Heterogeneity of colorectal cancers and extraction of discriminator gene signatures for personalized prediction of prognosis

M. MIYAKE[1*], I. TAKEMASA[1*], R. MATOBA[2], M. TANINO[2], S. NIIJIMA[3], M. IKEDA[1], H. YAMAMOTO[1], M. SEKIMOTO[1], S. KUHARA[3], T. OKAYAMA[4], M. MONDEN[1], K. MATSUBARA[2] and M. MORI[1]

[1]Department of Surgery, Graduate School of Medicine, Osaka University, Osaka; [2]DNA Chip Research Inc., Yokohama; [3]Department of Bioinformatics, Graduate School of Systems Life Sciences, Kyushu University, Fukuoka; [4]Bits Co. Ltd., Tokyo, Japan

Received February 16, 2011; Accepted April 12, 2011

DOI: 10.3892/ijo.2011.1092

Abstract. Dissected specimens of colorectal cancer (CRC) have been intensively studied using molecular sketches (gene signatures) to obtain a set of discriminator gene signatures for accurate prognosis prediction in individual patients. The discriminators obtained so far are not universally applicable, as the gene sets reflect the method and site of the study. In this study, we show that dissected stage II and III CRC samples are significantly heterogeneous in molecular sketches, and are not appropriate sources for discriminator extraction unless handled individually. To search for an accurate discriminator gene set for prediction of metastases, we need to start with less heterogeneous stage II CRC. We examined 198 (92 stage II and 106 stage III) CRC dissected samples for the predictability of discriminator gene signatures by analyzing stage II CRC alone, stage III alone, or in combination. The best predictive power of discriminator genes was obtained only when these genes were extracted and validated with stage II CRC samples. An accurate discriminator gene set for the prediction of CRC metastases can be obtained by focusing on stage II CRC samples.

Introduction

Colorectal cancer (CRC) is the second most common malignancy in developed countries (1,2). Colon carcinogenesis is a multi-step process involving progressive changes in genomes and regulatory pathways in colonic epithelial cell proliferation, differentiation and survival, followed by extravasation into lymphatic or blood flow systems. It is categorized into stages I, II, III and IV, depending on the extent of tumor penetration into the intestinal wall, the number of locoregional lymph nodes containing metastases, and distal organ metastases. According to the guidelines of the American Society of Clinical Oncology (ASCO) and National Comprehensive Cancer Network (NCCN), it is recommended that, in most cases, stage II and III CRC tumors be surgically removed. Additional chemotherapy is also recommended for the prevention of metastases for all stage III CRC patients, as well as for some patients with high-grade stage II CRC. However, no adjuvant chemotherapy is recommended on many low-grade stage II CRC patients (3-5). However, some of the stage III patients do not develop recurrence, and some stage II patients do develop recurrence. Therefore, whether or not adjuvant chemotherapy is used after surgery can affect patient survival and quality of life. The cost of medical care is also debatable.

The ongoing individual prediction of CRC prognosis relies mostly upon the clinicopathological observation of the patient. However, the factors available so far, including the depth of tumor invasion (T-stage), and lymph nodal (N-stage) and lymphovascular invasion, are not fully discriminative for correctly predicting the prognosis. Microarray technology with genome-wide analyses of gene expression profiles (gene signatures) of the dissected tumor specimen can help to discover a set of discriminator genes for accurate prognosis prediction in individual patients. Some of these markers may be used to elucidate the genes that are active (or inactive) in metastatic tissues. Remarkable success has been demonstrated in breast cancer (6-8). Similar attempts have been made in CRC (9-21), but the gene sets differ between different laboratories (22) and the predictive power of these discriminators is not as high as expected. It has been argued that these problems may have arisen from the differences in patient treatment, such as adjuvant chemotherapy application, limited number of samples (17,23), differences between work platforms, handling of samples, and statistical methods used. In addition, high-dimensional prediction may yield many models with the same fit (10). We have been working with nearly 700 CRC samples collected in a consortium under a unified protocol to obtain a reliable set of discriminators for metastasis prediction, but so far without success.

Correspondence to: Dr Ichiro Takemasa, Department of Gastroenterological Surgery, Graduate School of Medicine, Osaka University, Yamadaoka 2-2, Suita, Osaka 565-0871, Japan
E-mail: alfa-t@sf6.so-net.ne.jp

*Contributed equally

Key words: colorectal cancer, recurrence, expression profile

The difficulty may have stemmed from the hitherto unidentified complexity or heterogeneity of molecular sketches of CRC. We therefore attempted to reduce the complexity of the samples by focusing on stage II, rather than working with a mixture of stages II and III CRC cases. In this study, we report the results of analyses along this line, and show that focusing on stage II CRC allowed us to obtain a better discriminator gene set for the personalized prognosis of metastasis.

Materials and methods

Patients and clinical materials. In over 5 years, we collected 1978 fresh CRC samples. Data of these patients were added to the registry between 2003 and 2005 (Fig. 1A). All the patients were registered without omission and their prognostic data were up-dated once a year. Post-operative surveillance included clinical evaluation, laboratory tests (including serum carcinoembryonic antigen assay), abdominal CT/US, chest radiography, and colonoscopy in all patients. None of these patients had received chemotherapy.

The study protocol was approved by the Human Ethics Review Committee of Osaka University Graduate School of Medicine, and a signed consent form was obtained from each subject.

RNA preparation, labeling, hybridization and data management. The dissected samples were handled for RNA preparation and analysis at one laboratory station using a standard protocol. Tumor samples were collected within 30 min from the time of resection. They were immediately cut into 5-mm cubes, and after the addition of an RNAase inhibitor, they were stored at -85˚C until RNA extraction. For RNA preparation, samples were dissected from the frozen tissues and homogenized by hand. Total RNA was purified with a TRIzol reagent (Invitrogen, San Diego, CA, USA) using the protocol supplied by the manufacturer. The RNA integrity was assessed with an Agilent 2100 Bioanalyzer and RNA 6000 LabChip kits (Yokokawa Analytical Systems, Tokyo, Japan). Only high-quality RNA with intact 18S and 28S ribosomal RNAs was used for subsequent analysis. Reference colorectal RNA was prepared by combining 40 RNA preparations from normal colorectal mucosa.

The RNA samples were amplified with T7 RNA polymerase using an Amino Allyl MessageAmp™ aRNA kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions, and the quality of each Amino Allyl-aRNA sample was checked with the Agilent 2100 Bioanalyzer. Control and experimental aRNA samples (5 μg) were labeled with Cy3 and Cy5, respectively, and were mixed and hybridized on an oligonucleotide microarray covering 30,000 human probes (AceGene Human 30K; Hitachi Software Engineering Co., Yokohama, Japan). The experimental protocol is described in detail online (http://www.dna-chip.co.jp/thesis/AceGeneProtocol.pdf). The microarrays were scanned using ScanArray 4000 (GSI Lumonics, Billerica, MA, USA). The signal values were calculated by DNASISArray software (Hitachi Software Inc., Tokyo). Following background subtraction, data with low signal intensities were excluded from subsequent analysis. In each sample, the Cy5/Cy3 ratio values were log-transformed and global equalization was applied. Genes with missing values in >10% of samples were excluded from further analysis.

Table I. Post-operative recurrence and clinicopathological features of 693 patients with stage II and III CRC.

A, Stage II

	No recurrence n=331	Recurrence n=40
Age (median)	68 (21-97)	68 (44-80)
Gender male/female	184/147	24/16
Location colon/rectum	216/115	25/15
Tumor size <5cm/≥5cm	145/186	23/17
Depth of tumor ss/se, si[a]	233/98	32/8
Lymphatic invasion ly 0/ly 1, 2, 3[b]	150/181	11/29
Vascular invasion v 0/v 1, 2, 3[c]	195/136	17/23
Histology wel, mod/por, sig, muc[d]	310/21	38/2
Number of examined LNs <13/≥13	131/200	20/20

B, Stage III

	No recurrence n=259	Recurrence n=63
Age (median)	66.5 (28-92)	68 (35-88)
Gender male/female	130/129	34/29
Location colon/rectum	163/96	35/28
Tumor size <5cm/≥5cm	130/129	29/34
Depth of tumor ss/se, si[a]	178/81	31/32
Lymphatic invasion ly 0/ly 1, 2, 3[b]	62/197	2/61
Vascular invasion v 0/v 1, 2, 3[c]	121/138	16/47
Histology wel, mod/por, sig, muc[d]	234/25	59/4
Number of examined LNs[e] <13/≥13	79/180	24/39
Number of LN[e] metastasis <3/≥4	201/58	38/25

Colon (included ascending, transverse, descending and sigmoid colon), rectum (included RS). [a]ss, subserosa; se, serosa exposed; si, serosa infiltrating. [b]ly, lymphatic invasion. [c]v, venous invasion. [d]wel, well; mod, moderate; por, poor; sig, signet ring cell; muc, mucinous. [e]LN, lymph node.

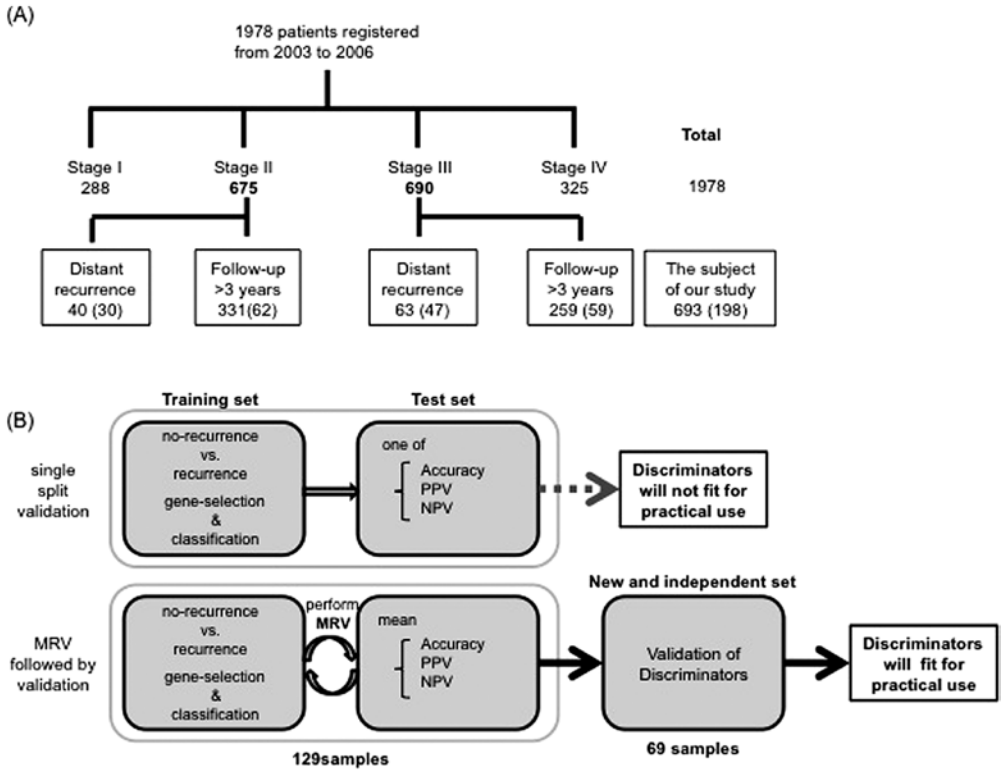

Figure 1. (A) Patients recruited in the present study. Data of 1978 CRC patients were registered from 2003 to 2006. From this collection, 331 stage II and 259 stage III cases, both free of recurrence, along with 40 stage II and 63 stage III cases that developed distance metastases (altogether 693) were taken for clinicopathological analyses. Among these, 198 cases were taken for expression profile analyses. (B) Schemes of discriminator extraction and validation (verification). Discriminator genes are obtained by comparing gene signatures between non-metastasis samples and metastasis samples in a training set. They are then tested (validated) using an independent set of samples for the test. In the single split validation (upper blocks), the training set and test (validation) set are the same. In MRV (lower blocks), as described by Barrier *et al* (10), the extracted discriminators are tested using a new and independent set of samples. We used the MRV for our study. We used 129 samples (70 non-metastasis, 59 metastasis) for discriminator extraction, and 69 samples for the test.

Analyses of the expression data. For the gene signature assay, samples (in this particular experiment n=129) were divided into two groups, one for the training set and another for the validation (test) set. The ratio of recurrence/recurrence-free cases was selected to be equal in the training set and validation set. Calculations within the training set were carried out repeatedly by changing the number of samples (from 10 to 60) in the training set (training set size). We performed 600 random tests in each training set size. Improvement in the mean accuracy was observed with increasing training set size. In this particular experimental system, 600 random tests with a training set size of 50 samples yielded 73% mean accuracy and 6.3% standard deviation, allowing us to adopt a training set size of 50 samples for further analyses. The discriminator genes were selected using the signal-to-noise ratio (SNR) as computed for each gene in the training set. Calculation of the mean accuracy, positive predictive value (PPV, i.e., positive recurrence), and negative predictive value (NPV, i.e., negative relapse for each stage) was carried out based on 100x6 gene selections and validations conducted for each training set size. Missing values were imputed by a simple k-nearest neighbor algorithm (24) and the value of k was set to 5, as described previously (25).

For each gene i, the SNR was defined as SNR (i) = $|\mu_i^{(1)} - \mu_i^{(2)}| / (\sigma_i^{(1)} + \sigma_i^{(2)})$, where $\mu_i^{(2)}$, $\mu_i^{(2)}$ and $\sigma_i^{(1)}$, $\sigma_i^{(2)}$ denote the means and standard deviations of two classes, respectively. The nearest mean classifier was employed in a previous study (7), and confirmed to be effective for outcome prediction despite its simplicity.

Results

Patients and clinicopathological analyses for personalized predictors of metastasis. To obtain as many patient samples as possible while keeping a constant sample quality, the vital requirements for this study, we set up a neo-paradigm of cancer treatment (NCT) consortium in the Kansai area in Japan. This collaborative effort allowed us to collect CRC samples under a unified system of diagnosis, categorization and description of clinicopathological findings. Among the 1978 samples registered (Fig. 1A), we obtained 693 CRC samples, 331 stage II and 259 stage III cases, free of recurrence, along with 40 stage II and 63 stage III cases that developed distant metastases (Table I). We attempted to sort the 693 cases into good or poor prognosis based solely on clinicopathological findings, such as depth of tumor, lymphatic invasion, and vascular invasion. Univariate analysis of stage II and III CRC patients suggested that recurrence could be correlated with certain factors, such as lymphatic invasion or vascular invasion. However, when we performed multivariate Cox proportional hazards regression analysis, no independent single predictor was found for recurrence in stage II CRC, while depth of tumor and lymphatic invasion were independent predictors for stage III CRC patients.

Table II. Clinicopathological features of 129 patients used in the training set to discover discriminators (A) and 69 patients used for the verification (B).

A, 129 patients for discovery study

	No recurrence n=70	Recurrence n=59
Age, mean	65.6	66.6
Gender male/female	37/33	40/19
Stage II/III	39/31	22/37
Differentiation wel, mod/por, sig, muc[a]	68/2	57/2
Location colon/rectum	49/21	43/26
No. of LN[b] harvested, mean	20.1	16.8
Adjuvant chemotherapy used/not used	30 (6)/40 (33)	30 (2)/29 (20)

B, 69 patients for verification study

	No recurrence n=51	Recurrence n=18
Age, mean	68.0	69.0
Gender male/female	29/22	15/3
Stage II/III	23/28	8/10
Differentiation wel, mod/por, sig, muc[a]	49/2	16/2
Location colon/rectum	33/18	12/6
No. of LN[b] harvested, mean	19.2	14.0
Adjuvant chemotherapy used/not used	26 (4)/25 (19)	7 (3)/11 (5)

[a]wel, well; mod, moderate; por, poor; sig, signet ring cell; muc, mucinous. [b]LN, lymph node.

Table III. Verification of the discriminators using 69 new samples.

Verification using stage II samples[a]

	No recurrence (n=23)	Recurrence (n=8)
Gene signature		
No recurrence	18	2
Recurrence	5	6

Verification using stage III samples[b]

	No recurrence (n=28)	Recurrence (n=10)
Gene signature		
No recurrence	10	2
Recurrence	18	8

[a]Accuracy, 77.4%; PPV, 54.5%; NPV, 90%. [b]Accuracy, 47.4%; PPV, 30.8%; NPV, 83.3%. Discriminators were obtained using stage II samples (Table IIA) and verified using stage II set or stage III set samples (Table IIB).

Discriminator extraction using expression profiles (gene signatures) and its verification. We then took 129 dissected specimens (70 patients with no metastasis and 59 with metastasis; Table IIA) from the 693 samples and analyzed the gene signatures. RNA extraction, microarray analyses and data management were carried out as described in Materials and methods. Expression profiles were then compared between the metastatic (recurrence) and non-metastatic samples to extract discriminator genes using a multiple random training-test strategy (MRV) system, as described by Barrier *et al* (10) (Fig. 1B). The resulting discriminators were used as those of the training set samples. We then took another independent set of 69 samples shown in Table IIB for validation (verification) of the discriminators. The validation with a new, independent set of samples guarantees the results to be free from over-fitting (Fig. 1B).

To test the quality of the discriminators, we extracted them from different combinations of stage II and stage III mixtures. The sources for this experiment are shown in Table II. Samples in each stage were randomly selected, admixed at a ratio of 100/0, 80/20, 60/40, 40/60, 20/80 and 0/100, and each of them was used as a training set. Six different discriminator sets were thus obtained, and were subjected to validation using two independent test sets (consisting of stage II alone or stage III alone) constructed from samples in Table IIA. The results are shown in Fig. 2A and B. The proportion of stage II and III samples in the training set greatly affected the accuracy of the resulting discriminators: The higher the proportion of stage II samples in the training set, the higher the accuracy, when evaluated with stage II validation samples (Fig. 2A). However, when we used the same set of discriminators and evaluated with stage III validation samples, we obtained a similar trend, but the highest mean accuracy was not high (Fig. 2B). The favorable result obtained with the stage II combination was not due to the matching of stages between training and validation sets, as seen in the stage III combination. The best performance was obtained with the combination of stage II training samples and stage II validation samples. For example, in an independent, but similar experiment in which both the training and test were performed with stage II samples (Table III), the results showed 18/23 (78.3%) non-recurrences and 6/8 (75%) recurrences; the prediction accuracy being (18+6)/(23+8) (77.4%), with a PPV of 6/(5+6) (54.5%) and a NPV of 18/(18+2) (90.0%). In contrast, when training and tests were performed with stage III samples, the correct prediction was 10/(10+18) (35.7%) non-recurrences and 8/(8+2) (80%) recurrences, and the prediction accuracy was (10+8)/(28+10)

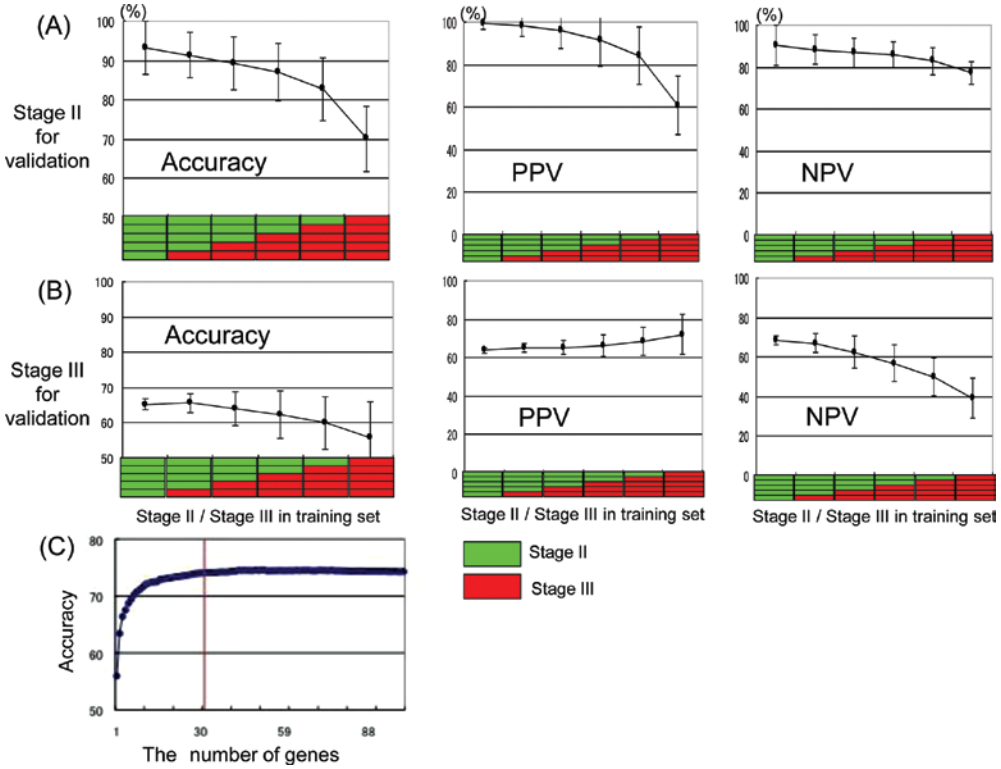

Figure 2. The proportion of stage II and stage III CRC cells affects the quality of the discriminators. Dissected stage II (green tiles) and III (red tiles) CRC specimens (see Table IA and B) were admixed in six different proportions (100/0, 80/20, 60/40, 40/60, 20/80 and 0/100). In each mixture the ratio of metastasis/non-metastasis was set to be identical in the stage II sample as well as in the stage III sample. Six different discriminator sets were extracted from each one of these mixtures using the MRV system. The resulting discriminators were then validated for accuracy, PPV and NPV (ordinate) using stage II samples (A) or stage III samples (B). Data are the means ± SD of accurate classification, which is equal to the percentage obtained from 100 random training-validation sets. (C) Number of genes in discriminators versus accuracy.

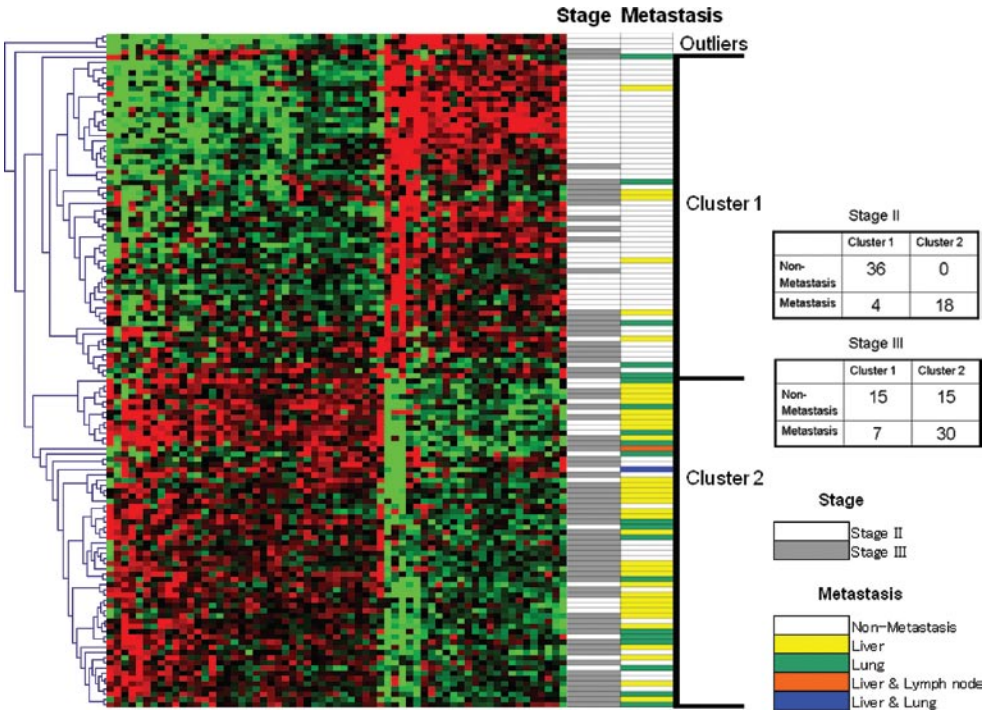

Figure 3. Cluster analyses of gene signatures of 129 CRC samples using a one-way Anova analysis. Four groups (non-metastatic stage II, metastatic stage II, non-metastatic stage III, metastatic of stage III) were expected, but essentially 3 clusters resulted. The clustering was carried out using unsupervised hierarchical clustering with Euclidean distance and the average linkage method using TIGR MultiExperimentViewer (MEV)-version 4.0.01. The left panel shows the heat map of the training samples based on 63 significant genes. Ordinate, samples; abscissa, genes. Overall threshold p-value=0.01. Significance was determined by standard Bonferroni correction. Columns at the right marked 'Stage' and 'Metastasis' show sample characteristics (stage II, white; stage III, gray). For description of metastasis, see accompanying index at lower right. The two contingency tables to the right show the composition of patients in each cluster and the metastasis status in stage II or III.

Table IV. List of 30 discriminator genes selected in the training set containing stage II CRC only.

AceGene ID	RefSeq ID	Symbol	S2N score	Number selected by MRV
Down-regulated genes in patients with recurrence				
AGhsA220222	NM_000112	SLC26A2	-1.187	98
AGhsC181602			-1.086	99
AGhsB151215	NM_022901	LRRC19	-0.914	76
AGhsC020219			-0.835	49
AGhsA071406		TP73L	-0.79	23
AGhsB190623	NM_152291	MUC7	-0.788	24
AGhsB260322		PSG3	-0.692	40
Up-regulated genes in patients with recurrence				
AGhsB261517	NM_003992	CLK3	0.723	24
AGhsC141104			0.733	24
AGhsC051216			0.737	29
AGhsB040907	NM_006694	JTB	0.741	27
AGhsB230319	NM_018406	MUC4	0.754	27
AGhsB061613	NM_182679	GPATC4	0.757	29
AGhsC141107			0.763	27
AGhsC251406			0.765	32
AGhsC160815			0.775	30
AGhsA100414	NM_016341	PLCE1	0.779	36
AGhsB161617	XM_370958		0.779	28
AGhsA061506	NM_024307	GDPD3	0.785	31
AGhsB051615	NR_002194	CSPG4LYP1	0.795	44
AGhsB140718	NM_021910	FXYD3	0.801	47
AGhsA141613			0.802	38
AGhsA030519	NM_021195	CLDN6	0.843	57
AGhsB200103	NM_023002	HAPLN4	0.843	65
AGhsB260314			0.847	54
AGhsA021018	NM_017636	TRPM4	0.85	56
AGhsC010106	NM_001001410	C16orf42	0.853	51
AGhsC160514			0.895	65
AGhsB181208		MEG3	0.936	84
AGhsB081519			0.967	88

(47.4%), with a PPV of 8/(18+8) (30.8%) and a NPV of 10/(10+2) (83.3%). These data strongly suggest that stage III samples are more heterogeneous than stage II samples.

The appropriate number of informative discriminator genes with the largest value was identified by changing their number (from 5 to 150) in each training-validation. We selected 30 top-ranking genes with the largest values as discriminators for prediction (Fig. 2C). When the number is small, there is underpowered predictability, and when large, there is a complication.

Difference in clusters of gene signature between stage II and stage III CRC. To obtain further evidence for the heterogeneity in stage III samples, we subjected the gene signatures of the 129 CRC samples to cluster analyses. Profiles of samples of non-metastatic stage II, metastatic stage II, non-metastatic stage III, and metastatic stage III were compared using ANOVA analysis. All the data were quantile normalized after assembly, and data with genes having >6 missing samples were eliminated to avoid possible bias resulting from microarray processing. The results are shown in Fig. 3. All non-metastatic stage II samples were located in cluster 1, whereas non-metastatic stage III samples were almost evenly distributed within clusters 1 and 2. Four and six metastasis samples from stage II and III, respectively, were located in cluster 1, while 48 metastasis samples from stage II and stage III were in cluster 2. These results showed that the stage III samples were more heterogeneous than the stage II samples (p-value <E-7). The metastatic samples from stage II and stage III were similar, if not identical, and were different from the non-metastatic stage II samples.

Genes selected by MRV. We selected 100x30 discriminator genes (3,000 genes) by 100 MRV in the training set that included only tumors of the same stage. We selected 491 genes from the stage II training set, wherein the number of genes selected 50 to 100 times was 11. However, 1,161 genes were selected from the stage III training set, and only one gene was selected in the window of 50 to 100 MRV, in conformity with the idea that stage III is more heterogeneous than stage II. Table IV shows the 30 discriminator genes for prognosis from stage II samples repeatedly selected by MRV.

Table V. List of 30 genes selected in the training set containing stage III CRC only.

AceGene ID	RefSeq ID	Symbol	S2N score	Number selected by MRV
Down-regulated genes in patients with recurrence				
AGhsC090622	NM_031449	ZMIZ2	-0.554	46
AGhsB260322		PSG3	-0.516	30
AGhsA030615	NM_014730	KIAA0152	-0.471	28
AGhsA021003	NM_018206	VPS35	-0.466	16
AGhsB011103	XM_031553		-0.464	16
AGhsA081007	NM_030581	WDR59	-0.46	18
AGhsB150212	NM_021217	ZNF77	-0.457	14
AGhsB180118	NM_007335	DLEC1	-0.455	16
AGhsA211612	NM_004520	KIF2A	-0.454	18
AGhsB030311	NM_032962	CCL15	-0.446	19
AGhsB050212			-0.435	13
AGhsB070303			-0.433	15
AGhsA120218	NM_016500	CXorf26	-0.433	16
AGhsA081515	NM_021963	NAP1L2	-0.416	23
AGhsB110823	NM_016625	RSRC1	-0.415	18
AGhsC010122	NM_213589	RAPH1	-0.412	15
AGhsB161223	NM_000046	ARSB	-0.402	17
Up-regulated genes in patients with recurrence				
AGhsB021202	NM_016264	ZNF44	0.36	17
AGhsB200718			0.385	16
AGhsA060812	NM_002970	SAT1	0.417	13
AGhsB040204	NM_018949	UTS2R	0.425	19
AGhsA131112	NM_003823	TNFRSF6B	0.448	16
AGhsA020705	NM_005345	HSPA1B	0.451	18
AGhsC051213	XM_001129756		0.465	14
AGhsA210414	NM_004568	SERPINB6	0.466	25
AGhsC261219			0.469	17
AGhsB180210	NM_002970	SAT1	0.48	25
AGhsA020911	NM_022873	IFI6	0.493	18
AGhsA240401	NM_004457	ACSL3	0.5	22
AGhsA221113	NM_005908	MANBA	0.588	50

For comparison, 30 genes selected similarly from stage III CRC samples are also listed in Table V.

Discussion

Although the current staging system of CRC based on TNM classification is useful for correlating groups of patients with different recurrence rates, it cannot be directly used for the personalized prediction of recurrence for stage II or III patients. Some of these patients will be cured, whereas others will develop recurrence, even though they had similar clinicopathological findings and were operated and treated similarly in after-care. However, another problem is that a significant fraction of the patients receive unnecessary chemotherapy. Personalized prediction of recurrence is awaited to avoid under-treatment or over-treatment, and therefore, the demand is high for expression signatures for their possible contribution on this issue. A set of discriminator genes to differentiate good prognosis and poor prognosis must be obtained. This is not just collecting one or a few metastasis-related genes for understanding the cause and consequence of functional change in a gene that is related to metastasis. Discriminators are extracted by a different approach. As noted in the Introduction however, discriminators obtained from several laboratories are not the same, and their performance has not been convincing.

We have been working along the same line, and noticed also the lack of robustness with the CRC discriminator set from different studies. This could be due, in part, to the difference in the platform used or sample handling, or to the small sample numbers. However, CRC might be more heterogeneous than expected. This would cause difficulty in defining the category of 'good prognosis CRC', or 'poor prognosis CRC' by assuming that each of them represents a relatively uniform single entity. CRC has been reported to be heterogeneous clinicopathologically, but attempts for sub-categorization have not been successful. Weiser *et al* found by nomogram studies great heterogeneity within CRC (26), especially the stage III cohort, and possible variability in the probability of recurrence.

Our expression signature studies showed that CRC samples from stage II are already heterogeneous, but that CRC samples from stage III patients are much more heterogeneous.

Thus, by focusing on stage II CRC, we obtained better discriminators than by working with mixtures of stages II and III CRC. It is clear that discriminators extracted from stage III CRC are less powerful. At this stage, we do not know how heterogeneous the CRC samples are. We would have to handle at least N times more CRC samples in power analyses with a mixture of stage II and stage III CRC samples together, assuming that the complexity or heterogeneity of the CRC mixture is N times that for breast cancer. We will analyze a larger number of stage II CRC in the near future. At least twice more samples should be handled, if the stage II CRC cases consist of two categories (26). Although the application is limited at this moment, we believe that focusing on stage II CRC is clinically important for the personalized care and management of the CRC patients.

The discriminator gene set is not complete even for stage II CRC (prediction accuracy 77.4%). Nevertheless, the following genes may be of some interest: Pregnancy-specific β-1-glycoprotein 3 (PSG3), which is a member of the carcinoembryonic antigen (CEA) gene family; TP73L (p63), a member of the p53 family related to the negative regulation of the cell cycle and apoptosis; DLEC1, a tumor suppressor gene identified in lung, esophageal and renal cancers; PLCE1, a regulator of small GTPases of the Ras superfamily; and Muc4, a regulator of p27, that is a cyclin-dependent kinase inhibitor involved in the control of G1 and S phases of the cell cycle, linked to poor prognosis in many cancers. Other genes include TNFRSF6B, HSPA1B and IFI6, which guard against apoptosis.

How did the heterogeneity of CRC, especially that of stage III CRC, evolve? To answer this question we need to subdivide stage II and III CRC using gene expression profiles. We also need to know whether or not metastatic CRC is similar, or very heterogeneous, as suggested by the profiles shown in Fig. 3 profiles, and also indirectly by Jorissen *et al* (14). In breast cancer, gene signature prediction of recurrence (7,8) has been approved by the Food and Drug Administration. The products are currently available (6), and have been recommended for use in the ASCO and NCCN guidelines. The accuracy of predicting prognosis in CRC cases is not as high as that in breast cancer cases, even when its application is limited to stage II CRC cases. Further studies to accurately categorize CRC by combining clinicopathological findings and molecular analyses will enable more strategic personalized prediction, and will allow for the appropriate validation and clinical application of these findings (22).

In summary, our diagnostic system using the gene signature of stage II CRC, however limited in capability, should be useful for the prediction of prognosis of individual patients after curative surgery. The expression profiling data and supplemental information are available at (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12032).

Acknowledgements

We thank the following physicians for providing the surgical specimens used in the present study. Dr Masakazu Ikenaga from the National Osaka Hospital, Dr Shingo Noura from the Osaka Medical Center for Cancer and Cardiovascular Diseases, Dr Hiroyoshi Takemoto from the Sakai Municipal Hospital, Dr Taishi Hata from the Toyonaka Municipal Hospital, Dr Yasuhiro Miyake from the Mino Municipal Hospital, Dr Tadashi Onishi from the NTT West Osaka Hospital, Dr Yoshio Uemura from the Kinki Central Hospital, Dr Kouhei Murata from the Suita Municipal Hospital, and Dr Shu Okamura from the Kansai Rosai Hospital.

References

1. Pisani P, Parkin DM and Ferlay J: Estimates of the worldwide mortality from eighteen major cancers in 1985. Implications for prevention and projections of future burden. Int J Cancer 55: 891-903, 1993.
2. Jemal A, Siegel R, Ward E, Hao Y, Xu J and Thun MJ: Cancer statistics, 2009. CA Cancer J Clin 59: 225-249, 2009.
3. Andre T, Boni C, Mounedji-Boudiaf L, *et al*: Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. N Engl J Med 350: 2343-2351, 2004.
4. Benson AB III, Schrag D, Somerfield MR, *et al*: American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. J Clin Oncol 22: 3408-3419, 2004.
5. Figueredo A, Charette ML, Maroun J, Brouwers MC and Zuraw L: Adjuvant therapy for stage II colon cancer: a systematic review from the Cancer Care Ontario Program in evidence-based care's gastrointestinal cancer disease site group. J Clin Oncol 22: 3395-3407, 2004.
6. Cronin M, Pho M, Dutta D, *et al*: Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. Am J Pathol 164: 35-42, 2004.
7. Van't Veer LJ, Dai H, van de Vijver MJ, *et al*: Gene expression profiling predicts clinical outcome of breast cancer. Nature 415: 530-536, 2002.
8. Van de Vijver MJ, He YD, van't Veer LJ, *et al*: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347: 1999-2009, 2002.
9. Arango D, Laiho P, Kokko A, *et al*: Gene-expression profiling predicts recurrence in Dukes' C colorectal cancer. Gastroenterology 129: 874-884, 2005.
10. Barrier A, Boelle PY, Roser F, *et al*: Stage II colon cancer prognosis prediction by tumor gene expression profiling. J Clin Oncol 24: 4685-4691, 2006.
11. Eschrich S, Yang I, Bloom G, *et al*: Molecular staging for survival prediction of colorectal cancer patients. J Clin Oncol 23: 3526-3535, 2005.
12. Fan J, Peng Z, Zhou C, *et al*: Gene-expression profiling in Chinese patients with colon cancer by coupling experimental and bioinformatic genomewide gene-expression analyses: identification and validation of IFITM3 as a biomarker of early colon carcinogenesis. Cancer 113: 266-275, 2008.
13. Fritzmann J, Morkel M, Besser D, *et a*l: A colorectal cancer expression profile that includes transforming growth factor beta inhibitor BAMBI predicts metastatic potential. Gastroenterology 137: 165-175, 2009.
14. Jorissen RN, Gibbs P, Christie M, *et al*: Metastasis-Associated Gene Expression Changes Predict Poor Outcomes in Patients with Dukes Stage B and C Colorectal Cancer. Clin Cancer Res 15: 7642-7651, 2009.
15. Ki DH, Jeung HC, Park CH, *et al*: Whole genome analysis for liver metastasis gene signatures in colorectal cancer. Int J Cancer 121: 2005-2012, 2007.
16. Kleivi K, Lind GE, Diep CB, *et al*: Gene expression profiles of primary colorectal carcinomas, liver metastases, and carcinomatoses. Mol Cancer 6: 2, 2007.
17. Shi L, Reid LH, Jones WD, *et al*: The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 24: 1151-1161, 2006.
18. Shih W, Chetty R and Tsao MS: Expression profiling by microarrays in colorectal cancer (Review). Oncol Rep 13: 517-524, 2005.
19. Wang Y, Jatkoe T, Zhang Y, *et al*: Gene expression profiles and molecular markers to predict recurrence of Dukes' B colon cancer. J Clin Oncol 22: 1564-1571, 2004.

20. Watanabe T, Kobunai T, Sakamoto E, *et al*: Gene expression signature for recurrence in stage III colorectal cancers. Cancer 115: 283-292, 2009.
21. Yamasaki M, Takemasa I, Komori T, *et al*: The gene expression profile represents the molecular nature of liver metastasis in colorectal cancer. Int J Oncol 30: 129-138, 2007.
22. Kopetz S and Abbruzzese JL: Barriers to Integrating Gene Profiling for Stage II Colon Cancer. Clin Cancer Res 15: 7451-7452, 2009.
23. Tan PK, Downey TJ, Spitznagel EL, *et al*: Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Res 31: 5676-5684, 2003.
24. Troyanskaya O, Cantor M, Sherlock G, *et al*: Missing value estimation methods for DNA microarrays. Bioinformatics 17: 520-525, 2001.
25. Dudoit S, Fridyand J and Speed TP: Comparison of discrimination methods for the classification of tumors using gene expression data. J Am Stat Assoc 97: 77-87, 2002.
26. Weiser MR, Landmann RG, Kattan MW, *et al*: Individualized prediction of colon cancer recurrence using a nomogram. J Clin Oncol 26: 380-385, 2008.