World Journal of Gastroenterology

Online Submissions: http://www.wjgnet.com/1007-9327office
wjg@wjgnet.com
doi:10.3748/wjg.v17.i15.2028

World J Gastroenterol 2011 April 21; 17(15): 2028-2036
ISSN 1007-9327 (print) ISSN 2219-2840 (online)

BRIEF ARTICLE

SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK are prognosis-related in colorectal cancer

Shu-Jing Yu, Jie-Kai Yu, Wei-Ting Ge, Han-Guang Hu, Ying Yuan, Shu Zheng

Shu-Jing Yu, Department of Radiation Oncology (Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education, Key Laboratory of Molecular Biology in Medical Science, Zhejiang Province, China), the Second Affiliated Hospital, Zhejiang University College of Medicine, Hangzhou 310009, Zhejiang Province, China
Jie-Kai Yu, Wei-Ting Ge, Shu Zheng, Cancer Institute (Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education, Key Laboratory of Molecular Biology in Medical Science, Zhejiang Province, China), the Second Affiliated Hospital, Zhejiang University College of Medicine, Hangzhou 310009, Zhejiang Province, China
Han-Guang Hu, Ying Yuan, Department of Medical Oncology, the Second Affiliated Hospital, Zhejiang University College of Medicine, Hangzhou 310009, Zhejiang Province, China
Author contributions: Yu SJ performed the majority of experiments and wrote the manuscript; Yu JK performed the bioinformatics analyses and was also involved in editing the manuscript; Ge WT was involved in editing the manuscript and analyzing experiment data; Yuan Y and Hu HG provided the collection of all the human material; Zheng S designed the study and edited the manuscript.
Supported by Grants from the Major State Basic Research Development Program of China, 973 Program No. 2004CB518707; the Zhejiang Provincial Natural Science Foundation of China, No. R2090353; the Fundamental Research Funds for the Central Universities, No. KYJD09007
Correspondence to: Shu Zheng, Professer, Cancer Institute, the Second Affiliated Hospital, Zhejiang University College of Medicine, 88 Jiefang Road, Hangzhou 310009, Zhejiang Province, China. zhengshu@zju.edu.cn
Telephone: +86-571-87784501 Fax: +86-571-87214404
Received: November 10, 2010 Revised: December 17, 2010
Accepted: December 24, 2010
Published online: April 21, 2011

Abstract

AIM: To investigate the expression of markers that are correlated with the prognosis of colorectal cancer (CRC) patients.

METHODS: One hundred and fifty-six CRC patients were followed up for more than 3 years after radical surgery. Immunohistochemical (IHC) analysis was performed to detect the expression of 14 pathway-related markers (p53, APC, p21ras, E-cadherin, endothelin-B receptor, Shp2, ADCY-2, SPARCL1, neuroligin1, hsp27, mmp-9, MAPK, MSH2 and rho) in specimens from these patients. Bioinformatics analysis involving a Support Vector Machine (SVM) was used to determine the best prognostic model from combinations of these markers.

RESULTS: Seven markers (SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK) were significantly related to the prognosis and clinical pathological features of the CRC patients ($P < 0.05$). Prognostic models were established through SVM from combinations of these 7 markers and proved able to differentiate patients with dissimilar survival, especially in stage Ⅱ/Ⅲ patients. According to the best prognostic model, the p53/SPARCL1 model, patients having high p53 and low SPARCL1 expression had about 50% lower 3-year survival than others ($P < 0.001$).

CONCLUSION: SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK are potential prognostic markers in CRC. A p53/SPARCL1 bioinformatics model may be used as a supplement to tumor-nodes-metastasis staging.

Key words: Colorectal cancer; Prognosis; SPARCL1; p53; Bioinformatics

Peer reviewers: Patrick O'Dwyer, MB, BCh, BAO, FRCS, MCh, FRCS (Glasg), University Department of Surgery, Western Infirmary, Glasgow, G11 6NT, United Kingdom; Antonio Basoli, Professor, General Surgery "Paride Stefanini", Università di Roma-Sapienza, Viale del Policlinico 155, Roma 00161, Italy; Dr. Benjamin Perakath, Professor, Department of Surgery Unit 5, Christian Medical College, Vellore 632004, Tamil Nadu, India

Yu SJ, Yu JK, Ge WT, Hu HG, Yuan Y, Zheng S. SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK are

prognosis-related in colorectal cancer. *World J Gastroenterol* 2011; 17(15): 2028-2036 Available from: URL: http://www.wjgnet.com/1007-9327/full/v17/i15/2028.htm DOI: http://dx.doi.org/10.3748/wjg.v17.i15.2028

INTRODUCTION

The incidence and mortality of colorectal cancer (CRC) are in the forefront of all cancers in western developed countries[1]. In China, the incidence of CRC has also increased in recent years, with 177 000 new cases and 99 000 deaths every year and a 5-year survival rate of 63.4%[2]. Tumor-nodes-metastasis (TNM) staging is helpful in predicting the survival of most patients. However, the heterogeneity of patients in their clinical outcome and their response to adjuvant chemotherapy calls for more useful prognostic pooled/panel molecular markers that will provide evidence for the choice of adjuvant therapy, especially for stage Ⅱ and Ⅲ patients.

Recently, Parsons *et al*[3] analyzed DNA mutations in CRC patients, and found that genetic changes of tumors are based on signaling pathways. Additionally, Wood *et al*[4] listed the number of mutations of all 140 genes included in 38 groups or pathways, which provided an impetus for the ongoing research on markers in CRC. After ranking these genes and pathways by the number of mutations listed by Wood *et al*[4], we selected three genes (p53, APC, ras) and 11 pathways which included genes with several mutations, and 14 genes were ultimately chosen from the pathways as candidate markers for our study. The genes are p53, APC, p21ras, E-cadherin, endothelin-B receptor, Shp2, ADCY-2, SPARCL1, neuroligin1, hsp27, mmp-9, MAPK, MSH2 and rho.

To identify prognosis-related markers of CRC, 156 patients who were followed up for more than 3 years after radical surgery were included in our survey. Immunohistochemical (IHC) analysis was performed to individually detect the expression of the 14 candidate markers in the specimens. The survival status of these patients was also analyzed. We found that seven tumor markers were found to be significantly related to the prognosis and clinical pathological features of these patients.

With the rapid development of the life sciences, bioinformatics has been developed and applied to collect, deposit and analyze large datasets and screen for useful information. In order to select molecular biomarkers more intelligently, we used a bioinformatics tool, the Support Vector Machine (SVM) classifier, to discriminate patients with different prognoses. SVM is based on the principles of Structure Risk Minimization and Vapnik-Chervonenkis Dimension as statistical learning theory, and thus provides a good generalization control[5]. SVM applications are actively used in various areas, from face recognition to genomics[6], and SVM is also a powerful tool for analyzing multiple markers. In this study, the seven prognostic markers were randomly combined, and SVM was used to evaluate which combination model was the best for predicting the prognosis of CRC patients.

MATERIALS AND METHODS

Ethics

This work has been carried out in accordance with the Declaration of Helsinki (2000) of the World Medical Association and approved by the Ethics Committee of the Second Affiliated Hospital of Zhejiang University, College of Medicine, along with the patients' informed consent.

Patients and specimens

Tumor specimens included in this study were from 156 CRC patients who underwent a radical resection operation in the Second Affiliated Hospital of Zhejiang University, College of Medicine, between 1999 and 2004, with a median age of 60 years (range 20-92 years) at diagnosis. The clinical data of all patients are presented in Table 1. Tumor specimens for IHC were from filed blocks in the histopathological department.

Living patients were all followed up for > 36 mo after the radical operation, with a median follow-up of 62 mo (range 36-108 mo). The follow-ups were performed by history and physical surveillance every 3-6 mo for 2 years, then every 6 mo up to 5 years and every year after 5 years (conforming to NCCN V.2.2010). No patient was lost during the follow-up.

IHC

All 156 specimens in paraffin blocks were made into tissue arrays using a ZM-1 tissue array machine[7]. Sections (4-μm thick) were cut, and immunostaining for each antigen was conducted using the avidin-biotin peroxidase complex technique (MaxVision™ HRP-Polymer IHC Kit, MAIXIN-Bio), following the manufacturer's instructions. The antibodies used were p53 (monoclonal mouse, ZhongShan), APC (polyclonal rabbit, ZhongShan), p21ras (monoclonal mouse, MAIXIN), E-cadherin (monoclonal mouse, ZhongShan), endothelin-B receptor (polyclonal rabbit, CHEMICON), Shp2 (monoclonal rabbit, Abcam), ADCY-2 (monoclonal rabbit, Abcam), SPARCL1 (polyclonal goat, R & D), neuroligin (polyclonal rabbit, CHEMICON), HSP27 (monoclonal mouse, ZhongShan), mmp9 (polyclonal rabbit, ZhongShan), ERK1 + ERK2 (monoclonal mouse, ZhongShan), MSH2 (monoclonal mouse, ZhongShan) and Rho(-A,-B,-C) (monoclonal rabbit, MILLIPORE).

The IHC results were assessed using a semi-quantitative system, as previously described[8]. According to the percentages of positive cells (0: none, 1: < 25%, 2: 25%-50%, 3: 50%-75% and 4: > 75%) and staining intensity (0: negative, 1: weak, 2: moderate and 3: strong), the expression levels of the proteins were divided into four groups by the sum of the two scores above: 0 (0, negative expression), 1 (2-3, low expression), 2 (4-5, medium expression) and 3 (6-7, high expression).

Bioinformatics analysis

Experimental data were then analyzed by the Zhejiang University ProteinChip Data Analysis System (ZUCIPDAS, www.zlzx.net). We constructed a non-linear SVM classifier (with a radial based function kernel, a parameter Gamma of

Table 1 Clinicopathologic data of patients

Terms	*n* (%)
Sex	
Male	85 (54.5)
Female	71 (45.5)
Location	
Right hemicolon	45 (30.1)
Transverse colon	3 (1.9)
Left hemicolon	8 (5.8)
Sigmoid colon	32 (20.5)
Rectum	67 (41.7)
Differentiation	
Well	95 (60.9)
Moderately	40 (25.6)
Poorly	17 (10.9)
Unknown	4 (2.6)
Bowel wall invasion (pT)	
T1	7 (4.5)
T2	30 (19.2)
T3	116 (74.4)
T4	3 (1.9)
Lymph node metastasis (pN)	
N0	82 (52.6)
N1	43 (27.5)
N2	31 (19.9)
Distant metastasis (pM)	
M0	144 (92.3)
M1	12 (7.7)
TNM staging	
Ⅰ	29 (18.6)
Ⅱ	52 (33.3)
Ⅲ	63 (40.4)
Ⅳ	12 (7.7)
Post-surgery event	
Recurrence or metastasis	51 (32.7)
Survival status	
Dead	51 (32.7)[1]
Alive	105 (67.3)

[1]Among patients who have died, 40 patients died from recurrence or metastasis, while 11 patients died from causes such as heart or lung failure, or reasons unknown. TNM: Tumor-nodes-metastasis.

0.6, and a cost of the constraint violation of 19) to distinguish groups with different prognoses, and validated results by a 10-fold cross validation method.

One hundred and thirty-one patients with complete data were then filtered for the ongoing bioinformatics analysis. The seven prognostic biomarkers (SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK) were combined randomly to build 127 SVM models. For each model, the expression of these markers was the input, and the 3-year survival status of each patient was the evaluation criteria. The model with the highest accuracy for predicting the 3-year survival of the patients was selected as the best prognostic model, and the accuracy of the models was then validated by 10-fold cross validation between training sets and test sets.

Of the 131 patients, 44 died within 3 years after surgery, but the other 87 were still alive after 3 years. Because the number of patients dead at 3 years was about half of the number of living ones, the model showed low sensitivity due to the unbalanced data. Obviously, sensitivity is important for a prognostic model, so we next defined an adjusted accuracy [accuracy = (sensitivity+specificity)/2 + sensitivity]/2; sensitivity = true positive/(true positive + false negative), specificity = true negative/(true negative + false positive). This increased the weight of sensitivity and allowed SVM to select models with higher sensitivity.

Table 2 Expression of candidate markers in 156 colorectal cancer patients

Markers	Numbers of patients with different expression[1]			
	Negative	Low	Medium	High
P53	42	36	32	40
APC	80	37	21	10
MAPK	114	27	8	0
E-cadherin	44	47	36	22
Mmp9	109	40	6	1
Hsp27	96	33	18	6
MSH2	17	52	47	39
P21ras	102	43	8	2
ADCY-2	45	67	36	3
Shp2	108	34	13	0
ETB	59	60	30	5
Neuroligin	53	52	43	4
Rho	28	68	45	9
SPARCL1	23	52	61	20

The immunohistochemical (IHC) results were assessed using a semi-quantitative system, as previously described[8]. According to the percentages of positive cells and staining intensity, the expression levels of the proteins were divided into four groups as negative, low, medium and high expression. [1]Data were missing because some specimens were lost during sectioning and staining of tissue arrays.

Statistical analyses

Kaplan-Meier survival analysis (log-rank test) was used to evaluate the relationship between marker expression and the survival of patients. Kruskal-Wallis test was used to evaluate the relationship between the expression of candidate markers and some pathologic features in IHC analyses. SPSS Version 13.0 software (SPSS Inc., Chicago, IL) was used for all statistical analyses. $P < 0.05$ was considered to be statistically significant, and all P values were two-sided.

RESULTS

Association between the expression of candidate markers and the survival of CRC patients

The expression of candidate markers in CRC was investigated by IHC (listed in Table 2). It should be noted that some specimens were lost during sectioning and staining of tissue arrays, resulting in an average of 3.6 specimens per marker (2.3%). Representative examples of immunoassayed slides for each marker are shown in Figure 1.

Kaplan-Meier survival analysis revealed that markers significantly related with survival were SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK. The higher protein expression of SPARCL1, Shp2, MSH2, E-cadherin, and MAPK in CRC patients was related to better survival, while the higher expression of p53 and ADCY-2 was related to worse survival. The Kaplan-Meier survival curves of these markers are shown in Figure 2.

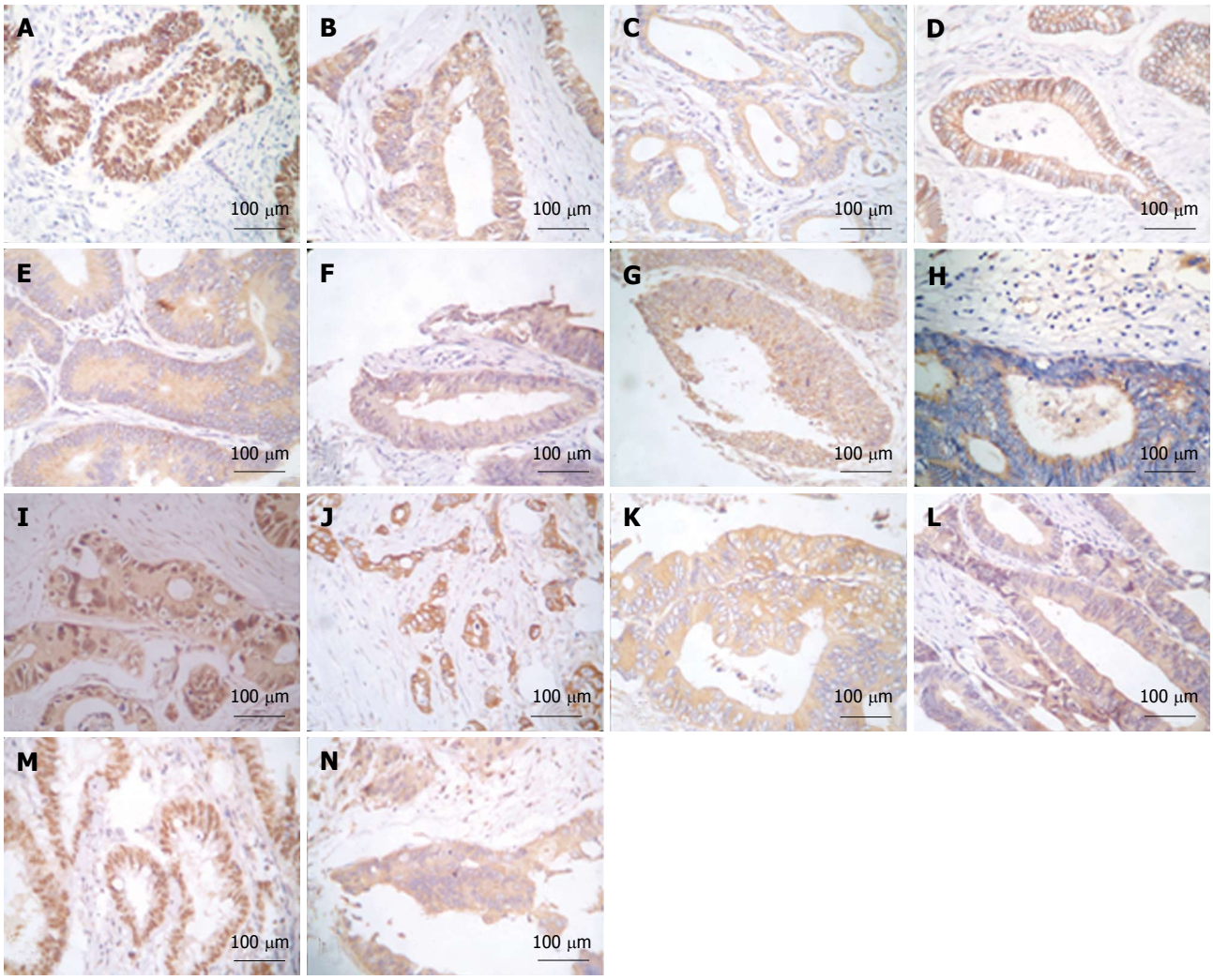

Figure 1 Immunohistochemical expression of 14 markers. The markers and the cellular location of positive staining are listed below: A: P53: nuclear; B: APC: cytoplasm; C: P21ras: cytoplasm; D: E-cadherin: membrane or cytoplasm; E: Endothelin B receptor: cytoplasm; F: Shp2: cytoplasm; G: ADCY-2: cytoplasm; H: SPARCL1: cytoplasm; I: neuroligin1: nuclear or cytoplasm; J: hsp27: nuclear or cytoplasm; K: MMP9: cytoplasm; L: MAPK: cytoplasm; M: MSH2: nuclear; N: Rho: cytoplasm. (Scale bar = 100 μm).

Kruskal-Wallis tests revealed that among these markers, SPARCL1, Shp2 and MSH2 were noticeably associated with the most clinical pathological features of CRC patients, including differentiation, bowel wall invasion (pT), lymph node metastasis (pN), distant metastasis (pM), TNM stage, post-surgery recurrence or metastasis. P53 was mainly related to TNM staging; E-cadherin and MAPK were mainly related to post-surgery recurrence and metastasis (Table 3). However, other markers, such as endothelin B receptor, APC and rho, were just related to differentiation or stages (data not shown).

Prognostic bioinformatics model established by combining the seven markers and evaluated by survival analysis

By SVM, the seven markers can randomly form 127 combinations. After being validated by 10-fold cross validation, the model with the highest accuracy (65.3) was the p53/SPARCL1 combination among all these combinations.

According to the prediction result (PR) given by the p53/SPARCL1 model, patients can be divided into two groups: "high risk" (PR > 0) and "low risk" (PR < 0). Three-year survival of the low risk group (88.30%) was more than twice as high as that of the high risk group (37.84%). Kaplan-Meier analysis revealed that the difference of survival was significant between these two groups ($P < 0.001$) (Figure 3A).

Prognostic value of the p53/SPARCL1 model for stage II and III CRC patients

Among these 131 patients, 99 patients were classified as stage Ⅱ or Ⅲ. We found that the difference in 3-year survival was not great between stage Ⅱ ($n = 43$) and Ⅲ ($n = 56$) patients, i.e. 88.40% *vs* 62.50% ($P = 0.039$) (Figure 3B). However, when these 99 patients were grouped by the PR of the p53/SPARCL1 model, the 3-year survival rate was very different between the low risk ($n = 70$) and high risk ($n = 29$) groups, i.e. 87.14% *vs* 37.93% ($P < 0.001$) (Figure 3C). Thus, the survival difference was much greater between low and high risk groups than between stage Ⅱ and

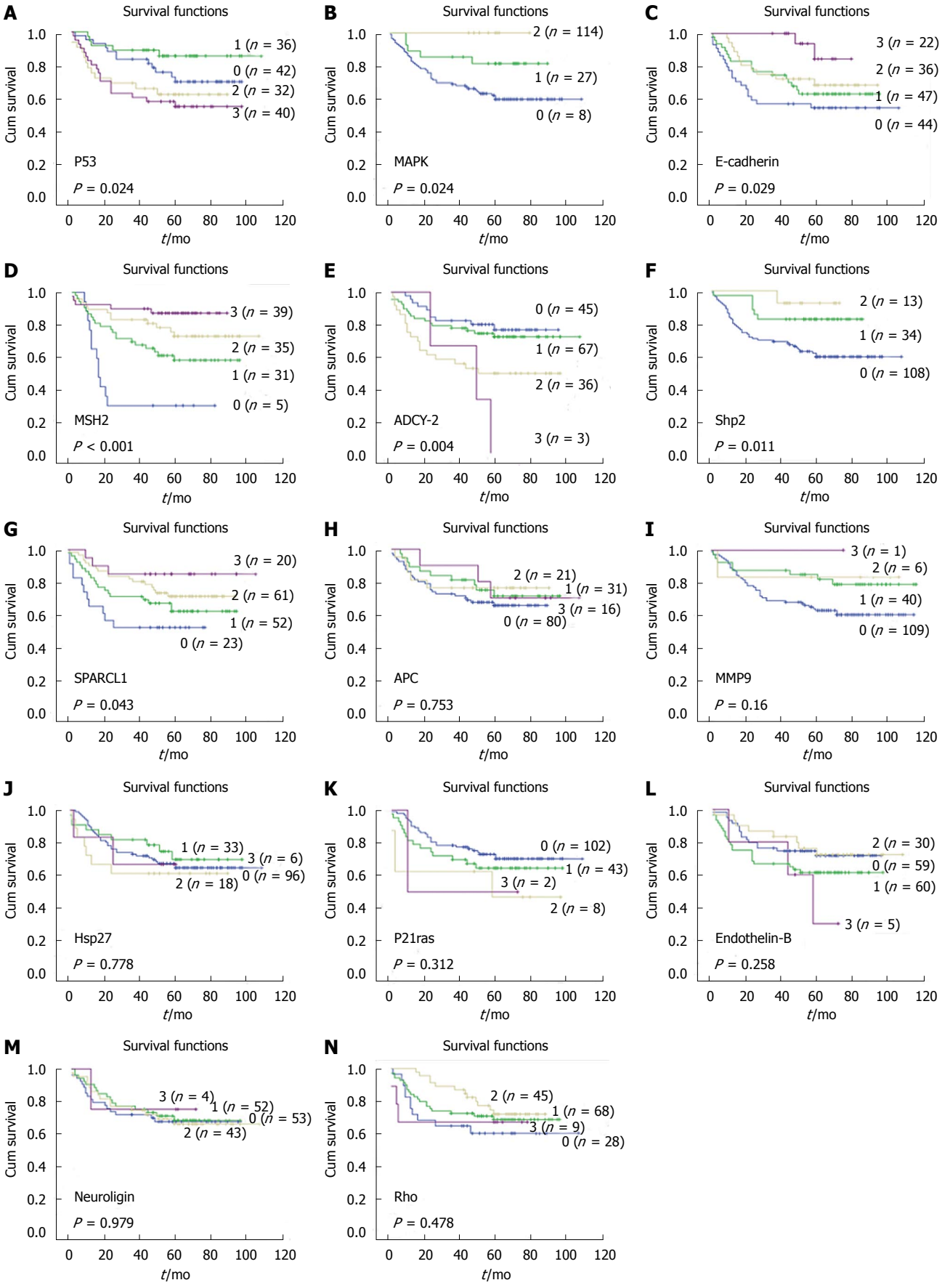

Figure 2 Kaplan-Meier curves of 14 markers. A: P53; B: MAPK; C: E-cadherin; D: MSH2; E: ADCY-2; F: Shp2; G: SPARCL1; H: APC; I: MMP9; J: Hsp27; K: P21ras; L: endothelin-B receptor; M: Neuroligin1; N: Rho.

Table 3 Relationship between marker expression and clinical features (*P* values)

Markers	*P* values					
	Differentiation	pT	pN	pM	TNM	Post-surgery
P53	0.671	0.654	0.003[b]	0.119	0.024[a]	0.207
MAPK	0.186	0.597	0.230	0.188	0.182	0.001[b]
E-cadherin	0.028[a]	0.342	0.080	0.041[a]	0.223	0.004[b]
MSH2	0.964	0.006[b]	0.012[a]	0.061	0.001[b]	0.001[b]
ADCY-2	0.458	0.430	0.779	0.470	0.878	0.082
Shp2	0.020[a]	0.004[b]	0.035[a]	0.849	0.006[b]	0.006[b]
SPARCL1	0.002[b]	0.171	0.037[a]	0.021[a]	0.044[a]	0.014[a]

The relationship between the expression of candidate markers and some pathologic features was evaluated by Kruskal-Wallis test (SPSS Version 13.0 software) in immunohistochemistry analyses. All *P* values are two-sided ([a]$P < 0.05$, [b]$P < 0.01$). The pathologic features in the table were differentiation, bowel wall invasion (pT), lymph node metastasis (pN), distant metastasis (pM), TNM stage (TNM), post-surgery recurrence or metastasis (Post-surgery).

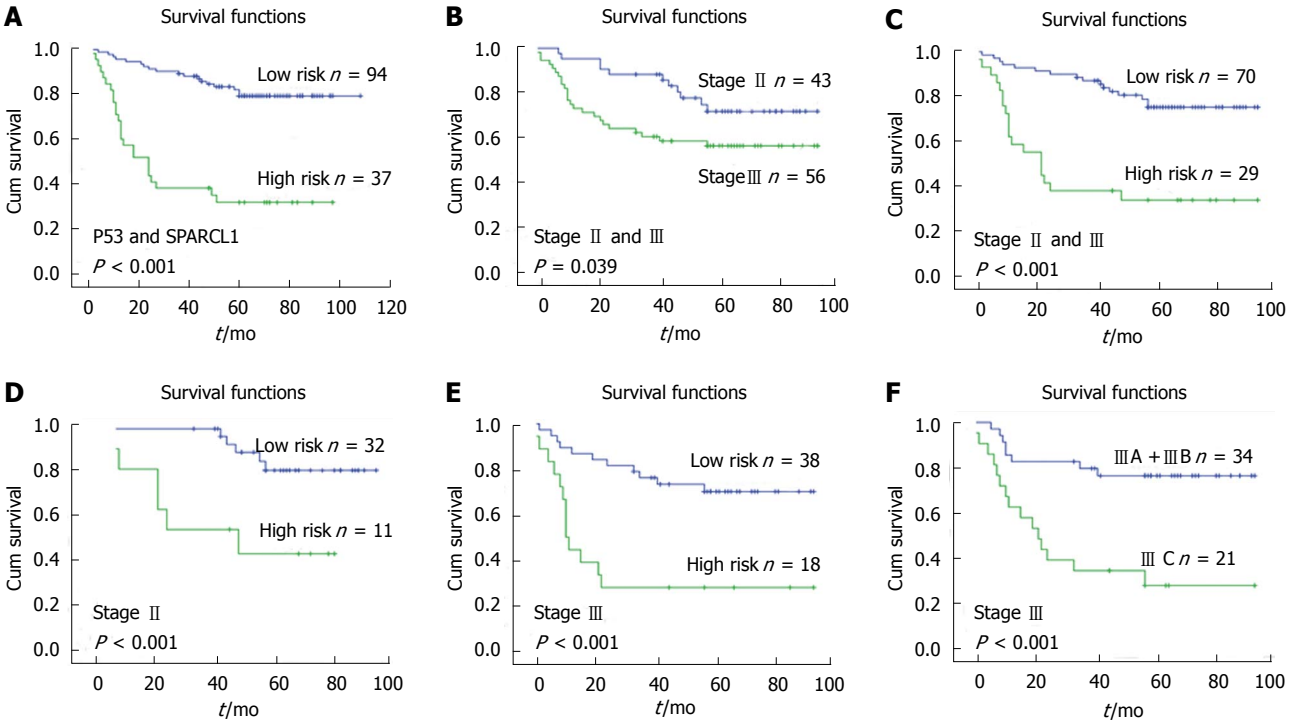

Figure 3 Prognostic value of p53/SPARCL1 model in colorectal cancer patients. According to the prediction result (PR) given by the p53/SPARCL1 model, patients could be divided into two groups: "high risk" (PR > 0) and "low risk" (PR < 0). A: 3-year survival of the "low risk" group was 88.30%, significantly higher at twice that of the "high risk" group, which was only 37.84% ($P < 0.001$). B: The 3-year survival of stage Ⅱ ($n = 43$) and Ⅲ ($n = 56$) patients was 88.40% *vs* 62.50% ($P = 0.039$), with an only 15.90% survival difference ($P = 0.039$); C: The same 99 stage Ⅱ/Ⅲ patients, when divided by the PR of the p53/SPARCL1 model: the 3-year survival of "low risk" ($n = 70$) and "high risk" ($n = 29$) group was 87.14% vs 37.93%, with a survival difference of 49.21% ($P < 0.001$), much more than the difference between stage Ⅱ and Ⅲ patients; D: According to the PR of the p53/SPARCL1 model, the 3-year survival of "low risk" and "high risk" patients at stage Ⅱ was 100% and 54.55%, respectively, with a significant difference of 45.45% ($P < 0.001$); E: At stage Ⅲ ($n = 56$), the 3-year survival was 78.95% of "low risk" patients and 27.78% of "high risk" patients, with a 51.17% higher survival rate ($P < 0.001$); F: At stage Ⅲ ($n = 56$), the 3-year survival was different between stage ⅢA/ⅢB ($n = 34$) and ⅢC ($n = 22$) patients: 82.36% *vs* 31.82% ($P < 0.001$).

Ⅲ patients.

Among the 99 stage Ⅱ/Ⅲ patients, 43 patients were of stage Ⅱ, all of whom were classified as stage ⅡA. According to the PR of the p53/SPARCL1 model, the 3-year survival rates of low risk and high risk stage Ⅱ patients were 100% and 54.55%, respectively. This 45.45% difference between the survival rates of low risk and high risk stage Ⅱ patients was significant ($P < 0.001$) (Figure 3D).

Among the 56 stage Ⅲ patients, the 3-year survival was 78.95% for low risk patients and 27.78% for high risk ones, and this 51.17% difference was statistically significant ($P < 0.001$) (Figure 3E). Similar survival difference was found between stage ⅢA/ⅢB ($n = 34$) and ⅢC ($n = 22$) patients, i.e. 82.36% *vs* 31.82% in 3-year survival rates ($P < 0.001$) (Figure 3F).

DISCUSSION

Which CRC patients should receive adjuvant chemotherapy after radical resection? Currently, it is a standard recommendation for stage Ⅲ but not stage Ⅱ patients. However, the 5-year survival rate of stage ⅡB (T4N0M0)

patients is even lower than stage ⅢA (T1-2N1M0)[9]. One explanation for this may be due to not dissecting enough lymph nodes during surgery. Another potential cause is that stage ⅡB tumors penetrate to the surface of the visceral peritoneum or directly invade the adjacent organs, which indicates that the biological behavior of the tumor is poor. Further, we do not know which of the stage Ⅱ patients are at high risk and should receive adjuvant chemotherapy to improve their survival. Therefore, better molecular tumor markers are urgently needed to predict which patients may potentially benefit from adjuvant chemotherapy.

In the network of cancer-related genes, pathways are the frame by which we can understand the network logically. In the present study, 14 candidate markers were selected based on the most frequently mutated genes and pathways listed in the study of Wood *et al*[4], and their expression levels in CRC specimens were detected by IHC, which is generally used for regular pathological detection. Among these 14 markers, seven markers (SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK) were significantly prognosis-related.

Shp2 is an essential component in several oncogene signaling pathways[10]. Here, we surprisingly found that Shp2 is a predictive marker for good prognosis, which is in stark contrast to previous studies indicating a role for Shp2 in promoting carcinogenesis in other cancers[11-13]. MSH2 is a vital mismatch repair gene. Patients with high MSH2 expression had better survival in CRC[14,15], and higher gene expression of MSH2 in responders to 5-fluorouracil-based chemotherapy indicates a predictive value of MSH2 in chemotherapy[16,17]. The MAPK signal pathway is associated with proliferation, survival and apoptosis of tumor cells and therefore plays a very important role in carcinogenesis[18,19]. E-cadherin, a member of the cadherins, is related to invasion and metastasis in many cancers[20]. Loss or low expression of E-cadherin is more frequent in CRC patients with liver metastasis[21], demonstrating that loss of E-cadherin is related to poor prognosis. ADCY is involved in the G-protein system-related GnRH signal pathway. The proliferation of rat pancreatic tumoral AR4-2J cells can be stimulated by pituitary ADCY-activating peptide through the ADCY pathway[22,23], which suggests that ADCY promotes the growth of tumor cells. Among these markers in the present study, Shp2, MSH2, MAPK and E-cadherin were significant markers for predicting good prognosis, but ADCY-2 was not.

P53 is an indispensible tumor suppressor that plays an important role in several carcinogenic processes. Previous studies suggest that p53 has an influence on the prognosis of patients in many cancers[24] including CRC, and is associated with tumor staging, multi-drug resistance, response to chemotherapy or radiotherapy, post-surgery recurrence and metastasis[25-30]. Recently, research has shown that mutant p53 proteins not only lose their tumor suppressive functions but may also gain new abilities that enhance tumorigenesis[31]. Indeed, the p53 mutation is linked with chemo-resistance and transformation to a more aggressive disease in many tumor types[32]. The p53 codon 72 polymorphism causes an increased risk for liver metastases in CRC patients positive for p53 overexpression[33]. In the present study, we found that a high expression of mutant p53 protein was associated with more frequent lymph node metastasis, advanced TNM stage and poor survival (Table 2), which is consistent with other reports.

SPARCL1, also known as hevin[34], belongs to the matricellular protein family. SPARCL1 is down-regulated in transformed prostate epithelial cell line P69SV40T[35,36], and tissues of metastatic prostate adenocarcinoma, non-small cell lung cancer, bladder and pancreatic ductal carcinoma, but up-regulated in liver cancer tissues[35-39]. Additional work by our group has revealed that SPARCL1 expression is significantly different between CRC specimens with and without liver metastasis (to be published). In the present study, the expression of SPARCL1 was not only significantly associated with histological differentiation and survival but also with distant and lymph node metastasis, suggesting that SPARCL1 is likely to be an important negative regulator in the progression or metastasis of CRC.

In the present study, SVM was utilized to analyze and establish prognostic models of CRC from the combinations of the 7 prognostic biomarkers mentioned above. For SVM, the right balance is struck between the accuracy attained on a particular training set and the "capacity" of the machine, i.e. the ability of the machine to learn any training set without error to achieve the best generalization ability. The remarkably robust performance of SVM with respect to sparse and noisy data has made it the system of choice in a number of applications. When used for classification, SVM separates a given set of binary labeled training data and can work in combination with the kernels technique for cases in which no linear separation is possible. The accuracy of our models was evaluated by 10-fold cross validation.

Ultimately, the combination of p53 and SPARCL1 was found to be the best prognostic model of those tested. Survival analysis proved that the prediction result of the p53/SPARCL1 model was a statistically significant prognostic factor for CRC patients in all stages or only stage Ⅱ/Ⅲ (Figure 3).

Other researchers have attempted to identify biomarkers to further stratify stage Ⅱ or stage Ⅲ patients. Prognostic advantages were found in patients with MSI-high tumors and stage Ⅱ and Ⅲ CRC patients treated with 5-fluorouracil-based adjuvant therapy[40,41]. In the PETACC-3 study, the prognostic value of MSI status was found to be more significant in patients with stage Ⅱ disease than in stage Ⅲ cases[42]. However, value of MSI status as a prognostic or predictive marker may be affected by mutations in other genes involved in cancer etiology, such as the BRAF gene[43]. Additionally, chromosome 18q loss of heterozygosity (LOH) has been associated with poor prognosis in stage Ⅱ and stage Ⅲ CRC patients in some studies[40,44] but not others[45,46]. Differences in the methodologies used possibly explained the contradictory

findings reported. In a large prospective study of patients with non-MSI-high CRC, 18q LOH was also not associated with patient survival, indicating that 18q LOH is not an independent survival predictive marker[47].

In our study, according to the p53/SPARCL1 model, the survival rate of low risk stage ⅡA patients was 45.45% higher than that of high risk ones. Moreover, low risk stage Ⅲ patients had a 51.17% higher 3-year survival rate than high risk ones ($P < 0.001$), the same as the survival difference between stage ⅢA/ⅢB and ⅢC. Therefore, the p53/SPARCL1 model established in this study can likely be used to supplement TNM staging, especially in stage Ⅱ and Ⅲ patients.

In conclusion, we discovered that SPARCL1, Shp2, MSH2, E-cadherin, p53, ADCY-2 and MAPK are significant prognostic markers in CRC. The p53/SPARCL1 model is of predictive use in discriminating patients with high or low risk, especially at stage Ⅱ and Ⅲ. Patients may benefit from accurate valuation and realistic treatment strategies for their disease with the help of potential prognostic markers. Larger scale studies and those involving multiple centers are planned to confirm clinic applicability of this prognosis model.

COMMENTS

Background

The incidence and mortality of colorectal cancer (CRC) are in the forefront of all cancers in China and western developed countries. More useful prognostic markers are urgently needed to provide evidence for the strategy of adjuvant therapy, especially for stage Ⅱ and Ⅲ patients.

Research frontiers

There are many ways to find useful markers in cancers. Cancer genomic research from the Vogelstein group has provided an enormous amount of information on genetic alterations in colorectal cancers. In this study, the authors set out to utilize this information to help in choosing their candidate markers. Moreover, the bioinformatics tool, which is used more and more to screen useful information from a large data set, was used here to further build prognostic models in CRC patients.

Innovations and breakthroughs

Some biomarkers have been identified as being related to the prognosis of CRC patients, including MSH2, E-cadherin and p53. However, this is the first study to report that SPARCL1, Shp2, ADCY-2 and MAPK are also potential prognostic markers in CRC. Furthermore, survival analysis proved that the p53/SPARCL1 model, established by the bioinformatics tool, could differentiate CRC patients with different prognoses in all stages or only stage Ⅱ/Ⅲ.

Applications

By finding significant prognostic markers in CRC, patients could be discriminated with high or low risk, especially in stage Ⅱ and Ⅲ. Patients may benefit from accurate valuation and realistic treatment strategies for their disease with the help of potential prognostic markers and models.

Terminology

The Support Vector Machine (SVM) classifier is a kind of bioinformatics tool, which is considered to be powerful for identifying the best discriminator from a large data set. Therefore, SVM applications are actively used in various areas from face recognition to genomics and SVM is also a powerful tool for analyzing multiple markers.

Peer review

This paper reports some novel findings on patient outcome with colorectal cancer. The array of genetic markers is extensive and worthy of publication.

REFERENCES

1 **Espey DK**, Wu XC, Swan J, Wiggins C, Jim MA, Ward E, Wingo PA, Howe HL, Ries LA, Miller BA, Jemal A, Ahmed F, Cobb N, Kaur JS, Edwards BK. Annual report to the nation on the status of cancer, 1975-2004, featuring cancer in American Indians and Alaska Natives. *Cancer* 2007; **110**: 2119-2152

2 **Jemal A**, Clegg LX, Ward E, Ries LA, Wu X, Jamison PM, Wingo PA, Howe HL, Anderson RN, Edwards BK. Annual report to the nation on the status of cancer, 1975-2001, with a special feature regarding survival. *Cancer* 2004; **101**: 3-27

3 **Parsons DW**, Wang TL, Samuels Y, Bardelli A, Cummins JM, DeLong L, Silliman N, Ptak J, Szabo S, Willson JK, Markowitz S, Kinzler KW, Vogelstein B, Lengauer C, Velculescu VE. Colorectal cancer: mutations in a signalling pathway. *Nature* 2005; **436**: 792

4 **Wood LD**, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, Shen D, Boca SM, Barber T, Ptak J, Silliman N, Szabo S, Dezso Z, Ustyanksky V, Nikolskaya T, Nikolsky Y, Karchin R, Wilson PA, Kaminker JS, Zhang Z, Croshaw R, Willis J, Dawson D, Shipitsin M, Willson JK, Sukumar S, Polyak K, Park BH, Pethiyagoda CL, Pant PV, Ballinger DG, Sparks AB, Hartigan J, Smith DR, Suh E, Papadopoulos N, Buckhaults P, Markowitz SD, Parmigiani G, Kinzler KW, Velculescu VE, Vogelstein B. The genomic landscapes of human breast and colorectal cancers. *Science* 2007; **318**: 1108-1113

5 **Vapnik VN**. Statistical Learning Theory. New York: Wiley, 1998

6 **Liu Y**. Active learning with support vector machine applied to gene expression data for cancer classification. *J Chem Inf Comput Sci* 2004; **44**: 1936-1941

7 **Meng PQ**, Hou G, Zhou GY, Peng JP, Dong Q, Zheng S. Application of new tissue microarrayer-ZM-1 without recipient paraffin block. *J Zhejiang Univ Sci B* 2005; **6**: 853-858

8 **Elkhuizen PH**, Hermans J, Leer JW, van dE Vijver MJ. Isolated late local recurrences with high mitotic count and early local recurrences following breast-conserving therapy are associated with increased risk on distant metastasis. *Int J Radiat Oncol Biol Phys* 2001; **50**: 387-396

9 **O'Connell JB**, Maggard MA, Ko CY. Colon cancer survival rates with the new American Joint Committee on Cancer sixth edition staging. *J Natl Cancer Inst* 2004; **96**: 1420-1425

10 **Chan RJ**, Feng GS. PTPN11 is the first identified proto-oncogene that encodes a tyrosine phosphatase. *Blood* 2007; **109**: 862-867

11 **Kathpalia VP**, Mussak EN, Chow SS, Lam PH, Skelley N, Time M, Markelewicz RJ Jr, Kanduc D, Lomas L, Xiang Z, Sinha AA. Genome-wide transcriptional profiling in human squamous cell carcinoma of the skin identifies unique tumor-associated signatures. *J Dermatol* 2006; **33**: 309-318

12 **Tao XH**, Shen JG, Pan WL, Dong YE, Meng Q, Honn KV, Jin R. Significance of SHP-1 and SHP-2 expression in human papillomavirus infected Condyloma acuminatum and cervical cancer. *Pathol Oncol Res* 2008; **14**: 365-371

13 **Zhou X**, Coad J, Ducatman B, Agazie YM. SHP2 is up-regulated in breast cancer cells and in infiltrating ductal carcinoma of the breast, implying its involvement in breast oncogenesis. *Histopathology* 2008; **53**: 389-402

14 **Losi L**, Ponti G, Gregorio CD, Marino M, Rossi G, Pedroni M, Benatti P, Roncucci L, de Leon MP. Prognostic significance of histological features and biological parameters in stage I (pT1 and pT2) colorectal adenocarcinoma. *Pathol Res Pract* 2006; **202**: 663-670

15 **Lanza G**, Gafà R, Santini A, Maestri I, Guerzoni L, Cavazzini L. Immunohistochemical test for MLH1 and MSH2 expression predicts clinical outcome in stage II and III colorectal cancer patients. *J Clin Oncol* 2006; **24**: 2359-2367

16 **Jensen LH**, Danenberg KD, Danenberg PV, Jakobsen A. Predictive value of MSH2 gene expression in colorectal cancer treated with capecitabine. *Clin Colorectal Cancer* 2007; **6**: 433-435

17 **Bendardaf R**, Lamlum H, Ristamäki R, Syrjänen K, Pyrhönen S. Oncoprotein Bcl-2 and microsatellite instability are associated with disease-free survival and treatment response in colorectal cancer. *Oncol Rep* 2008; **20**: 999-1004

18 **Platanias LC**. Map kinase signaling pathways and hematologic malignancies. *Blood* 2003; **101**: 4667-4679
19 **Ajenjo N**, Cañón E, Sánchez-Pérez I, Matallanas D, León J, Perona R, Crespo P. Subcellular localization determines the protective effects of activated ERK2 against distinct apoptogenic stimuli in myeloid leukemia cells. *J Biol Chem* 2004; **279**: 32813-32823
20 **Beavon IR**. The E-cadherin-catenin complex in tumour metastasis: structure, function and regulation. *Eur J Cancer* 2000; **36**: 1607-1620
21 **Delektorskaya VV**, Perevoshchikov AG, Golovkov DA, Kushlinskii NE. Expression of E-cadherin, beta-catenin, and CD-44v6 cell adhesion molecules in primary tumors and metastases of colorectal adenocarcinoma. *Bull Exp Biol Med* 2005; **139**: 706-710
22 **Zia F**, Fagarasan M, Bitar K, Coy DH, Pisegna JR, Wank SA, Moody TW. Pituitary adenylate cyclase activating peptide receptors regulate the growth of non-small cell lung cancer cells. *Cancer Res* 1995; **55**: 4886-4891
23 **Buscail L**, Cambillau C, Seva C, Scemama JL, De Neef P, Robberecht P, Christophe J, Susini C, Vaysse N. Stimulation of rat pancreatic tumoral AR4-2J cell proliferation by pituitary adenylate cyclase-activating peptide. *Gastroenterology* 1992; **103**: 1002-1008
24 **Steele RJ**, Lane DP. P53 in cancer: a paradigm for modern management of cancer. *Surgeon* 2005; **3**: 197-205
25 **Oka M**, Kounoura K, Narasaki F, Sakamoto A, Fukuda M, Matsuo I, Ikeda K, Tsurutani J, Ikuno N, Omagari K, Mizuta Y, Soda H, Gudas JM, Kohno S. P-glycoprotein is positively correlated with p53 protein accumulation in human colorectal cancers. *Jpn J Cancer Res* 1997; **88**: 738-742
26 **Swisher SG**, Roth JA, Komaki R, Gu J, Lee JJ, Hicks M, Ro JY, Hong WK, Merritt JA, Ahrar K, Atkinson NE, Correa AM, Dolormente M, Dreiling L, El-Naggar AK, Fossella F, Francisco R, Glisson B, Grammer S, Herbst R, Huaringa A, Kemp B, Khuri FR, Kurie JM, Liao Z, McDonnell TJ, Morice R, Morello F, Munden R, Papadimitrakopoulou V, Pisters KM, Putnam JB Jr, Sarabia AJ, Shelton T, Stevens C, Shin DM, Smythe WR, Vaporciyan AA, Walsh GL, Yin M. Induction of p53-regulated genes and tumor regression in lung cancer patients after intratumoral delivery of adenoviral p53 (INGN 201) and radiation therapy. *Clin Cancer Res* 2003; **9**: 93-101
27 **Fujimoto K**, Yamada Y, Okajima E, Kakizoe T, Sasaki H, Sugimura T, Terada M. Frequent association of p53 gene mutation in invasive bladder cancer. *Cancer Res* 1992; **52**: 1393-1398
28 **Esrig D**, Elmajian D, Groshen S, Freeman JA, Stein JP, Chen SC, Nichols PW, Skinner DG, Jones PA, Cote RJ. Accumulation of nuclear p53 and tumor progression in bladder cancer. *N Engl J Med* 1994; **331**: 1259-1264
29 **Sarkis AS**, Bajorin DF, Reuter VE, Herr HW, Netto G, Zhang ZF, Schultz PK, Cordon-Cardo C, Scher HI. Prognostic value of p53 nuclear overexpression in patients with invasive bladder cancer treated with neoadjuvant MVAC. *J Clin Oncol* 1995; **13**: 1384-1390
30 **Noske A**, Lipka S, Budczies J, Müller K, Loddenkemper C, Buhr HJ, Kruschewski M. Combination of p53 expression and p21 loss has an independent prognostic impact on sporadic colorectal cancer. *Oncol Rep* 2009; **22**: 3-9
31 **Adhikari AS**, Iwakuma T. Mutant p53 gain of oncogenic function: in vivo evidence, mechanism of action and its clinical implications. *Fukuoka Igaku Zasshi* 2009; **100**: 217-228
32 **Al-Joudi FS**, Iskandar ZA, Rusli J. The expression of p53 in invasive ductal carcinoma of the breast: a study in the North-East States of Malaysia. *Med J Malaysia* 2008; **63**: 96-99
33 **Zhu ZZ**, Liu B, Wang AZ, Jia HR, Jin XX, He XL, Hou LF, Zhu GS. Association of p53 codon 72 polymorphism with liver metastases of colorectal cancers positive for p53 overexpression. *J Zhejiang Univ Sci B* 2008; **9**: 847-852
34 **Girard JP**, Springer TA. Cloning from purified high endothelial venule cells of hevin, a close relative of the antiadhesive extracellular matrix protein SPARC. *Immunity* 1995; **2**: 113-123
35 **Nelson PS**, Plymate SR, Wang K, True LD, Ware JL, Gan L, Liu AY, Hood L. Hevin, an antiadhesive extracellular matrix protein, is down-regulated in metastatic prostate adenocarcinoma. *Cancer Res* 1998; **58**: 232-236
36 **Schraml P**, Shipman R, Colombi M, Ludwig CU. Identification of genes differentially expressed in normal lung and non-small cell lung carcinoma tissue. *Cancer Res* 1994; **54**: 5236-5240
37 **Esposito I**, Kayed H, Keleg S, Giese T, Sage EH, Schirmacher P, Friess H, Kleeff J. Tumor-suppressor function of SPARC-like protein 1/Hevin in pancreatic cancer. *Neoplasia* 2007; **9**: 8-17
38 **Bendik I**, Schraml P, Ludwig CU. Characterization of MAST9/Hevin, a SPARC-like protein, that is down-regulated in non-small cell lung cancer. *Cancer Res* 1998; **58**: 626-629
39 **Lau CP**, Poon RT, Cheung ST, Yu WC, Fan ST. SPARC and Hevin expression correlate with tumour angiogenesis in hepatocellular carcinoma. *J Pathol* 2006; **210**: 459-468
40 **Popat S**, Houlston RS. A systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC status and colorectal cancer prognosis. *Eur J Cancer* 2005; **41**: 2060-2070
41 **Westra JL**, Schaapveld M, Hollema H, de Boer JP, Kraak MM, de Jong D, ter Elst A, Mulder NH, Buys CH, Hofstra RM, Plukker JT. Determination of TP53 mutation is more relevant than microsatellite instability status for the prediction of disease-free survival in adjuvant-treated stage III colon cancer patients. *J Clin Oncol* 2005; **23**: 5635-5643
42 **Roth AD**, Tejpar S, Delorenzi M, Yan P, Fiocca R, Klingbiel D, Dietrich D, Biesmans B, Bodoky G, Barone C, Aranda E, Nordlinger B, Cisar L, Labianca R, Cunningham D, Van Cutsem E, Bosman F. Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. *J Clin Oncol* 2010; **28**: 466-474
43 **French AJ**, Sargent DJ, Burgart LJ, Foster NR, Kabat BF, Goldberg R, Shepherd L, Windschitl HE, Thibodeau SN. Prognostic significance of defective mismatch repair and BRAF V600E in patients with colon cancer. *Clin Cancer Res* 2008; **14**: 3408-3415
44 **Lanza G**, Matteuzzi M, Gafá R, Orvieto E, Maestri I, Santini A, del Senno L. Chromosome 18q allelic loss and prognosis in stage II and III colon cancer. *Int J Cancer* 1998; **79**: 390-395
45 **Halling KC**, French AJ, McDonnell SK, Burgart LJ, Schaid DJ, Peterson BJ, Moon-Tasson L, Mahoney MR, Sargent DJ, O'Connell MJ, Witzig TE, Farr GH Jr, Goldberg RM, Thibodeau SN. Microsatellite instability and 8p allelic imbalance in stage B2 and C colorectal cancers. *J Natl Cancer Inst* 1999; **91**: 1295-1303
46 **Popat S**, Zhao D, Chen Z, Pan H, Shao Y, Chandler I, Houlston RS. Relationship between chromosome 18q status and colorectal cancer prognosis: a prospective, blinded analysis of 280 patients. *Anticancer Res* 2007; **27**: 627-633
47 **GJ**, Meyerhardt JA, Fuchs CS. Prognostic significance and molecular associations of 18q loss of heterozygosity: a cohort study of microsatellite stable colorectal cancers. *J Clin Oncol* 2009; **27**: 4591-4598

S- Editor Tian L **L- Editor** Logan S **E- Editor** Ma WH