available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Prediction of liver metastasis after colorectal cancer using reverse transcription-polymerase chain reaction analysis of 10 genes

Toshiaki Watanabe [a,*], *Takashi Kobunai* [a,c], *Yoko Yamamoto* [a], *Takamitsu Kanazawa* [b], *Tsuyoshi Konishi* [b], *Toshiaki Tanaka* [b], *Keiji Matsuda* [a], *Soichiro Ishihara* [a], *Keijiro Nozawa* [a], *Kiyoshi Eshima* [c], *Tetsuichiro Muto* [d], *Hirokazu Nagawa* [b]

[a] *Department of Surgery, Teikyo University School of Medicine, 2-11-1 Kaga, Itabashi-ku, Tokyo 173-8605, Japan*
[b] *Department of Surgical Oncology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan*
[c] *Tokushima Research Center, Taiho Pharmaceutical Co., Ltd., 224-2, Ebisuno, Hiraishi, Kawauchi-cho, Tokushima 771-0194, Japan*
[d] *The Cancer Institute Hospital of JFCR, 3-10-6, Ariake, Kotoh-ku, Tokyo 135-8550, Japan*

ARTICLE INFO

Article history:
Received 8 February 2010
Received in revised form 19 April 2010
Accepted 20 April 2010
Available online 1 June 2010

Keywords:
Liver metastasis
Colorectal cancer
RT-PCR
Microarray
Prediction

ABSTRACT

Purpose: Liver metastasis is one of the major types of recurrence after surgery for colorectal cancer. Traditional methods of predicting liver metastasis are limited in their accuracy, suggesting the need to develop new predictors. We developed a 10-gene signature that is closely associated with the development of liver metastasis after colorectal cancer.

Patients and methods: We examined a total of 189 frozen specimens of primary colorectal cancers using both microarray and quantitative real-time reverse transcriptase-polymerase chain reaction (RT-PCR) analysis. Initially, we studied gene expression in colorectal cancer tissue from 160 randomly selected patients who had undergone surgical resection of colorectal cancer and evaluated the association between the level of gene expression and the occurrence of liver metastasis. We developed a gene-expression model for the prediction of liver metastasis based on the RT-PCR findings. We then used specimens from 29 other patients for validation.

Results: The expression of 14 genes was correlated with liver metastasis according to both microarray and RT-PCR analysis. We constructed an accurate predictive model based on the results for 10 of these genes, which included epiregulin (*EREG*), amphiregulin (*AREG*), cyclooxygenase 2 (*COX-2*) and lymphocyte-specific protein tyrosine kinase (*LCK*). The 10-gene signature was an independent predictor of liver metastasis. The model was validated in the independent set of 29 patients. The predictive accuracy of the model in a test set of patients was 86.2%.

Conclusion: The 10-gene signature identified in this study is closely associated with the occurrence of liver metastasis in colorectal cancer patients.

1. Introduction

Colorectal cancer is the third most common type of cancer worldwide.[1] Studies have shown that adjuvant chemotherapy significantly reduces the recurrence rate and improves survival after surgery for colorectal cancer.[2,3] However, adverse events occur in some patients who receive adjuvant chemotherapy. We previously demonstrated that expression of

* *Corresponding author:* Tel.: +81 3 3964 1231; fax: +81 3 5375 6097.
E-mail address: toshwatanabe@yahoo.co.jp (T. Watanabe).
0959-8049/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2010.04.019

molecular markers could select those patients who would benefit from adjuvant chemotherapy[4]; however, we did not focus on specific sites of recurrence. As liver metastasis is one of the major forms of recurrence after surgery for colorectal cancer, the identification of high-risk patients could reduce its development by allowing intensive adjuvant chemotherapy or curative surgical resection of the disease to be carried out. However, no specific markers for the prediction of liver metastases after surgery have yet been identified.

We previously reported on the use of DNA microarrays in predicting the development of colitic cancer and classifying molecular subtypes of microsatellite instability in colorectal cancer.[5–7] Microarrays have also been used in several studies to predict the development of liver metastasis by colorectal cancer.[8–12] However, microarrays lack reproducibility for the quantitative analysis of gene expression and therefore require verification by an alternative approach such as the real-time polymerase chain reaction (PCR).[13–18] In addition, a large number of genes are used in prediction models. Because of these difficulties, the use of microarrays in clinical practice is still limited and recent studies have shown that reverse transcription (RT)-PCR analyses of a small number of selected genes are capable of accurately predicting the outcome of patients with diffuse large-B-cell lymphoma lung cancer.[14,15]

Accordingly, in the present study, we aimed to establish a predictive model of liver metastasis based on RT-PCR analysis. Using the expression levels of 10 genes analysed by RT-PCR, our predictive model gave an accuracy rate of 86.2%. To the best of our knowledge, this is the first study to predict recurrence in the form of liver metastasis based on gene-expression levels determined by RT-PCR.

2. Materials and methods

Amongst colorectal cancer patients who underwent surgery between 2001 and 2007, 189 patients whose RNA was available for microarray analysis were included in this study. Patients were evaluated at 3-month intervals for the first year after surgery and at 6-month intervals thereafter. Liver metastases were identified by clinical and laboratory examination and verified by liver ultrasound or computerised tomography (CT) scan. The follow-up was standardised amongst all patients. In total, 53 of the patients had metachronous liver metastasis with or without synchronous liver metastasis, whereas the other 136 patients remained liver metastasis-free throughout the follow-up period. Patients were divided randomly into a training set (160 patients) and a test set (29 patients). Because the main purpose of the study was to predict the development of recurrence in the liver, we selected patients with metachronous liver metastasis alone for inclusion in the test set. The training set consisted of 45 patients with metachronous and/or synchronous liver metastasis and 115 patients without. The test set consisted of eight patients with metachronous liver metastasis alone and 21 patients without. Patient and tumour characteristics are shown in Table 1.

Tumour specimens were taken from surgically resected specimens and used for RNA extraction as described previously.[5–7] Informed consent was obtained from patients for the collection of specimens, and the study protocol was approved by the local ethics committee.

2.1. RNA isolation, microarray expression profiling and its validation

Total RNA was isolated from frozen samples using the RNeasy Mini Kit (Qiagen, Chatsworth, GA) for gene-expression analysis. Gene-expression profiles were determined using Affymetrix HG-U133 Plus 2.0 GeneChips (Affymetrix, Santa Clara, CA) according to the manufacturer's recommendations, as described previously.[6,7] Differences in gene-expression levels identified by microarray analysis were validated by Taqman real-time PCR (Applied Biosystems, Foster City, CA) as described previously.[19]

2.2. Microarray data analysis

Gene-expression profiles of samples were prepared with GeneSpring software version 7.3 (Agilent Technologies, Santa Clara, CA). To identify probes the expression of which differed significantly between patients with and without liver metastasis, we examined 160 training samples. Expression profiles were compared using unpaired *t*-tests (Benjamini and Hochberg false discovery rate (FDR) controlling procedure) and a fold-change value. Identified probes were then used to perform supervised class prediction by the k-nearest-neighbour method (KNN) and a 10-fold cross-validation.[20] The significance of classification accuracy was confirmed by the permutation test in the *R* computing environment. Using selected probes, two-dimensional hierarchical clustering was applied to the log-transformed data obtained by GeneChip, and variation in multigene expression between patients with and without liver metastasis was compared by principal component analysis (PCA).

2.3. Quantitative RT-PCR analysis

The relationship was determined between the microarray expression ratio and RT-PCR analysis of differentially expressed probes that could be identified by UniGene in all of the samples that had been subjected to microarray analysis. We then established a prediction model for liver metastases in the training set as described in the GeneChip analysis and applied the model to an independent test set to validate its accuracy.

2.4. Statistical analysis

Categorical data were statistically analysed using the χ^2 test or Fisher's exact test. Continuous data were analysed using the Student's *t*-test. The Kaplan–Meier survival model was used to estimate the outcome of patients stratified by the prediction model and clinicopathological variables. Patients who were predicted to become positive for liver metastasis were classified as 'high-risk signature' patients and those predicted to be negative were classified as 'low-risk signature' patients. In the analysis of the probability that patients would remain liver metastasis-free, we defined liver metastasis as the first event. Patient data were analysed from the date of surgery

Table 1 – Patient and tumour characteristics.

Characteristic		Training patients (n = 160)	Test patients (n = 29)	p Value
Age	Mean (SD)	63.3 (12.9)	65.1 (10.6)	0.50[a]
	Range (years)	22–86	33–92	
Gender	Male	99	22	0.15[b]
	Female	61	7	
T classification*	*T*1, 2	39	6	0.96[c]
	*T*3	87	17	
	*T*4	34	6	
Lymph-node metastasis	Present	69	15	0.95[b]
	Absent	91	14	
Liver metastasis	Present	45	8	0.95[c]
	Absent	115	21	
Size (mm)	Mean (SD)	46.8 (25.1)	45.0 (24.9)	0.59[a]
	Range	11–130	15–125	
Site	Colon	96	19	0.58[b]
	Rectum	64	10	
Follow up period (years)	Mean (SD)	3.6(1.4)	3.9(1.7)	0.33[a]
	Range	1.4–6.7	1.6–8.2	
Histological type	Well	119	22	0.68[c]
	Moderate	34	5	
	Poor	5	2	
	Others	2	0	

* *T* classification: TNM classification
[a] Student's *t*-test.
[b] χ^2 test.
[c] Fisher's exact test.

to the time of the first event or the date on which data were censored. The log-rank test was used to assess the significance of differences between survival curves. All differences were considered statistically significant if the p value was <0.05.

The Cox proportional hazard model was used for univariate and multivariate survival analyses and the hazard ratios were calculated. Variables with p values <0.4 in a univariate analysis were included in the Cox proportional hazard model for multivariate analysis. All p values were two-tailed, and those <0.05 were considered statistically significant.

3. Results

3.1. Microarray data analysis

We identified 26 probes (*t*-test with FDR $p < 7 \times 10^{-5}$) that were differentially expressed between patients with and without liver metastasis (Table 2). Next, we established a prediction model for liver metastasis in the training set with the 26 discriminating probes. The classification accuracy was calculated as follows. First, the average classification accuracy was measured by a 100 times-repeated 10-fold cross validation using a 26-probe predictor. The average prediction accuracy was 86.1%. Then, the predictive performance of the model was evaluated by a permutation-based procedure. The class labels of the samples were permutated 5000 times, yielding a new signature, and the 10-fold cross-validation accuracy for each permutated data set was calculated. Finally, we measured the random chance of obtaining a signature with greater accuracy than the 26 signature. The false-finding rate was 0.0028 (that is, 14 amongst 5000 permutations exhibited higher prediction accuracy than in the actual data). These 26 probes were used to perform a hierarchical cluster analysis and PCA (Fig. 1).

3.2. Quantitative RT-PCR analysis and prediction of liver metastasis

Amongst the 26 discriminating probes, 17 genes (21 probes) could be identified by UniGene (Table 2). Quantitative RT-PCR was performed to validate the microarray expression level of probes in the training set. Samples in two patients were insufficient for quantitative RT-PCR analysis, so this was successful in 158 patients. The analyses showed significant differences in the expression of 14 of the 17 genes (17 probes) between patients with and without liver metastasis (FDR <0.05; Table 2). RT-PCR analyses showed higher expression in four of these genes and lower expression in 10 of these genes in patients with liver metastasis compared with those without, which was in agreement with the results obtained by the GeneChip analysis (Supplementary Fig. S1).

Next, based on the RT-PCR data, we established a prediction model in the training set. To determine the number of probes that provided the best separation between patients with and without liver metastasis, we ranked the 14 genes on the basis of the significance of their FDR p values and in decrements of one starting at the bottom of the rank-ordered list (14, 13, 12 and so on). Fig. 2 shows the different prediction rates when the number of discriminating probes was changed. The best accuracy rate, 86.2%, was obtained when we used the 10 top-ranked genes (12 probes) (Table 2). The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were 87.5%, 85.7%, 70.0% and 94.7%, respectively.

Table 2 – Probes with significantly different expression between patients with and without liver metastasis by GeneChip analysis and their validation results with TaqMan real-time PCR.

	GeneChip					TaqMan low density array		
Probe ID	Fold change (LM-N*/LM-P**)	FDR *p* value	Common name	Genebank accession	UniGene ID	Taqman assay ID	Fold change (LM-N*/LM-P**)	FDR *p* value
228854_at	0.23	3.10E−05		AI492388				
1569583_at	0.34	6.03E−05	EREG	BC035806	Hs.115263	Hs00914313_m1	0.36	0.0029[a]
235652_at	0.36	1.86E−05	SCML1	AI431345	Hs.662047	Hs00232467_m1	0.81	0.136
236340_at	0.37	4.53E−05		AI769947				
243588_at	0.43	6.03E−05		N74058				
209781_s_at	0.43	1.09E−05	KHDRBS3	AF069681	Hs.444558	Hs00198065_m1	0.53	0.0031[a]
1557285_at	0.44	9.30E−06	AREG	AI891075	Hs.270833	Hs00155832_m1	0.50	0.0031[a]
232535_at	0.45	1.28E−05		AL133570				
213577_at	0.45	9.30E−06	SQLE	AA639705	Hs.71465	Hs00162288_m1	0.77	0.1190
1570441_at	0.48	9.30E−06	NAPB	BC008463	Hs.269471	Hs00397268_m1	0.64	0.0015[a]
222366_at	0.50	4.22E−05		W86781				
215719_x_at	2.09	6.03E−05	FAS	X83493	Hs.244139	Hs00163653_m1	1.31	0.0436
204821_at	2.14	6.04E−05	BTN3A3	NM_006994	Hs.167741	Hs00757230_m1	1.38	0.0166
204891_s_at	2.17	1.09E−05	LCK	NM_005356	Hs.470627	Hs00178427_m1	2.35	0.0002[a]
216252_x_at	2.21	2.98E−05	FAS	Z70519	Hs.244139	Hs00163653_m1	1.31	0.0436
217478_s_at	2.28	6.03E−05	HLA-DMA	X76775	Hs.351279	Hs00185435_m1	1.87	0.0031
209813_x_at	2.38	2.87E−05	TRGV9	M16768	Hs.534032	Hs00233330_m1	2.11	0.0015[a]
205488_at	2.41	1.83E−05	GZMA	NM_006144	Hs.90708	Hs00196206_m1	2.17	0.0008[a]
204890_s_at	2.70	2.30E−05	LCK	U07236	Hs.470627	Hs00178427_m1	2.35	0.0002[a]
209312_x_at	2.79	4.40E−05	HLA-DRB3	U65585	Hs.534322	Hs00734212_m1	0.51	0.2380
204748_at	2.80	6.03E−05	PTGS2	NM_000963	Hs.196384	Hs00153133_m1	3.23	0.0002[a]
215193_x_at	2.90	1.98E−05	HLA-DRB3	AJ297586	Hs.534322	Hs00734212_m1	0.51	0.2380
211991_s_at	3.08	2.49E−05	HLA-DPA1	M27487	Hs.347270	Hs00410276_m1	1.63	0.0099
1567628_at	3.27	1.79E−05	CD74	M28590	Hs.436568	Hs00269961_m1	1.92	0.0016[a]
208894_at	3.42	2.17E−05	HLA-DRA	M60334	Hs.520048	Hs00219575_m1	2.16	0.0004[a]
210982_s_at	3.79	5.28E−05	HLA-DRA	M60333	Hs.520048	Hs00219575_m1	2.16	0.0004[a]

* LM-N: liver metastasis negative.
** LM-P: liver metastasis positive.
[a] Top-ranked 10 genes (12 probes) that gave the best predictive accuracy rate for liver metastasis.

3.3. *Prognostic value of gene-expression signatures and Clinicopathological Factors*

We calculated the probability of remaining free of liver metastasis and of overall survival, according to the gene-expression signature and clinicopathological factors in the test set. The Kaplan–Meier curves showed a significant difference between patients with high-risk and low-risk signatures (Fig. 3). Moreover, there was a significant difference between patients with and without lymph-node metastasis. However, no differences were observed in regard to other clinicopathological factors. The estimated hazard ratio for liver metastases as a first event in patients with high-risk signatures in comparison with low-risk signatures was 4.17 (95% confidence interval, 1.75–18.09; $p = 0.0005$). The estimated hazard ratios according to the clinicopathological variables are shown in Tables 3.1 and 3.2.

Multivariable analysis found the high-risk signature to be the only independent predictive factor (Table 3.2)

4. Discussion

Determining the expression levels of 10 genes by quantitative RT-PCR allowed us to establish a prediction model for recurrence in the form of liver metastasis after surgical treatment of colorectal cancer. The prediction accuracy rate in the training set was 77.0%, whilst the overall accuracy rate was 86.2% as validated in an independent test set. The sensitivity and specificity were 87.5% and 85.7%, respectively.

Several previous studies have attempted to use gene-expression profiles to construct a predictive model for liver metastasis after colorectal cancer.[8–12] However, these studies were based on DNA microarray data alone, and none confirmed gene expression of selected genes by quantitative methods such as RT-PCR. In addition, the numbers of colorectal cancer samples were small (20–86)[8–12] or the results were not validated in a different test set of patients.[9–11] Furthermore, as with all microarray analyses, they required a large number of genes (46–119).[11,12]

By contrast, quantitative RT-PCR analysis has been found to be useful in predicting the outcome of patients with malignancies such as lymphoma and lung cancer.[14–18] These studies have shown that RT-PCR is superior to microarray analysis in terms of reproducibility for quantitative analyses of gene expression, and that a smaller number of genes is required for prediction.[14,15] Therefore, in the present study, we established a prediction model for liver metastasis after colorectal cancer based on RT-PCR analysis, with an accuracy of 86.1% in a training set. The predictive performance of the model was evaluated by a permutation-based procedure (false-finding

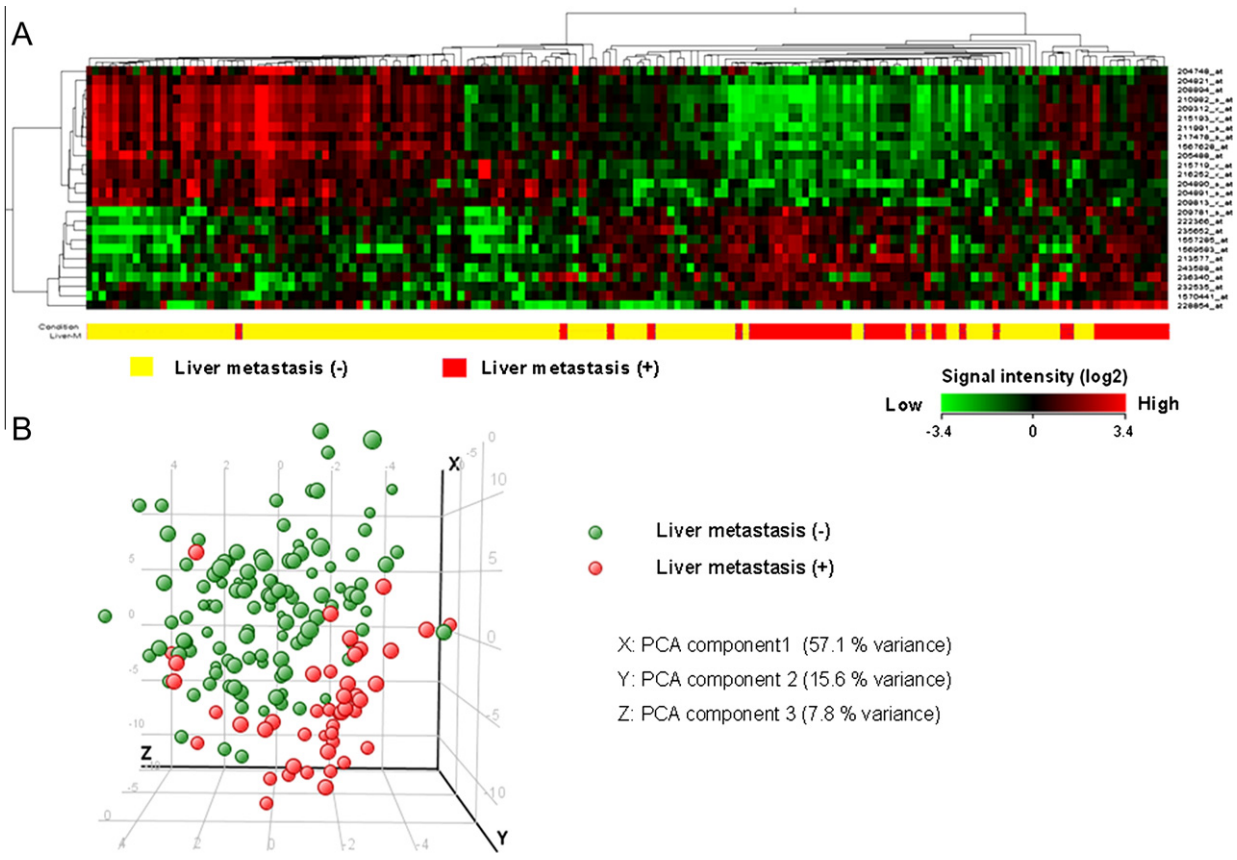
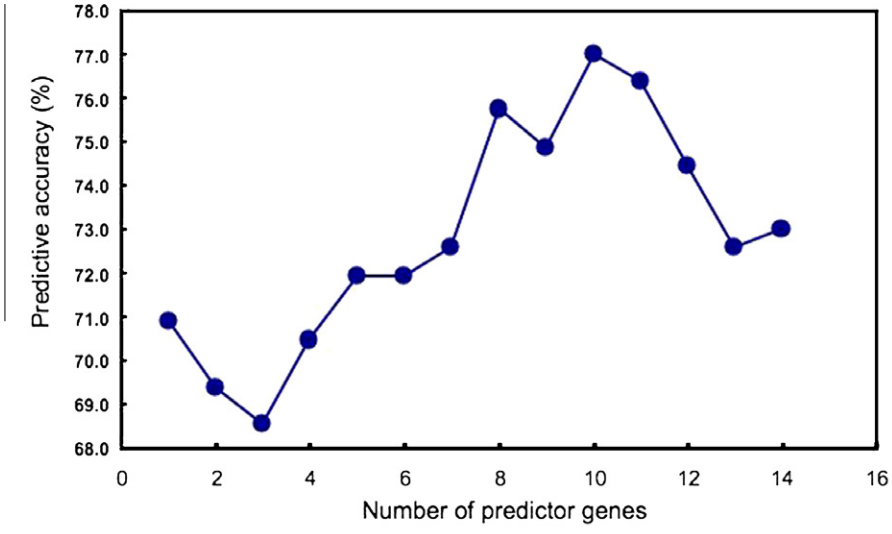


Fig. 1 – (A) Two-way hierarchical clustering analysis used to order samples (columns) and array targets (rows). (B) PCA used to generate a three-dimensional data plot from 26 probes.



No. of predictor genes	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Accuracy (%)	70.9	69.4	68.6	70.5	71.9	71.9	72.6	75.7	74.9	77.0	76.4	74.5	72.6	73.0

Fig. 2 – Predictive accuracy of genes. Predictive models were established based on different numbers of genes (14, 13 and so on) to determine the number that provided the best separation between patients with and without liver metastasis. The best accuracy rate was obtained when the 10 top-ranked genes were used.

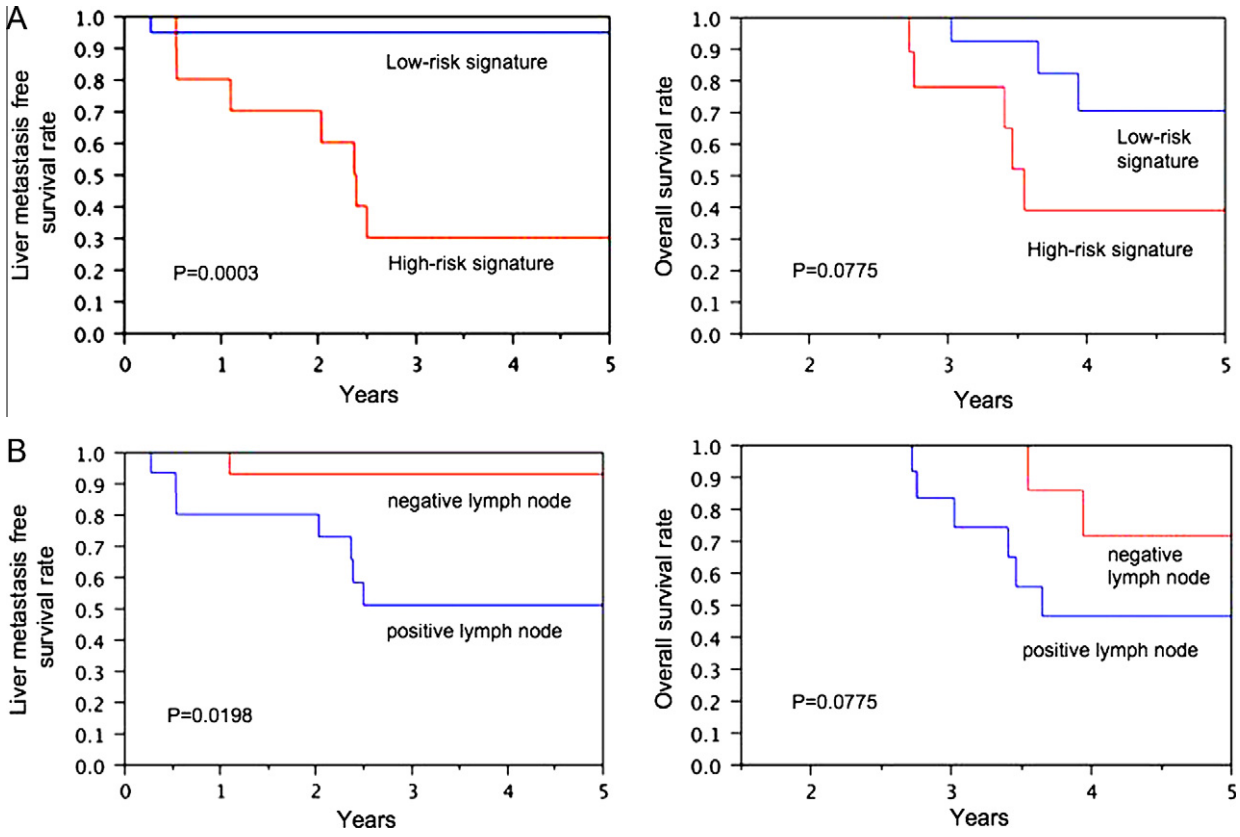


Fig. 3 – Kaplan–Meier analysis of liver-metastasis-free survival rate and overall survival rate. (A) Patients with high-risk signatures had a significantly lower liver metastasis-free rate than those with low-risk signatures ($p = 0.0003$). (B) Patients with lymph-node metastasis (lymph node-positive) had a significantly lower liver metastasis-free rate than those without lymph-node metastasis (lymph node-negative) ($p = 0.0198$).

Table 3.1 – Univariate proportional-hazard analysis of liver metastases risk.

Variable	Hazard ratio	95% CI*	*p* value
Prognosis signature (high-risk versus low-risk signature)	4.17	1.75–18.09	0.0005
Age (>65.1 versus ⩽65.1)	3.37	0.77–22.99	0.1087
Gender (male versus female)	1.02	0.15–4.45	0.9778
Tumour site (colon versus rectum)	1.16	0.24–4.72	0.8420
Tumour size (>45 mm versus ⩽45 mm)	0.94	0.19–3.83	0.9302
Tumour stage (T1-3 versus T4)	1.59	0.23–6.95	0.5898
Histological type (well versus others)	1.28	0.19–5.59	0.7662
Venous invasion (positive versus negative)	2.05	0.47–13.99	0.3554
Lymphatic invasion (positive versus negative)	1.64	0.34–6.67	0.5116
Lymph-node metastases (positive versus negative)	8.12	1.44–151.90	0.0146
Number of lymph nodes examined (>12 versus ⩽12)	0.79	0.18–5.38	0.7746
Number of positive lymph nodes (>4 versus ⩽4)	0.86	0.19–1.35	0.6549
Adjuvant chemotherapy** (treatment versus no treatment)	3.01	0.74–14.72	0.1244

* CI: confidence interval.

** Adjuvant chemotherapy: oral fluorouracil for one year after surgery.

rate = 0.0028). In total, 10 genes were selected that demonstrated significant differences in expression between patients with and without liver metastasis, and which provided the best predictive accuracy of 86.2%.

Lymph-node metastasis and venous invasion have previously been identified as prognostic factors for colorectal cancer.[21–26] In the present study, univariate analysis showed that the prognostic signature of the predictive model and the presence of lymph-node metastasis were significant factors associated with liver metastasis; however, only the prognostic signature remained significant following multivariate analysis. The liver metastasis-free survival analysis also showed a significant difference according to the prediction model and the presence of lymph-node metastasis. The reason

Table 3.2 – Multivariate proportional-hazard analysis of liver metastases risk.

Variable	Hazard ratio	95% CI*	*p* value
Prognosis signature (high-risk versus low-risk signature)	6.98	1.78–53.5	0.0027
Age (>65.1 versus ⩽65.1)	3.51	0.46–49.31	0.2399
Venous invasion (positive versus negative)	0.39	0.04–4.60	0.4352
Lymph-node metastases (positive versus negative)	7.49	0.92–164.98	0.0613
Adjuvant chemotherapy (treatment versus no treatment)**	0.35	0.03–2.92	0.3295

* CI: confidence interval.
** Adjuvant chemotherapy: oral fluorouracil for one year after surgery.

why adjuvant chemotherapy was not a significant factor in the present study may be due to a small number of patients. These findings might be useful in clinical settings; the present model helps to select patients who are at high-risk of developing liver metastasis and who should receive intensive follow-up and adjuvant chemotherapy.

The list of 14 genes (17 probes) the expression of which differed significantly between patients with and without liver metastasis both in the microarray and RT-PCR analyses included epiregulin (*EREG*), amphiregulin (*AREG*), lymphocyte-specific protein tyrosine kinase (*LCK*), tumour necrosis factor-receptor superfamily member 6 (*TNFRSF6*) and prostaglandin-endoperoxide synthase 2 (*PTGS2*). *EREG* and *AREG* are ligands for the epidermal growth factor receptor (EGFR) and previous studies have shown that they play an important role in the development of metastasis in malignancies, such as lung, breast and colorectal cancers.[27–32] In the present study, both genes showed higher expression in patients with liver metastasis, suggesting that they might play an important role in the development of liver metastasis in colorectal cancer. *LCK* plays a key role in the activation of T cells and is also expressed in many cancers. Indeed, *LCK* was recently included in a five-gene signature identified by RT-PCR that predicts clinical outcomes in non-small-cell lung cancer.[17] In the present study, *LCK* was selected as one of the final 10 genes used to predict the development of liver metastasis by colorectal cancer. Loss of *Fas* (*TNFRSF6*) expression correlates with disease progression and metastasis in oesophageal, colorectal and breast cancers.[33–35] Our results showed that *Fas* expression was reduced in patients with liver metastasis by both microarray and RT-PCR analysis. Expression of *PTGS2* (also known as cyclooxygenase-2 or *COX-2*) has also been shown to associate with survival in various cancers.[36–40] *PTGS2* (*COX-2*) was included amongst the final predictive genes in the present study.

There are some limitations to the current work. As the aim was to predict recurrence in the liver, only metachronous liver metastasis cases should have been included in our study. However, eliminating patients with synchronous liver metastasis would have reduced the total number of cases; therefore, we included patients with metachronous liver metastasis who also had synchronous metastasis in the training set. In the test set, we selected patients with metachronous metastasis alone and could predict liver metastasis with an accuracy of 86.2%. These results imply that patients with synchronous and metachronous liver metastasis might share a common genetic signature. Another point is the proportion of the test set. In the present study, the number of patients in the test set was smaller than in the training set. A large number of test set patients are ideal for the validation of the predictive model. However, on the other hand, Michilels S showed that in DNA microarray analysis the predictive ability improves with large training set sizes.[41] Therefore, in the present study, we assigned a higher proportion of patients to the training set than to the test set.

In conclusion, the 10-gene expression signature is closely associated with the development of liver metastasis after surgical resection of colorectal cancer. This signature might be useful for selecting patients that require adjuvant treatment and intensive follow-up. To our knowledge, this is the first study to establish a prediction model for the recurrence of colorectal cancer in the form of liver metastasis based on gene expression by RT-PCR analysis.

Conflict of interest statement

None declared.

Acknowledgements

The authors thank Eriko Hashimoto and Riyo Kakimoto for technical and secretarial support.

This study was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan and by a grant from the Ministry of Health, Labor and Welfare of Japan.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ejca.2010.04.019.

REFERENCES

1. Weitz J, Koch M, Debus J, et al. Colorectal cancer. *Lancet* 2005;**365**:153–65.
2. Andre T, Boni C, Mounedji-Boudiaf L, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. *N Engl J Med* 2004;**350**:2343–51.
3. Kuebler JP, Wieand HS, O'Connell MJ, et al. Oxaliplatin combined with weekly bolus fluorouracil and leucovorin as

surgical adjuvant chemotherapy for stage II and III colon cancer: results from NSABP C-07. *J Clin Oncol* 2007;**25**:2198–204.
4. Watanabe T, Wu TT, Catalano PJ, et al. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. *N Engl J Med* 2001;**344**:1196–206.
5. Watanabe T, Komuro Y, Kiyomatsu T, et al. Prediction of sensitivity of rectal cancer cells in response to preoperative radiotherapy by DNA microarray analysis of gene-expression profiles. *Cancer Res* 2006;**66**:3370–4.
6. Watanabe T, Kobunai T, Toda E, et al. Gene expression signature and the prediction of ulcerative colitis-associated colorectal cancer by DNA microarray. *Clin Cancer Res* 2007;**13**:415–20.
7. Watanabe T, Kobunai T, Toda E, et al. Distal colorectal cancers with microsatellite instability (MSI) display distinct gene expression profiles that are different from proximal MSI cancers. *Cancer Res* 2006;**66**:9804–8.
8. Li M, Lin YM, Hasegawa S, et al. Genes associated with liver metastasis of colon cancer, identified by genome-wide cDNA microarray. *Int J Oncol* 2004;**24**:305–12.
9. Bertucci F, Salas S, Eysteries S, et al. Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters. *Oncogene* 2004;**23**:1377–91.
10. D'Arrigo A, Belluco C, Ambrosi A, et al. Metastatic transcriptional pattern revealed by gene expression profiling in primary colorectal carcinoma. *Int J Cancer* 2005;**115**:256–62.
11. Ki DH, Jeung HC, Park CH, et al. Whole genome analysis for liver metastasis gene signatures in colorectal cancer. *Int J Cancer* 2007;**121**:2005–12.
12. 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* 2007;**30**:129–38.
13. Ding Y, Xu L, Jovanovic BD, et al. The methodology used to measure differential gene expression affects the outcome. *J Biomol Tech* 2007;**18**:321–30.
14. Lossos IS, Czerwinski DK, Alizadeh AA, et al. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *N Engl J Med* 2004;**350**:1828–37.
15. Chen HY, Yu SL, Chen CH, et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med* 2007;**356**:11–20.
16. Endoh H, Tomida S, Yatabe Y, et al. Prognostic model of pulmonary adenocarcinoma by expression profiling of eight genes as determined by quantitative real-time reverse transcriptase polymerase chain reaction. *J Clin Oncol* 2004;**22**:811–9.
17. Chen CN, Lin JJ, Chen JJ, et al. Gene expression profile predicts patient survival of gastric cancer after surgical resection. *J Clin Oncol* 2005;**23**:7286–95.
18. Takata R, Katagiri T, Kanehira M, et al. Predicting response to methotrexate, vinblastine, doxorubicin, and cisplatin neoadjuvant chemotherapy for bladder cancers through genome-wide gene expression profiling. *Clin Cancer Res* 2005;**11**:2625–36.
19. Liu R, Wang X, Chen GY, et al. The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* 2007;**356**:217–26.
20. Cover TM, Hart PE. Nearest neighbor pattern classification. IEEE Trans Inf Theory 1967;**IT-13**:21–7.
21. Le Voyer TE, Sigurdson ER, Hanlon AL, et al. Colon cancer survival is associated with increasing number of lymph nodes analyzed: a secondary survey of intergroup trial INT-0089. *J Clin Oncol* 2003;**21**:2912–9.
22. Chang GJ, Rodriguez-Bigas MA, Skibber JM, et al. Lymph node evaluation and survival after curative resection of colon cancer: systematic review. *J Natl Cancer Inst* 2007;**99**:433–41.
23. Johnson PM, Malatjalian D, Porter GA. Adequacy of nodal harvest in colorectal cancer: a consecutive cohort study. *J Gastrointest Surg* 2002;**6**:883–90.
24. Ohji Y, Yao T, Eguchi T, et al. Evaluation of risk of liver metastasis in colorectal adenocarcinoma based on the combination of risk factors including CD10 expression: multivariate analysis of clinicopathological and immunohistochemical factors. *Oncol Rep* 2007;**17**: 525–30.
25. Krasna MJ, Flancbaum L, Cody RP, et al. Vascular and neural invasion in colorectal carcinoma. Incidence and prognostic significance. *Cancer* 1988;**61**:1018–23.
26. Talbot IC, Ritchie S, Leighton MH, et al. The clinical significance of invasion of veins by rectal cancer. *Br J Surg* 1980;**67**:439–42.
27. Gupta GP, Nguyen DX, Chiang AC, et al. Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. *Nature* 2007;**446**:765–70.
28. Eltarhouny SA, Elsawy WH, Radpour R, et al. Genes controlling spread of breast cancer to lung "gang of 4". *Exp Oncol* 2008;**30**:91–5.
29. Shigeishi H, Higashikawa K, Hiraoka M, et al. Expression of epiregulin, a novel epidermal growth factor ligand associated with prognosis in human oral squamous cell carcinomas. *Oncol Rep* 2008;**19**:1557–64.
30. Berasain C, Garcia-Trevijano ER, Castillo J, et al. Amphiregulin: an early trigger of liver regeneration in mice. *Gastroenterology* 2005;**128**:424–32.
31. Brown CL, Meise KS, Plowman GD, et al. Cell surface ectodomain cleavage of human amphiregulin precursor is sensitive to a metal-loprotease inhibitor. Release of a predominant *N*-glycosylated 43-kDa soluble form. *J Biol Chem* 1998;**273**:17258–68.
32. Yamada M, Ichikawa Y, Yamagishi S, et al. Amphiregulin is a promising prognostic marker for liver metastases of colorectal cancer. *Clin Cancer Res* 2008;**14**:2351–6.
33. Botti C, Buglioni S, Benevolo M, et al. Altered expression of FAS system is related to adverse clinical outcome in stage I–II breast cancer patients treated with adjuvant anthracycline-based chemotherapy. *Clin Cancer Res* 2004;**10**:1360–5.
34. Ogawa S, Nagao M, Kanehiro H, et al. The breakdown of apoptotic mechanism in the development and progression of colorectal carcinoma. *Anticancer Res* 2004;**24**:1569–79.
35. Mottolese M, Buglioni S, Bracalenti C, et al. Prognostic relevance of altered Fas (CD95)-system in human breast cancer. *Int J Cancer* 2000;**89**:127–32.
36. Fux R, Schwab M, Thon KP, et al. Cyclooxygenase-2 expression in human colorectal cancer is unrelated to overall patient survival. *Clin Cancer Res* 2005;**11**:4754–60.
37. Masunaga R, Kohno H, Dhar DK, et al. Cyclooxygenase-2 expression correlates with tumor neovascularization and prognosis in human colorectal carcinoma patients. *Clin Cancer Res* 2000;**6**:4064–8.
38. Sheehan KM, Sheahan K, O'Donoghue DP, et al. The relationship between cyclooxygenase-2 expression and colorectal cancer. *JAMA* 1999;**282**:1254–7.
39. Zhang H, Sun XF. Overexpression of cyclooxygenase-2 correlates with advanced stages of colorectal cancer. *Am J Gastroenterol* 2002;**97**:1037–41.
40. Helms MW, Kemming D, Pospisil H, et al. Squalene epoxidase, located on chromosome 8q24.1, is upregulated in 8q+ breast cancer and indicates poor clinical outcome in stage I and II disease. *Br J Cancer* 2008;**99**:774–80.
41. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;**365**:488–92.