 42: 393-402, 2011.
59. Vakiani E, Janakiraman M, Shen R, Sinha R, Zeng Z, Shia J, Cercek A, Kemeny N, D'Angelica M, Viale A, *et al*: Comparative genomic analysis of primary versus metastatic colorectal carcinomas. *J Clin Oncol* 30: 2956-2962, 2012.
60. Schrijver WAME, Selenica P, Lee JY, Ng CKY, Burke KA, Piscuoglio S, Berman SH, Reis-Filho JS, Weigelt B, van Diest PJ, *et al*: Mutation profiling of key cancer genes in primary breast cancers and their distant metastases. *Cancer Res* 78: 3112-3121, 2018.
61. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, Ritchey JK, Young MA, Lamprecht T, McLellan MD, *et al*: Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481: 506-510, 2012.
62. Johnson BE, Mazar T, Hong C, Barnes M, Aihara K, McLean CY, Fouse SD, Yamamoto S, Ueda H, Tatsuno K, *et al*: Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* 343: 189-193, 2014.

63. Harada K, Okamoto W, Mimaki S, Kawamoto Y, Bando H, Yamashita R, Yuki S, Yoshino T, Komatsu Y, Ohtsu A, *et al*: Comparative sequence analysis of patient-matched primary colorectal cancer, metastatic, and recurrent metastatic tumors after adjuvant FOLFOX chemotherapy. *BMC Cancer* 19: 255, 2019.
64. Gagnière J, Dupré A, Gholami SS, Pezet D, Boerner T, Gönen M, Kingham TP, Allen PJ, Balachandran VP, De Matteo RP, *et al*: Is hepatectomy justified for BRAF mutant colorectal liver metastases?: A multi-institutional analysis of 1497 patients. *Ann Surg* 271: 147-154, 2020.
65. Wu Y, Denhardt DT and Rittling SR: Osteopontin is required for full expression of the transformed phenotype by the ras oncogene. *Br J Cancer* 83: 156-163, 2000.
66. Nemoto H, Rittling SR, Yoshitake H, Furuya K, Amagasa T, Tsuji K, Nifuji A, Denhardt DT and Noda M: Osteopontin deficiency reduces experimental tumor cell metastasis to bone and soft tissues. *J Bone Miner Res* 16: 652-659, 2001.
67. Wu XL, Lin KJ, Bai AP, Wang WX, Meng XK, Su XL, Hou MX, Dong PD, Zhang JJ, Wang ZY, *et al*: Osteopontin knockdown suppresses the growth and angiogenesis of colon cancer cells. *World J Gastroenterol* 20: 10440-10448, 2014.
68. Wei R, Wong JPC and Kwok HF: Osteopontin - a promising biomarker for cancer therapy. *J Cancer* 8: 2173-2183, 2017.
69. Rose AAN, Biondini M, Curiel R and Siegel PM: Targeting GPNMB with glebatumumab vedotin: Current developments and future opportunities for the treatment of cancer. *Pharmacol Ther* 179: 127-141, 2017.