Int J Clin Exp Med 2017;10(6):9009-9018
www.ijcem.com /ISSN:1940-5901/IJCEM0052269

Original Article

Identification of a cytokine-cytokine receptor interaction gene signature for predicting clinical outcomes in patients with colorectal cancer

Chuanpeng Dong[1*], Xing Wang[1*], Huilin Xu[1*], Xiaohui Zhan[2,3], He Ren[1,4], Zhenhao Liu[1], Gang Liu[1], Lei Liu[1]

[1]Shanghai Public Health Clinical Center, Institute of Biomedical Sciences, Fudan University, Shanghai 200032, China; [2]Key Lab of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China; [3]University of Chinese Academy of Sciences, Beijing, China; [4]School of Health Informatics Technology and Management, Shanghai University of Medicine & Health Sciences, Shanghai 200032, China. [*]Equal contributors.

Received March 6, 2017; Accepted April 7, 2017; Epub June 15, 2017; Published June 30, 2017

Abstract: The role of chronic inflammation in colorectal cancer development and progression is an issue of high interest. Cytokine and cytokine receptor interaction networks are regarded as crucial aspects of inflammation and tumor immunology, particularly for colorectal cancer. However, cytokine expression pattern genes have never been established for colorectal cancer prognosis prediction. Using a gene-mining approach with TCGA and GEO database, we identified an eight-gene signature in cytokine and cytokine receptor interaction signaling pathway, which is significantly associated with the colorectal patients' prognosis. Based on this eight-gene model, we classified patients for training sets into two groups, patients that with high-risk and low risk group, due to significant difference on overall survival (HR=2.72; P=0.00001) and RFS (HR=1.87; P=0.0007). The prognostic value of these eight cytokine-cytokine receptor interaction gene panels was validated in independent testing sets. Our results might provide an aid for colorectal cancer prognosis evaluation and preselect patients suitable for more effective therapy strategies.

Keywords: Prognosis, cytokines, gene signature, colorectal cancer

Introduction

Colorectal cancer (CRC) is one of the most common cancers in the world and also one of the leading causes of cancer-related death worldwide [1]. The role of chronic inflammation in colorectal cancer development and progression is an issue of high interest. Favorable prognosis in CRC patients was found associating with robust anti-tumor immunity, though CRC risk is significantly higher in patient with chronic intestinal inflammation [2].

Cytokines mixture produced in tumor microenvironment relevant to infection, and inflammation are known as important inflammatory effect mediators for oncogenesis both in inflammation-induced neoplasia and inflammation that follows tumor development [3]. Cytokine and cytokine receptor interaction networks are regarded as crucial effects on inflammation, as well as tumor immunology, particularly for colorectal cancer [4]. Cytokines and its receptors, such as tumor necrosis factor and interleukin-6, are regarded as central players in CRC. They drive key oncogenic transcription factors activation. STAT3 and NF-κB in intestinal epithelial cells have been reported to promote proliferation and function an anti-apoptotic role [5]. IL-11 induced by TGF-β-stimulated cancer-related fibroblasts could increase the efficiency of metastasis and colonization of CRC [6].

The correlation between inflammatory related cell pattern in CRC tumors and clinical outcome has further been elaborated as "Immunoscore" by Galon and his colleagues [7, 8]. Recently, gene signatures generated from microarray or RNA-seq data, such as Oncotype DX [9, 10], ColoPrint [11] and ColoGuideEx [12], have been

Table 1. Clinicopathologic characteristics of patients in three sets [N (%)]

Characteristic	TCGA (N=367)	GSE39582 (N=566)	GSE17536 (N=177)
Age (years)			
Median	66	68	66
Range	31-90	22-97	26-92
<65	170 (46.3)	211 (37.3)	78 (44.1)
≥65	197 (53.7)	354 (62.5)	99 (55.9)
Sex			
Male	202 (55.0)	310 (54.8)	96 (54.2)
Female	165 (45.0)	256 (45.2)	81 (45.8)
T stage			
T1-2	66 (18.0)	56 (10.0)	*NA*
T3-4	299 (81.5)	486 (85.9)	*NA*
N stage			
N0	201 (54.8)	302 (53.3)	*NA*
N1	97 (26.4)	134 (23.7)	*NA*
N2	67 (18.0)	98 (17.3)	*NA*
M stage			
M0	251 (68.4)	482 (85.2)	*NA*
M1	50 (13.6)	61 (10.8)	*NA*
MX	61 (16.6)	3 (0.5)	*NA*
Stage			
I	56 (15.3)	33 (5.8)	24 (13.6)
II	132 (36.0)	264 (46.6)	116 (65.5)
III	111 (30.2)	205 (36.2)	57 (32.2)
IV	51 (13.9)	60 (10.6)	39 (22)
Chemotherapy			
Yes	39 (10.6)	233 (41.2)	*NA*
No	30 (8.2)	316 (55.8)	*NA*
Unknown	298 (81.2)	17 (3.0)	*NA*

TCGA, The Cancer Genome Atlas; NA, Not Available.

used to identify patients with a more aggressive phenotype or poor outcomes in colorectal adenocarcinoma. Compared with conventional pathological criteria, genomic classifiers can provide more accurate risk of recurrence information and may facilitate selection of patients who can get more benefit from chemotherapy. However, these genomic-based molecular classification method might not help to develop patient pre-selection tool for specific inhibitor targeted therapy. Thus, the necessity to establish a pathway or multi pathway based gene expression signature has been raised.

In this study, we included genes that are related to all major cytokines and cytokine receptors, like interleukins (IL), chemokines, interferons (IFN) and transforming growth factor beta (TGF-β). By analyzing the association between gene expression profiling and clinical outcome of colorectal patients, we identified an eight-gene signature, which is associated with cytokine and cytokine receptor interaction and significantly correlated with the prognosis of colorectal patients. Our findings may provide useful aid for the development of novel colorectal cancer therapies.

Materials and methods

Data processing

Gene expression profile and corresponding clinical data were retrieved from the Cancer Genome Atlas (TCGA) thorough Cancer Genomics Browser of University of California Santa Cruz (UCSC) (https://genome-cancer.ucsc.edu) [13]. 367 colorectal cancer samples with detailed gene expression and survival information were selected for study [14]. Microarray studies from the GEO database are downloaded from GEO website (http://www.ncbi.nlm.nih.gov/geo). Two independent datasets GSE39582 (N=566) and GSE17536 (N=177) with either overall survival or recurrence-free survival information were selected as testing sets. Totally, 1110 colorectal cancer patients were analyzed and their clinical data are presented in **Table 1**.

264 genes associated with cytokine-cytokine receptor interaction were obtained from the KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.ad.jp/kegg), including chemokines, hematopoietins, Interferon family, TNF family, TNF-beta family and FDGF family members. **Figure 1** depicts the workflow of this study.

Statistical analysis

The correlation between gene expression and patient's overall survival was calculated with univariate Cox proportional hazards regression analysis. Genes with p value <0.01 were selected as candidate genes. Then the candidate genes were fitted in a multivariable Cox regression model in the TCGA colorectal cancer cohorts [15]. A risk score formula was constructed by including each of these selected

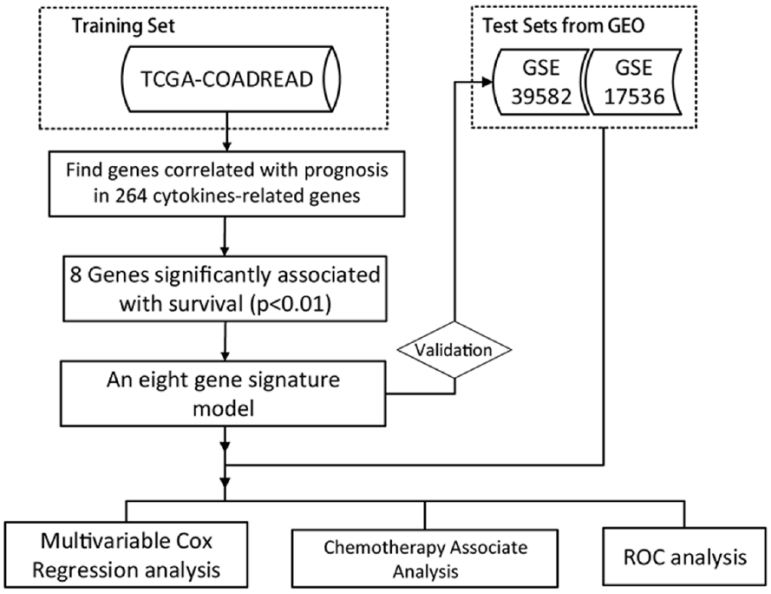

Figure 1. Workflow chart of our study. TGGA, the Cancer Genome Atlas; GEO, Gene Expression Ominus.

Table 2. Eight gene significantly associated with the overall survival in the training-set patients

Gene symbol	Full name	*p* value	Hazard ratio	Coefficient
VEGFC	Vascular endothelial growth factor C	0.00261	1.306	0.1323
TGFB3	Transforming growth factor beta-3	0.00453	1.195	-0.0280
TGFB2	Transforming growth factor-beta 2	0.00197	1.254	0.1255
LEP	Leptin	0.00055	1.199	0.0566
INHBC	Inhibin beta C	0.00725	1.549	0.3941
INHBB	Inhibin beta B	0.00662	1.172	0.0429
IL20RB	Interleukin 20 Receptor Beta	0.00055	1.340	0.2613
GDF5	Growth/differentiation factor 5	0.00822	1.182	0.0377

Coefficient, Coefficients in the multivariate cox regression analysis.

genes, weighted by their estimated regression coefficients from the multivariable Cox regression [16]. By using the median of risk scores as the cutoff point, patients were divided into high risk and low risk group. Kaplan-Meier analysis was executed to estimate the survival time in the two different groups [17, 18]. Log-rank test was used to evaluate the statistical significance between the two survival groups. A two-sided *P* value <0.05 was considered statistical significant.

Receiver operating characteristic (ROC) curves were used to estimate the sensitivity and specificity of the survival prediction of the genes risk model and other indicators. Area under the curve (AUC) values were measured for the ROC curves. All the data were analyzed with R 3.2.1 (www.rproject.org).

Results

Identification of prognostic cytokine and cytokine receptor interaction genes from the TCGA COADREAD data set

Totally, 264 genes involved in cytokine and cytokine receptor interaction pathway were identified, and mRNA expression profile of these genes were retracted from the TCGA and the GEO datasets. The TCGA colorectal cancer cohort was taken as training set, using for detecting prognostic genes. Eight cytokine receptor interaction genes showed significant associations with overall survival (P≤0.01). Hazard ratio of all these genes (GDF5, IL20RB, INHBB, INHBC, LEP, TGFB2, TGFB3 and VEGFC) were above 1, indicating that elevating mRNA expression of these genes were associated with poor prognosis. More details were described in **Table 2**.

The association of eight cytokine and cytokine receptor interaction gene signatures and patient's survival in training set and testing sets

As showed in materials and methods, the risk scores of colorectal cancer patient in training sets was calculated using the follow formula: Risk score = $\sum_{n=1}^{8}$ Gene (n) expression * coefficient (n). Patients were divided into low-risk and high-risk groups by using median as cut-off. The risk scores distribution, survival status and gene expression profiles for TCGA training sets were shown in **Figure 2**. Patients with high risk score revealed significant shorter overall survival time than those with low risk score (log-rank

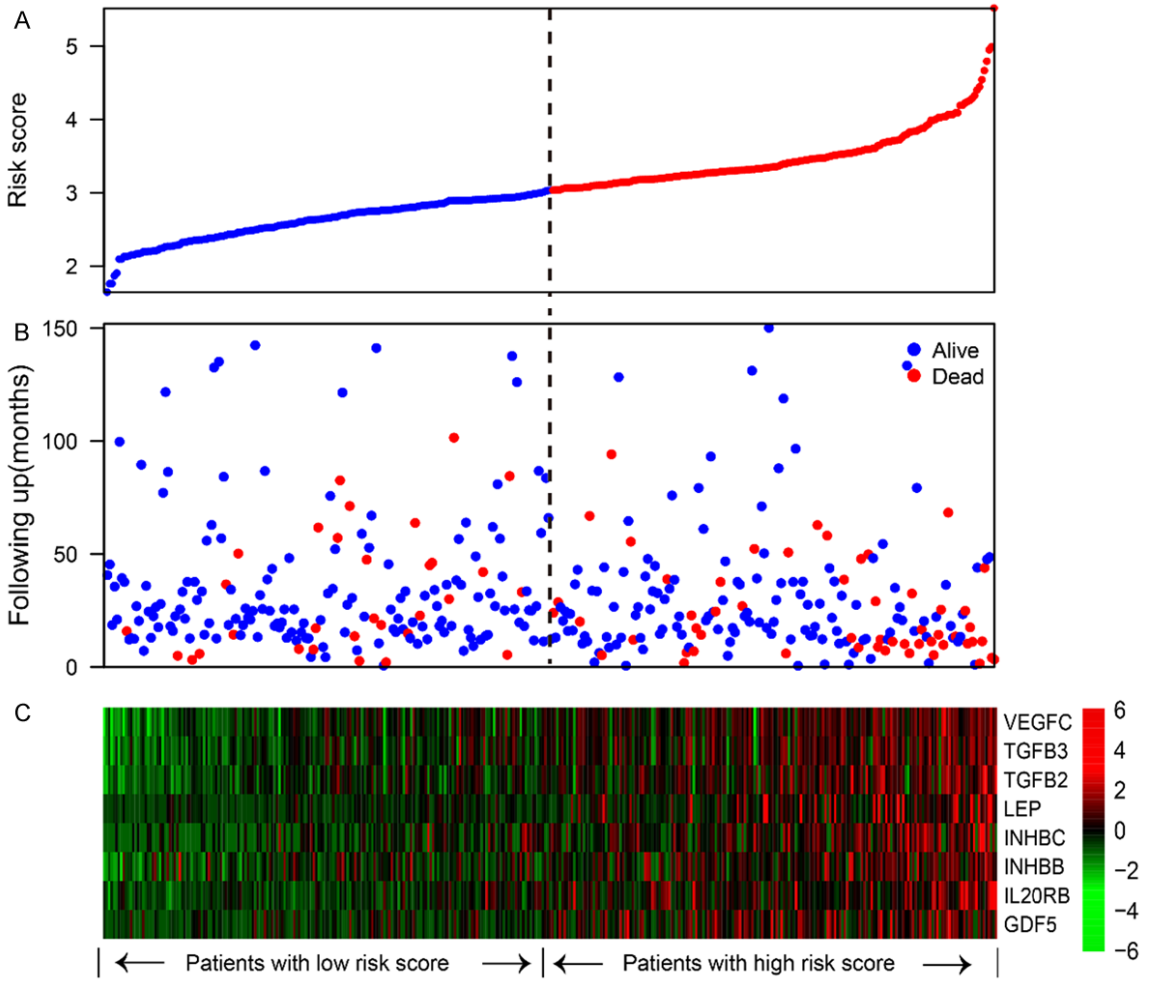

Figure 2. The risk scores distribution, survival status and gene expression profiles for TCGA training sets. A. Eight gene based risk score distribution; B. Overall survival status and survival time; C. Heatmap of the selected gene expression profiles.

P=0.00045) (**Figure 3A**). Apart from that, we also performed univariate Cox regression analysis to estimate the correlation of the eight gene risk score with overall survival (HR=2.72; P<0.00001) and recurrence free survival (HR= 1.87; P=0.00073).

Two colorectal cancer datasets GSE39582 and GSE17536 was used as the testing sets, using the same methods and confidents as training sets. In consistence with the results described above, patients with high risk score had shorter survival time than those with low risk score (**Figure 3B-F**). According to univariate Cox regression analysis, patients with high risk score were significantly associated with shorter overall survival (**Tables 3** and **4**).

Prognostic value of the eight gene expression signature is independent of clinical parameters

Further investigation was performed in training set, to assess whether our eight-gene model could predict prognosis independent of other clinical factor, such as age, gender stage and tumor primary site. The result showed that within female or male CRC patients, those with high risk score both likely to have poor prognosis (**Figure 4A** and **4B**); while high risk score can predict poor outcome in stage III&IV patient (**Figure 4D**), not in those of stage I&II (**Figure 4C**). We also found the risk scores were positive associated with AJCC stage (**Figure 4E**). More subtype analysis was shown in Figure S1. Receiver operating characteristic (ROC) analysis

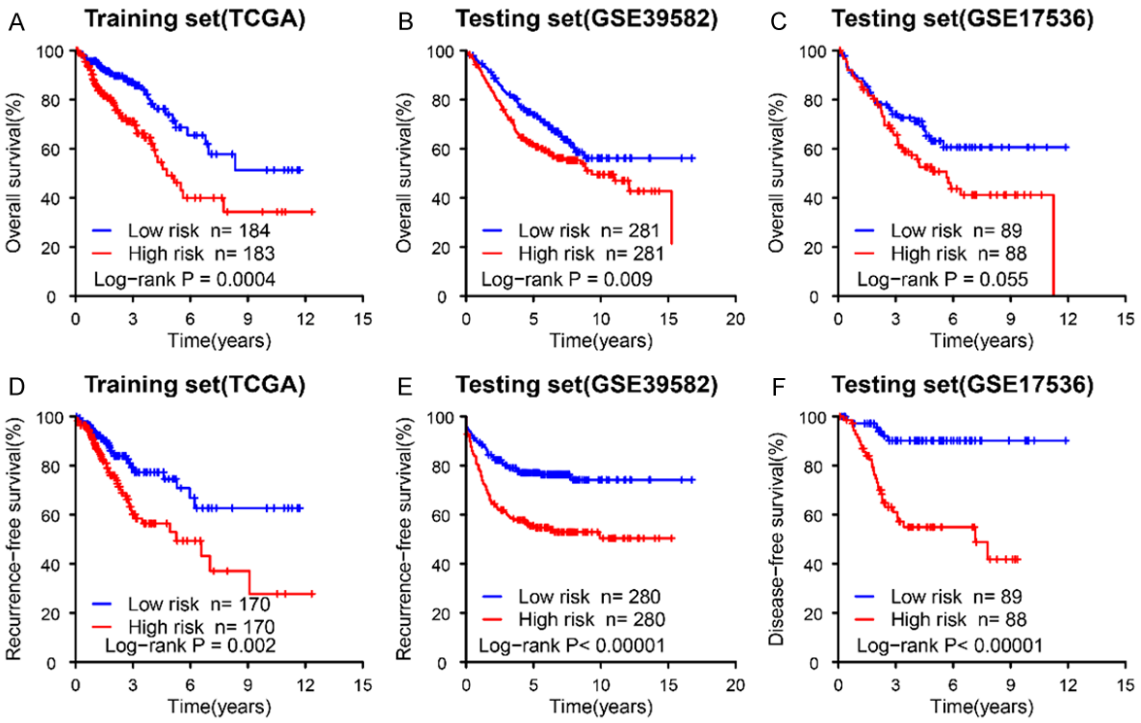

Figure 3. The association of eight-gene signatures and patient's survival in training set and testing sets. Kaplan-Meier curves of (A-C) overall survival of TCGA, GSE39582 and GSE17536; (D, E) Recurrence-free survival of TCGA and GSE39582; (F) Disease-free survival of GSE17536.

Table 3. Univariate and multivariate cox regression analyses of overall survival in each data set

Variables	Univariate model		*p* value	Multivariate model		*p* value
	HR	95% CI		HR	95% CI	
Training set (TCGA)						
8 gene risk score (high risk vs. low risk)	2.17	1.39-3.38	0.0006	1.89	1.17-3.05	0.009
Age (≥65 vs. <65)	2.15	1.34-3.44	0.0015	2.77	1.66-4.62	9.31E-05
Gender (male vs. female)	1.22	0.79-1.88	0.366	1.05	0.66-1.66	0.8459
Stage (I/II/III/IV)	1.94	1.49-2.52	6.55E-07	1.97	1.50-2.60	1.37E-06
Testing set (GSE39582)						
8 gene risk score (high risk vs. low risk)	1.46	1.10-1.95	0.0098	1.36	1.02-1.82	0.037
Age (≥65 vs. <65)	1.47	1.08-2.00	0.0138	1.73	1.27-2.36	0.0005
Gender (male vs. female)	1.31	0.98-1.75	0.0684	1.41	1.05-1.88	0.0213
Stage (I/II/III/IV)	1.92	1.57-2.34	1.93e-10	2.01	1.63-2.46	3.35e-11
Testing set (GSE17536)						
8 gene risk score (high risk vs. low risk)	1.58	0.99-2.52	0.0565	1.51	0.93-2.45	0.092
Age (≥65 vs. <65)	0.96	0.60-1.52	0.848	1.25	0.77-2.02	0.360
Gender (male vs. female)	1.105	0.69-1.76	0.674	0.95	0.58-1.55	0.837
Stage (I/II/III/IV)	2.85	2.11-3.86	9.20E-12	2.97	2.17-4.08	1.24E-11

TCGA, The Cancer Genome Atlas; HR, Hazards ratio; CI, Confidence interval.

also indicated that the risk score may have as similar prediction effect as AJCC stage for CRC patients (**Figure 4F**).

We then carried out univariate and multivariate Cox regression analysis that included eight-gene risk score and other clinical factors upon

Table 4. Univariate and multivariate cox regression analyses of disease free survival in each data set

Variables	Univariate model		*p* value	Multivariate model		*p* value
	HR	95% CI		HR	95% CI	
Training set (TCGA)						
8 gene risk score (high risk vs. low risk)	2.03	1.29-3.21	0.0023	1.76	1.10-2.82	0.0191
Age (≥65 vs. <65)	0.84	0.54-1.31	0.444	0.84	0.53-1.34	0.4674
Gender (male vs. female)	1.55	0.98-2.45	0.062	1.58	0.98-2.55	0.0591
Stage (I/II/III/IV)	2.01	1.54-2.63	3.93E-07	1.88	1.43-2.46	5.72E-06
Testing set (GSE39582)						
8 gene risk score (high risk vs. low risk)	2.24	1.64-3.06	4.37e-07	1.96	1.43-2.69	2.5e-05
Age (≥65 vs. <65)	0.84	0.62-1.13	0.252	1.01	0.75-1.36	0.9546
Gender (male vs. female)	1.27	0.94-1.71	0.123	1.42	1.05-1.92	0.0247
Stage (I/II/III/IV)	2.69	2.17-3.32	<2e-16	2.71	2.17-3.39	<2e-16
Testing set (GSE17536)						
8 gene risk score (high risk vs. low risk)	5.93	2.47-14.27	7.05E-05	5.31	2.20-12.82	0.0002
Age (≥65 vs. <65)	0.66	0.34-1.27	0.214	0.88	0.45-1.71	0.7017
Gender (male vs. female)	1	0.52-1.92	0.9999	0.94	0.48-1.84	0.8617
Stage (I/II/III/IV)	2.05	1.35-3.10	0.0007	2.01	1.26-3.21	0.0036

TCGA, The Cancer Genome Atlas; HR, Hazards ratio; CI, Confidence interval.

training and testing sets (**Tables 3** and **4**). As gene risk score, age, stage and gender are available in all cohorts, they were defined as covariates in the multivariate cox regression model (chemotherapy only available in partial TCGA patients and all GSE39582 samples). Taken together, these results showed that eight-gene risk score may serve as a powerful predictor of patient survival when adjusted by age, gender or stage in three independent datasets.

The eight gene signature for survival prediction in patient with adjuvant chemotherapy

We also assessed the prognostic value of the signature for the patients with or without postoperative chemotherapy. Adjuvant chemotherapy information were available in the GSE39582 series. There are 233 patients who were get adjuvant chemotherapy, and 313 patients without, respectively. We found that high-risk score was significantly correlated with an unfavorable overall survival (**Figure 5A-C**) in patients with chemotherapy (HR=2.35, P=0.07) and those without (HR=4.12, P=0.006); recurrence-free survival (**Figure 5D-F**) of patients with chemotherapy (HR=2.71, P=0.015) and those without (HR=11.96, P<0.001). The result indicated that the eight gene signature model may have a more powerful predict effect in predicting recurrence free survival.

Discussion

In the present study, we mined cytokine-cytokine receptor interaction pathway gene expression profiles of 1110 colorectal cancer patients from TCGA and GEO database. By analyzing the association between gene expression profiling and prognosis of colorectal patients, we identified an eight genes, which are in cytokine and cytokine receptor interaction pathway, significantly related to the overall survival of colorectal patients. Further, we investigated the prognostic value of the eight-gene signature in different datasets and demonstrated the eight-gene signature is an independent predictor for colorectal cancer.

Among the eight genes identified in the present study, five of TGF-β superfamily members: TGFB2, TGFB3, INHBB, INHBC and GDF5 can bind to its TGF-β ligands and initiated TGF-β signaling. They were considered to play important roles in several biological processes, including cell proliferation, migration and apoptosis [19]. Previous study revealed that elevated TGFB (TGFB2 and TGFB3) level in CRC can influence disease progression and is associated with poor-prognosis [20]. INHBA upregulation indicated a poorer outcome in CRC [21], as well as INHBB and INHBC. Previous studies have reported that VEGFC plays an important role in lymphangiogenesis. Survival time of VEGF-C

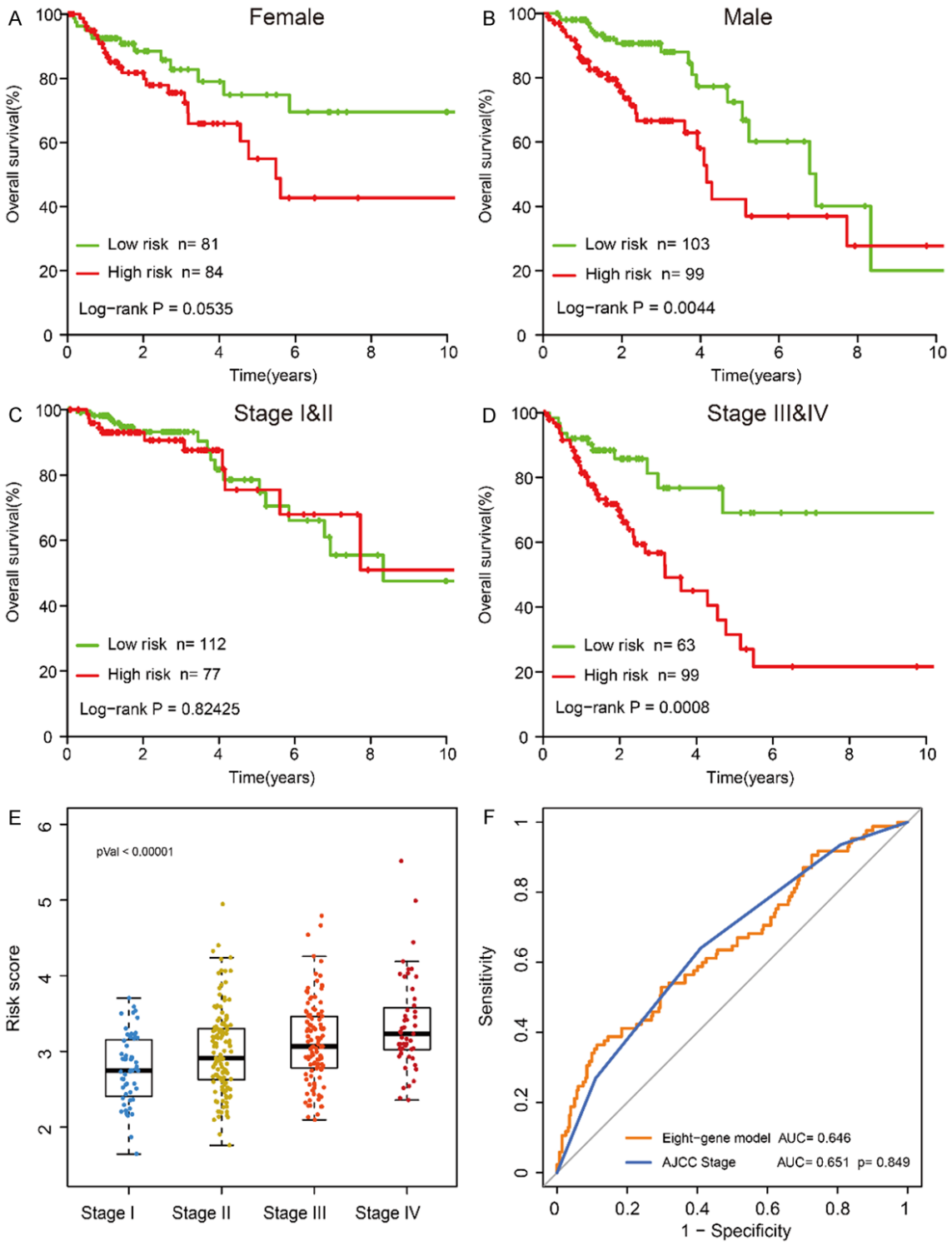

Figure 4. The association of eight-gene signatures and clinical parameters in training set. Kaplan-Meier curves of high and low risk group in patients of (A) female; (B) Male; (C) Stage I&II; (D) Stage III&IV; (E) Risk score base eight gene was associated with AJCC stage; (F) ROC analysis for eight gene signature and clinical stage.

positive patients was shorter than VEGF-C negative ones [22]. As Leptin may regulate colorectal carcinoma proliferation and apoptosis through PI3K/Akt/mTOR signaling pathway, it is

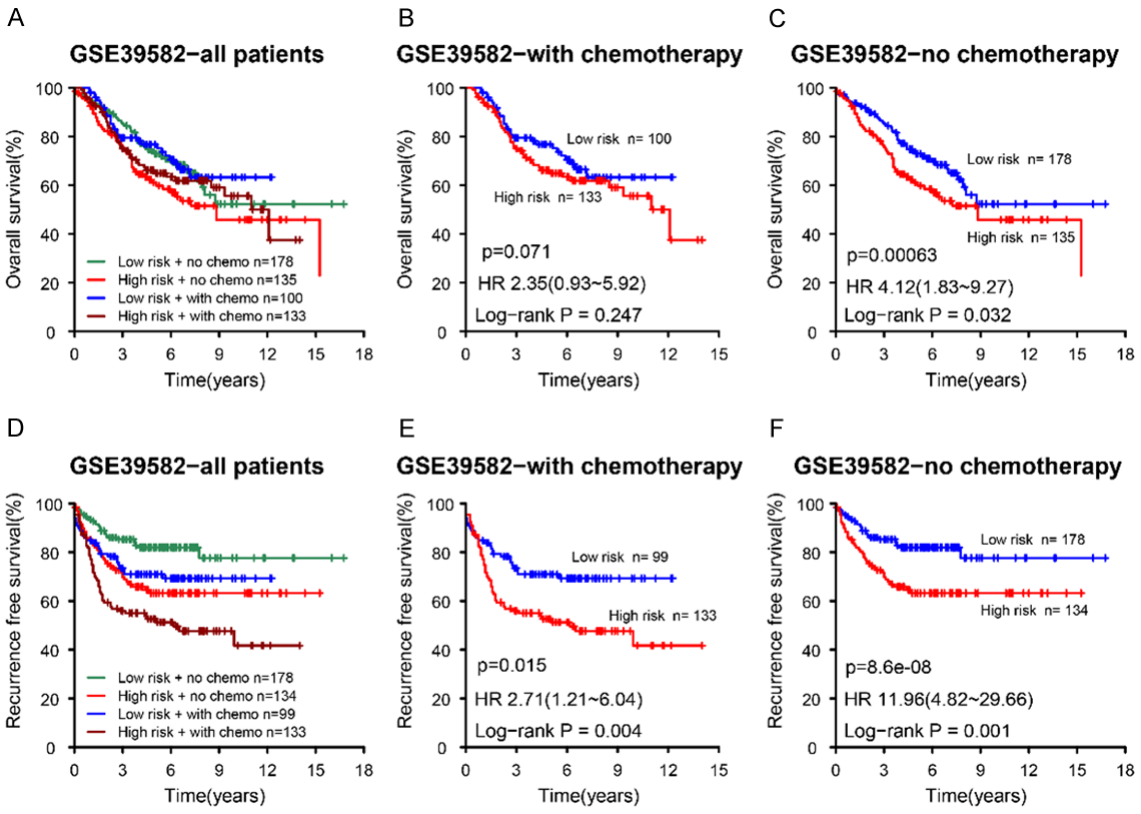

Figure 5. The eight gene signature for survival prediction in patient with adjuvant chemotherapy. Kaplan-Meier analysis of overall survival for (A) all patients (N=562); (B) Patients with chemotherapy (N=240); (C) Patients without chemotherapy (N=323). Kaplan-Meier analysis of recurrence free survival for (D) all patients (N=561); (E) Patients with chemotherapy (N=239); (F) Patients without chemotherapy (N=322).

considered to be a risk factor for colorectal cancer development [23]. GDF5 and IL20RB seems have no directly tumorgenesis function in colorectal cancer. Further investigations of these genes upon vivo and vitro level may enhance our understanding of their prognosis roles in colorectal cancer.

The eight-gene signature have predicative and prognostic effect in colorectal cancer patient under adjuvant chemotherapy

According to the result, the eight-gene signature was demonstrated as an independent and reliable prognostic predictor for colorectal cancer in both datasets. Apart from that, we noticed that the cytokine-cytokine receptor interaction gene signature can also predict an unfavorable prognosis of patients with postoperative chemotherapy in the GSE39582 cohorts. The association between high risk score patients and adjuvant chemotherapy resistance, indicating the possibility of chemotherapy response involvement and therapeutic outcome contribution of those signature genes. Thereby, we propose the crucial role of the eight-gene predictor in drug resistance and treatment outcomes, as they may reflect activation state of inflammation and tumor microenvironment.

Although this study retrieved large populations from different platform and have a long-time follow-up, the limitations should be acknowledged as well. First, mRNA expression profiles were taken from different center different platform: (TCGA) Illumina RNA-seq and (GEO) Affymetrix microarray. Second, the significance and robustness of the eight-gene signature as a prognostic predictor require large prospective patient cohorts confirmation. Further validation will contribute to a more powerful and reliable prognosis predictor for CRC.

In conclusion, we have developed a cytokine-cytokine receptor interaction gene signature model that can predict colorectal cancer prog-

nosis in three independent data sets. The cytokines gene expression pattern signature can also predict unfavorable prognosis of patients with postoperative chemotherapy, which means it can improve molecular classification and outcome prediction for colorectal cancer alone or combining with other molecular prognostic predictors. Specifically, it can help to identify patients who may benefit more from chemotherapies, or help to select the patients suitable for their use. However, further research is needed for better understanding of the role of cytokine-cytokine receptor interaction signaling networks in colorectal cancer development.

Acknowledgements

We thank The Cancer Genome Atlas (TCGA) project and Gene Expression Omnibus (GEO) database for providing their platforms, and contributors for their valuable data sets. This study was supported by grant from the National High Technology Research and Development Program of China (863 Program) (No. 2015AA020104).

Disclosure of conflict of interest

None.

Address correspondence to: Lei Liu and Gang Liu, Shanghai Public Health Clinical Center, Institute of Biomedical Sciences, Fudan University, 131 Dong'an Road, Shanghai 200032, China. E-mail: liulei_sibs@163.com (LL); liugang@fudan.edu.cn (GL)

References

[1] Siegel R, Naishadham D and Jemal A. Cancer statistics, 2013. CA Cancer J Clin 2010; 60: 10-29.

[2] West NR, Mccuaig S, Franchini F and Powrie F. Emerging cytokine networks in colorectal cancer. Nat Rev Immunol 2015; 15: 615-629.

[3] Dranoff G. Cytokines in cancer pathogenesis and cancer therapy. Nat Rev Cancer 2004; 4: 11.

[4] Lasry A, Zinger A and Ben-Neriah Y. Inflammatory networks underlying colorectal cancer. Nat Immunol 2016; 17: 230.

[5] Grivennikov SI, Greten FR and Karin M. Immunity, inflammation, and cancer. Cell 2010; 140: 883-899.

[6] Calon A, Espinet E, Palomo-Ponce S, Tauriello DF, Iglesias M, Céspedes M, Sevillano M, Nadal C, Jung P and Zhang XF. Dependency of colorectal cancer on a TGF-β-driven program in stromal cells for metastasis initiation. Cancer Cell 2012; 22: 571.

[7] Galon J, Costes A, Sanchezcabo F, Kirilovsky A, Mlecnik B, Lagorcepagès C, Tosolini M, Camus M, Berger A and Wind P. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science 2006; 313: 1960-1964.

[8] Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, Lugli A, Zlobec I, Hartmann A and Bifulco C. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. J Pathol 2014; 232: 199-209.

[9] Gray RG, Quirke P, Handley K, Lopatin M, Magill L, Baehner FL, Beaumont C, Clarklangone KM, Yoshizawa CN and Lee M. Validation study of a quantitative multigene reverse transcriptase-polymerase chain reaction assay for assessment of recurrence risk in patients with stage II colon cancer. J Clin Oncol 2011; 29: 4611.

[10] Jayadevi K, Chithra S, Clark-Langone KM and Drew W. Translating tumor biology into personalized treatment planning: analytical performance characteristics of the Oncotype DX® colon cancer assay. BMC Cancer 2010; 10: 691.

[11] Salazar R, Roepman P, Capella G, Moreno V, Simon I, Dreezen C, Lopezdoriga A, Santos C, Marijnen C and Westerga J. Gene expression signature to improve prognosis prediction of stage II and III colorectal cancer. J Clin Oncol 2011; 29: 17-24.

[12] Ågesen TH, Sveen A, Merok MA, Lind GE, Nesbakken A, Skotheim RI and Lothe RA. ColoGuideEx: a robust gene classifier specific for stage II colorectal cancer prognosis. Gut 2012; 61: 1560-1567.

[13] Cline MS, Craft B, Swatloski T, Goldman M, Ma S, Haussler D and Zhu J. Exploring TCGA Pan-cancer data at the UCSC cancer genomics browser. Sci Rep 2013; 3: 2652.

[14] Lu X, Wan F, Zhang H, Shi G and Ye D. ITGA2B and ITGA8 are predictive of prognosis in clear cell renal cell carcinoma patients. Tumor Biol 2016; 37: 253-262.

[15] Lossos IS, Czerwinski DK, Alizadeh AA, Wechser MA, Tibshirani R, Botstein D and Levy R. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med 2004; 350: 1828-1837.

[16] Kawaguchi A, Iwadate Y, Komohara Y, Sano M, Kajiwara K, Yajima N, Tsuchiya N, Homma J, Aoki H and Kobayashi T. Gene expression signature-based prognostic risk score in patients with primary central nervous system lymphoma. Clin Cancer Res 2012; 18: 5672-5681.

[17] Jin M, Li P, Zhang Q, Yang Z and Shen F. A four-long non-coding RNA signature in predicting

breast cancer survival. J Exp Clin Cancer Res 2014; 33: 84.
[18] Zhang XQ, Sun S, Lam KF, Kiang KM, Pu JK, Ho AS, Lui WM, Fung CF, Wong TS and Leung GK. A long non-coding RNA signature in glioblastoma multiforme predicts survival. Neurobiol Dis 2013; 58: 123.
[19] Syed V. TGF-β signaling in cancer. J Cell Biochem 2016; 117: 1279-1287.
[20] Calon A, Lonardo E, Berenguerllergo A, Espinet E, Hernandomomblona X, Iglesias M, Sevillano M, Palomoponce S, Tauriello DV and Byrom D. Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. Nat Genet 2015; 47: 320.
[21] Okano M, Yamamoto H, Ohkuma H, Kano Y, Kim H, Nishikawa S, Konno M, Kawamoto K, Haraguchi N and Takemasa I. Significance of INHBA expression in human colorectal cancer. Oncol Rep 2013; 30: 2903-2908.
[22] Akagi K, Ikeda Y, Miyazaki M, Abe T, Kinoshita J, Maehara Y and Sugimachi K. Vascular endothelial growth factor-C (VEGF-C) expression in human colorectal cancer tissues. Br J Cancer 2000; 83: 887.
[23] Wang D, Chen J, Chen H, Duan Z, Xu Q, Wei M, Wang L and Zhong M. Leptin regulates proliferation and apoptosis of colorectal carcinoma through PI3K/Akt/mTOR signalling pathway. J Biosci 2012; 37: 91.

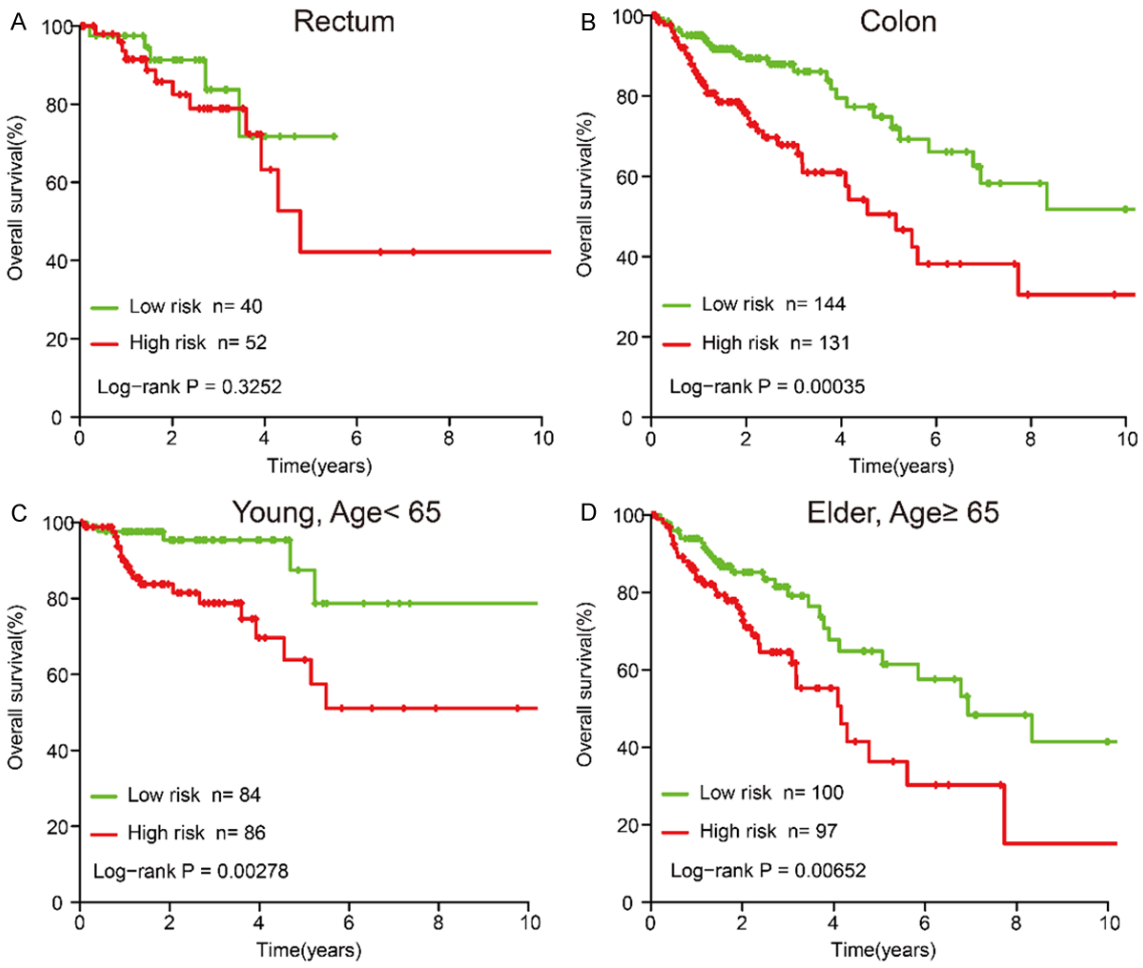

Figure S1. The association of eight-gene signatures and other clinicalparametersin TCGA. Kaplan-Meier curves of high and low risk group in patients of (A) rectum; (B) Colon; (C) Youngpatients (age <65); (D) Elderpatients (age ≥65).