IJC
International Journal of Cancer

A three-gene signature and clinical outcome in esophageal squamous cell carcinoma

Ling-Ling Sun[1,2], Jian-Yi Wu[3], Zhi-Yong Wu[4], Jin-Hui Shen[5], Xiu-E Xu[1], Bo Chen[1], Shao-Hong Wang[5], En-Min Li[3] and Li-Yan Xu[1]

[1] Institute of Oncologic Pathology, The Key Laboratory of Molecular Biology for High Cancer Incidence Coastal Chaoshan Area, Shantou University Medical College, Shantou, Guangdong, China
[2] Department of Pathology, Shaoyang Central Hospital, Affiliated Shaoyang Hospital of University of South China, Shaoyang, Hunan, China
[3] Department of Biochemistry and Molecular Biology, The Key Laboratory of Molecular Biology for High Cancer Incidence Coastal Chaoshan Area, Shantou University Medical College, Shantou, Guangdong, China
[4] Department of Oncology Surgery, Shantou Central Hospital, Affiliated Shantou Hospital of Sun Yat-sen University, Shantou, Guangdong, China
[5] Department of Pathology, Shantou Central Hospital, Affiliated Shantou Hospital of Sun Yat-sen University, Shantou, Guangdong, China

It is increasingly apparent that cancer development depends not only on genetic alterations, but also on epigenetic changes involving histone modifications. GASC1, member of the histone demethylases affecting heterochromatin formation and transcriptional repression, has been found to be dysregulation in many types of cancers including breast cancer, prostate cancer, metastatic lung sarcomatoid carcinoma, and leukemia. In this study, we examined the expression of GASC1 and certain GASC1-targeted genes (KLF4, MYC, SOX2, PPARG, MDM2, and NANOG) and identified a three-gene prognostic signature (PPARG, MDM2, and NANOG), using risk scores based on immunohistochemical analyses of 149 tumor specimens from patients with esophageal squamous cell carcinoma (ESCC). The presence of a high-risk three-gene signature in the ESCC tumors was significantly associated with decreased overall survival (OS) of the patients. We validated the predictive value of the three-gene signature in a second independent cohort of 101 patients with ESCC in order to determine whether it had predictive value. The results were similar to those in 149 patients. According to multivariate Cox proportional hazards analyses, the predictive model of a three-gene signature was an independent predictor for OS ($p = 0.005$ in cohort 1, $p = 0.025$ in cohort 2). In addition, ROC analysis indicated that the predictive ability of the three-gene model was more robust than that of a single biomarker. Therefore, our three-gene signature is closely associated with OS among patients with ESCC and may serve as a predictor for the poor prognosis of ESCC patients.

Esophageal carcinoma has a high mortality rate and is one of the most prevalent gastrointestinal cancers worldwide.[1] Among the various histological subtypes of esophageal carcinoma, esophageal squamous cell carcinoma (ESCC) occurs most often in Asia.[2] The 5-year overall survival (OS) rate of patients with ESCC is less than 10%.[3] The poor prognosis and low OS rate for patients with ESCC are due in part to the difficult nature of diagnosing early-stage ESCC and in part to the frequent occurrence of local invasion and lymph node metastasis in cases of advanced ESCC.[4] In addition, conventional chemotherapy and radiotherapy treatments are relatively ineffective.[5] Therefore, the identification of a

Key words: esophageal squamous cell carcinoma, histone demethylase, gene prognostic signature, histopathology

Abbreviations: ESCC: esophageal squamous cell carcinoma; GASC1: the gene amplified in squamous cell carcinoma; KLF4: kruppel-like factor 4; MDM2: mouse double minute 2 homolog; OS: overall survival; PPARG: peroxisome proliferator-activated receptor gamma; ROC: receiver operating characteristic; SOX2: sex determining region Y-box 2

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: National Basic Research Program; **Grant number:** 973 Program No. 2012CB526600; **Grant sponsors:** National High Technology Research and Development Program of China; **Grant numbers:** 2012AA02A503 and 2012AA02A209; **Grant sponsor:** Natural Science Foundation of China-GuangDong Joint Fund; **Grant numbers:** U0932001 and U1301227; **Grant sponsor:** National Science Foundation of China; **Grant number:** 81172264

DOI: 10.1002/ijc.29211

History: Received 30 Apr 2014; Accepted 9 Sep 2014; Online 13 Sep 2014

Correspondence to: Li-Yan Xu, Institute of Oncologic Pathology, Shantou University Medical College, No.22, Xinling Road, Shantou, Guangdong 515041, China, Tel.: +86-754-88900460, Fax: +86-754-88900847, E-mail: lyxu@stu.edu.cn or nmli@stu.edu.cn

What's new?

Epigenetic alterations that involve modifications to histones are thought to play critical roles in cancer, with effects on processes ranging from tumor development to metastasis. The present investigation focused on the expression of the histone demethylase GASC1 and its gene targets in tumors from patients with esophageal squamous cell carcinoma (ESCC). Using risk scores from immunohistochemical analyses, the authors developed a three-gene prognostic signature involving the genes *PPARG*, *MDM2*, and *NANOG*. The signature was associated with a reduction in overall survival of ESCC patients, suggesting that it is predictive for poor prognosis in ESCC.

sensitive and reliable method that would help identify patients at a higher or lower risk of death for ESCC progression is of critical importance, not only for appropriate treatment and improved prognosis, but also for a better understanding of the molecular and cellular processes involved in the tumorigenesis of ESCC.

It is increasingly apparent that cancer development depends not only on genetic alterations, but also on epigenetic changes involving histone modifications.[6] For example, methylation of histones is regarded as a stable modification that defines the epigenetic program of a cell, which regulates chromatin structure and transcription. However, the recent discovery of histone demethylases has challenged this view of the stable nature of histone methylation.[7] Recently, several histone demethylases were found to be involved in many types of tumors. Specifically, the histone demethylase GASC1 (the gene amplified in squamous cell carcinoma 1), a member of the JmjC-domain-containing proteins, has been shown to be overexpressed, amplified, and/or mutated in such human cancers as breast cancer, prostate cancer, metastatic lung sarcomatoid carcinoma, and leukemia.[6,8–11] More importantly, recent studies have demonstrated that GASC1 regulates (both positively and negatively) the expression of Kruppel-like factor 4 (KLF4), sex determining region Y-box 2 (SOX2), peroxisome proliferator-activated receptor gamma (PPARG), a regulator gene that encodes for a transcription factor (MYC), mouse double minute 2 homolog (MDM2), and a transcription factor critically involved with self-renewal of undifferentiated embryonic stem cells (NANOG) via histone lysine demethylase activity.[6,12–14] Previously, we found that histone demethylase GASC1 was overexpressed in a subset of primary ESCC samples and was significantly associated with lymph node metastasis and tumor-node-metastasis (TNM) classification of the International Union against Cancer.[15] Indeed, GASC1 expression was originally found to be up-regulated in several ESCC cell lines, suggesting that overexpressed GASC1 may play an important role in the development and/or progression of various types of cancer, including ESCC).[16–19] Therefore, we hypothesized that GASC1 and a set of key genes regulated by GASC1 may serve as a predictor for the poor prognosis of ESCC patients.

In the current study, we examined the expression of GASC1 in combination with several genes (KLF4, MYC, SOX2, PPARG, MDM2, and NANOG) in 149 surgical specimens with ESCC using immunohistochemistry. We then built a predictive model based on the genes that were correlated with OS, either positively or negatively, and validated the model by applying it to a set of such samples from an independent cohort of 101 patients with ESCC in order to determine whether it had predictive value. Our goal was to identify a gene signature that is correlated with the clinical outcome of patients with ESCC.

Material and Methods

Patients and specimens

For the retrospective study, 149 archival formalin-fixed, paraffin-embedded ESCC specimens between 1987 and 1997 were retrieved from the Department of Clinical Pathology at Center Hospital of Shantou City. We validated the three-gene risk-prediction model using an independent cohort of 101 randomly selected patients who underwent surgical resection of ESCC at the Department of Clinical Pathology of Center Hospital of Shantou City between 2007 and 2011. The clinicopathological characteristics of patients in the two cohorts are summarised in Table 1. The follow-up for patients after esophageal resection was continued until their deaths and only patients that died from ESCC were included in the tumor-related deaths. The patients, suffering from severe postoperative complications, other tumors, or died of other causes were excluded.

All the tumors were confirmed as ESCC by the pathologists in the Clinical Pathology Department of the Hospital, and the cases were classified according to the seventh edition of the tumor-node-metastasis (TNM) classification of the International Union against Cancer. Evaluation of tumor differentiation was based on histological criteria of the guidelines of the WHO Pathological Classification of Tumors. The study was approved by the Ethics committee of the Center Hospital of Shantou City, the local ethics committee, and only patients with written informed consent were included.

Tissue microarray construction and immunohistochemical analysis

Tissue microarray (TMA) construction of esophageal carcinoma tissue has been described earlier.[20] KLF4, MYC, SOX2, PPARG, MDM2, NANOG and GASC1 were analyzed in this study. The antibodies are shown in Supporting Information Table 1. After dewaxing in xylene and rehydration in a series of graded alcohols, TMA sections were sectioned (4 μm) and subjected to immunostaining in the SuperPicTureTM Polymer Detection Kit and the Liquid DAB Substrate Kit (Zymed/Invitrogen, San Francisco, CA).

Immunohistochemical staining was assessed by three independent pathologists (B.C., J.-H.S. and S.-H.W.) without

Table 1. Patient characteristics

Clinical Parameter	Cohort 1 (No.)	Cohort 2 (No.)
Specimens	149	101
Mean age	54.7	58.6
Age		
≤54	67	37
>54	82	64
Gender		
Male	112	80
Female	37	21
Tumor size (cm)		
≤3	39	29
3–5	79	47
>5	31	25
Histologic grade		
G1	29	13
G2	99	77
G3	21	11
Invasive depth		
T1	1	2
T2	21	11
T3	123	86
T4	4	2
Lymph node metastasis		
N0	87	44
N1 + N2+ N3	62	57
TNM classification		
I		
IA	1	1
IB	3	3
II		
IIA	31	12
IIB	51	31
III		
IIIA	50	29
IIIB	3	16
IIIC	1	0
IV		
IV	9	0

knowledge of patient characteristics. Discrepancies were resolved by consensus. The immunohistochemical staining results were assigned a maximum score considering both the intensity of staining and the proportion of tumor cells showing unequivocal positive reaction. Positive reactions were defined as those showing brown immunostaining in the cell cytoplasm and nucleus. For KLF4, MYC, SOX2, PPARG, MDM2, NANOG and GASC1, a staining index (values 0–12) was determined by multiplying the score for staining intensity with the score for positive area. The intensity of staining was determined as: 0 = no staining; 1 = weak staining; 2 = moderate staining; and 3 = strong staining. Tumor cells area: 0 = positive staining in less than 5% of tumor cells; 1 = positive staining in 5 to 25% of tumor cells; 2 = positive staining in 26 to 50% of tumor cells; 3 = positive staining in 51 to 75% of tumor cells; 4 = positive staining in 75 to 100% of tumor cells. For statistical analyses, a composite staining index was defined as the product of intensity and area scores, giving values from 0 to 12. Negative/positivity was defined as low/high-expression on the basis of scores of 0–12 by the X-tile.[21] For KLF4, MYC and GASC1, scores of 0 to 4 were considered "negative staining" (low-expression), and scores of 5 to 12 were considered "positive staining" (high-expression) while scores of 0 to 8 were considered "negative staining" (low-expression), and scores of 9 to 12 were considered "positive staining" (high-expression) for MDM2 and NANOG. For PPARG, we assumed as negative scores from 0 to 3 (low expression), and positive scores of 4 to 12 (high expression) while we defined as negative scores from 0 to 9 (low expression), and positive scores of 10 to 12 (high expression) for SOX2.

Construction of a weighted overall survival (OS) predictive score algorithm

We used univariate Cox proportional hazards regression analysis to evaluate the association between patients' OS and the expression of each biomarker. A patient's risk score was derived by the summation of the expression level (positive = 2, negative = 1) of each biomarker multiplied by its corresponding regression coefficient.[22] All patients were then divided into two groups (high-risk signature and low-risk signature) by the cut-off value that came from the median of the final risk scores.

Statistical analysis

Statistical analyses were performed using SPSS 13.0 for windows (IBM, Chicago, IL). Overall survival time was calculated by the Kaplan-Meier method and analysed by the log-rank test. The overall survival was defined as the time from the date of primary surgery to the date of death due to esophageal cancer and data on survivors were recorded at the last follow-up. Univariate and multivariate analyses were based on the Cox proportional hazards regression model. The correlation significance was analyzed by Kendall tau-b rank correlation analysis. Receiver operating characteristic (ROC) curve analysis was used to determine the predictive value of the parameters. A p value of less than 0.05 was considered statistically significant and each value is two-tailed.

Results

Expression of seven biomarkers in ESCC

Cytoplasmic and/or nuclear immunostaining patterns of seven biomarkers were successfully interpreted in ESCC

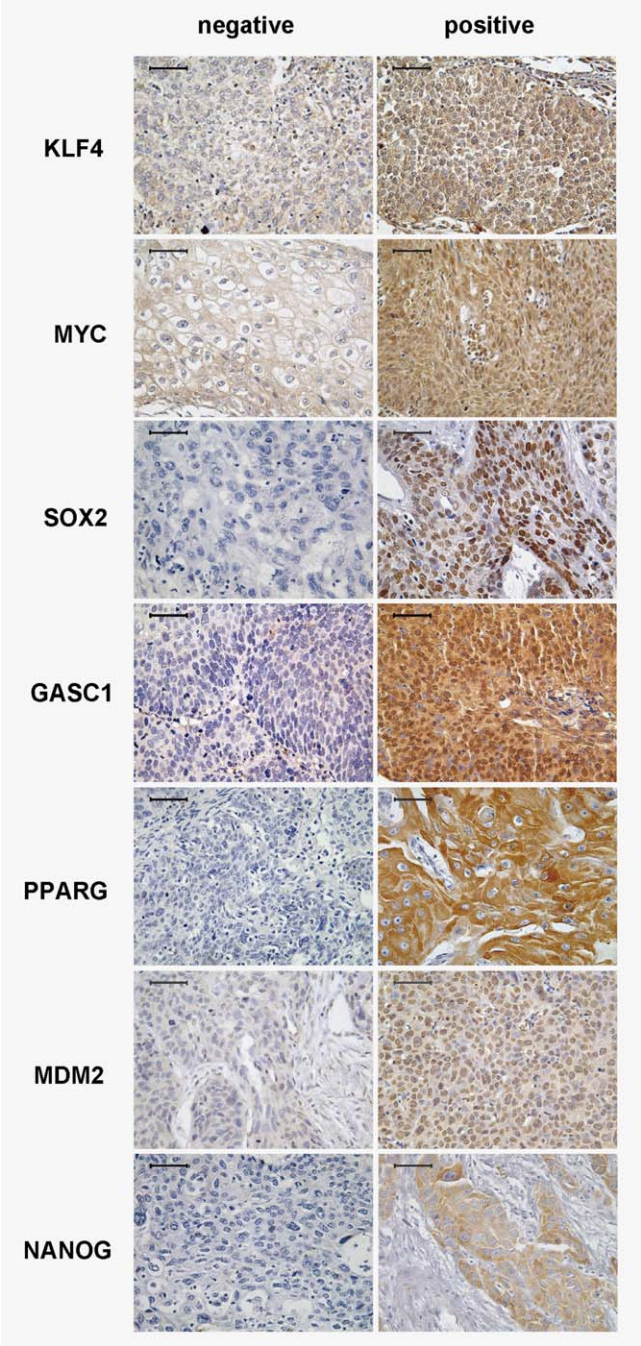

Figure 1. Representative positive/negative expression of KLF4, MYC, SOX2, GASC1, PPARG, MDM2, and NANOG by immunochemistry study in tissue microarrays. The bar indicates 50 μm.

tissues. Based on the staining intensity, all the biomarkers displayed two immunostaining phenotypes; that is, negative staining and positive diffuse staining (Fig. 1). The staining patterns of the biomarkers varied in staining intensity and percentage of positive cells. A duplicate set of spots for each tumor showed a good level of homogeneity for both intensity and stained cell percentages. The patterns were focal, scattered, or diffuse at different staining intensities. The staining patterns of seven biomarkers were also varied by location. PPARG and NANOG protein staining was primarily observed in the cytoplasm, SOX2 was primarily observed in the nucleus, and KLF4, MYC, MDM2, and GASC1 showed both positive cytoplasmic and strong nuclear immunostaining (Fig. 1).

Table 2. Independent index of prognosis assessment by clinical characteristics

			95.0% CI for Exp(B)	
Parameter	Sig.	Exp(B)	Lower	Upper
Cohort 1				
PPARG	0.001	2.458	1.448	4.172
MDM2	0.030	1.698	1.054	2.736
NANOG	0.015	1.924	1.737	3.255
Cohort 1				
Lymph node metastasis	0.000	2.651	1.546	4.544
Three-gene signature	0.005	1.987	1.232	3.204
Cohort 2				
Lymph node metastasis	0.014	2.315	1.183	4.531
Three-gene signature	0.025	2.081	1.095	3.956

Statistical analysis: the multivariate Cox proportional-hazards regression.

Prognostic significance of seven biomarkers and clinicopathological characteristics

The 5-year OS was 40.6% for the entire study population of cohort 1. The results of univariate analysis confirmed that three biomarkers (PPARG, MDM2, and NANOG) and two clinical factors (lymph node metastasis and TNM classification) were prognostic factors for OS, whereas KLF4, MYC, SOX2, GASC1, and other clinical indexes (age, gender, tumor size, differentiation grade, and invasive depth) had no prognostic significance for OS (Supporting Information Table 2 and Supporting Information Fig. 1). PPARG, MDM2, and NANOG were also independent factors for OS according to multivariable Cox proportional hazard regression analyses (Table 2).

A predictive model of the three-gene signature and survival

The risk score of the predictive model was calculated as follows: $(0.566 \times \text{PPARG}) + (0.708 \times \text{MDM2}) + (0.627 \times \text{NANOG})$. The coefficients were calculated by Cox regression, and the gene name represents its expression level (positive = 2, negative = 1). The median of the final risk scores was 2.467. All patients were divided into the high-risk signature (risk score >2.467) and low-risk signature (risk score $\leq$2.467, Fig. 2*a*).

The OS in the high-risk signature group was significantly shorter than that in the low-risk signature group ($p < 0.001$, Fig. 2*a*). In a subgroup analysis of 127 patients with invasive

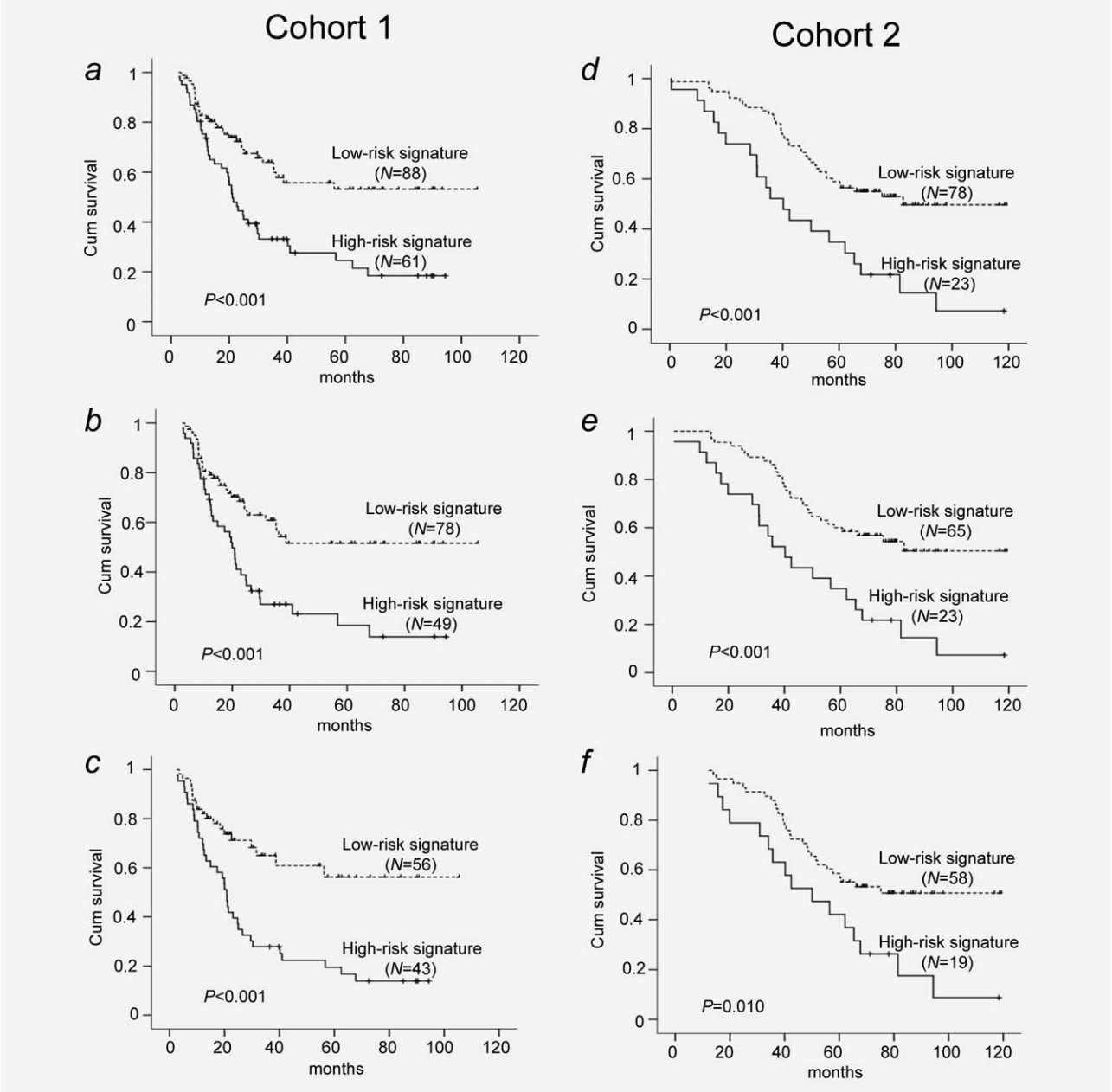

Figure 2. Kaplan-Meier estimates of survival of patients with ESCC according to the three-gene signatures as measured by immunohistochemistry. In cohort 1, overall survival is shown for the 149 patients with ESCC (*a*), for the 127 patients with invasive depth 3 (T3) or invasive depth 4 (T4) disease (*b*), and for the 99 patients with differentiation grade 2 (G2) disease (*c*). Overall survival is also shown for the independent cohort 2 of 101 patients (*d*), for the 88 patients in cohort 2 who had T3 or T4 disease (*e*), and for the 77 patients in cohort 2 who had G2 disease (*f*).

depth 3 (T3) or invasive depth 4 (T4) disease, those with a high-risk gene signature had a shorter OS than those with a low-risk gene signature ($p < 0.001$, Fig. 2*b*). In a subgroup analysis of 99 patients with differentiation grade 2 (G2) disease, patients with a high-risk gene signature had a lower OS than patients with a low-risk gene signature ($p < 0.001$, Fig. 2*c*). Using multivariate analyses, the predictive model of a three-gene signature was an independent predictor for OS ($p = 0.005$, Table 2).

Correlation of the predictive model with clinicopathological features

To obtain a better understanding of the clinical significance of the predictive model in patients with ESCC, we correlated it with a series of clinicopathological parameters. As shown in Table 3, a significant correlation was observed between the three-gene signature and lymph node metastasis ($p = 0.029$) and TNM classification ($p = 0.016$). The high-risk gene signature was found in 51.6% (32/62) of lymph node metastasis

Table 3. Association between the three-gene signature and clinical pathological parameters in ESCC in cohort 1

Clinical parameter	Three-gene signature status		
	Low-risk signature	High-risk signature	r/p
Age (yr)			
≤54	43	24	0.094/0.315
>54	45	37	
Gender			
Male	64	48	−0.068/0.446
Female	24	13	
Tumor size (cm)			
≤3	23	16	0.005/0.969
3–5	47	32	
>5	18	13	
Differentiation			
G1	18	11	−0.015/0.874
G2	56	43	
G3	14	7	
Invasive depth			
T1 + T2	10	12	−0.115/0.240
T3 + T4	78	49	
Lymph node metastasis			
N0	58	29	0.183/0.029
N1 + N2 + N3	30	32	
TNM classification			
I (IA + IB)	3	1	0.193/0.015
II (IIA + IIB)	55	27	
III (IIIA + IIIB + IIIC)	26	28	
IV	4	5	

Statistical analysis: the Kendall's tall-b test.

compared with 33.3% (29/87) of no-lymph node metastasis. In addition, the high-risk gene signature was found in 32.6% (28/86) of TNM-I or TNM-II disease, 51.9% (28/54) of TNM-III disease, and 55.6% (5/9) of TNM-IV disease. There was no significant difference in other clinicopathological features between high-risk gene signature and low-risk gene signature groups.

Validation of the predictive model

We validated the predictive value of the three-gene signature in another independent cohort (cohort 2) of 101 patients with ESCC. The 3-year OS was 47.2% for the entire study population of cohort 2. The results for cohort 2 were similar to those in cohort 1. On univariate analysis, three biomarkers (PPARG, MDM2, and NANOG) and two clinical factors (lymph node metastasis and TNM classification) were also confirmed as prognostic factors for OS (Supporting Information Table 3 and Supporting Information Fig. 2). Patients with a high-risk gene signature had a shorter OS than those with a low-risk gene signature ($p < 0.001$, Fig. 2*d*). According to multivariate Cox proportional hazards analyses, the predictive model was still an independent predictor of OS (Table 2). We also analyzed the three-gene signature in tumor specimens obtained from patients in the validation cohort with T3 or T4 disease and G2 disease. Among 88 patients with T3 or T4 disease, those with a high-risk gene signature had a shorter OS than those with a low-risk gene signature ($p < 0.001$, Fig. 2*e*). Among 77 patients with G2 disease, patients with a high-risk gene signature had a lower OS than patients with a low-risk gene signature ($p = 0.010$, Fig. 2*f*). Moreover, the three-gene signature significantly correlated with lymph node metastasis ($p < 0.001$) and TNM classification in cohort 2 ($p < 0.001$) (Supporting Information Table 4). Compared with the single biomarkers, the predictive power of the three-gene signature was higher than that of PPARG, MDM2, or NANOG ($p < 0.001$) as revealed by the ROC analysis in both cohort 1 and cohort 2 (Fig. 3).

Discussion

Recent studies using loss- or gain-of-functions approaches indicate that histone demethylases, including GASC1, modulate histone methylation status and affect the expression of a set of key genes that are critical for cancer invasion and metastasis.[8,14,23] These studies suggest that the genes modulated by histone demethylases may serve as predictors for the poor prognosis of cancer patients. In this study, we identified a three-gene prognostic signature (PPARG, MDM2, and NANOG) from seven genes (including GASC1), using risk scores based on immunohistochemical analyses of 149 tumor specimens from patients with ESCC. The presence of a high-risk three-gene signature in the ESCC tumors was significantly associated with decreased OS. We validated the predictive value of the three-gene signature in a second independent cohort of 101 patients with ESCC. The results were similar to those in cohort 1. In addition, ROC analysis indicated that the predictive ability of the three-gene model was more robust than that of a single biomarker.

Diagnosis of ESCC at its early stage remains difficult. As a result, a great majority of patients with ESCC are in the advanced stages of the disease, and conventional chemotherapy and radiotherapy treatments are relatively ineffective. In a subgroup analysis of patients with T3 or T4 disease in this study, those with a high-risk gene signature had a shorter OS than those with a low-risk gene signature. Construction of this three-gene signature may improve the classification of patients in the late phase of ESCC and help doctors use different interventions on the basis of the classification, thereby paving a way for the discovery of novel treatment modalities. Similarly, our analysis of a subgroup of patients with G2 disease showed that patients with a high-risk gene signature had a lower OS than patients with a low-risk gene signature.

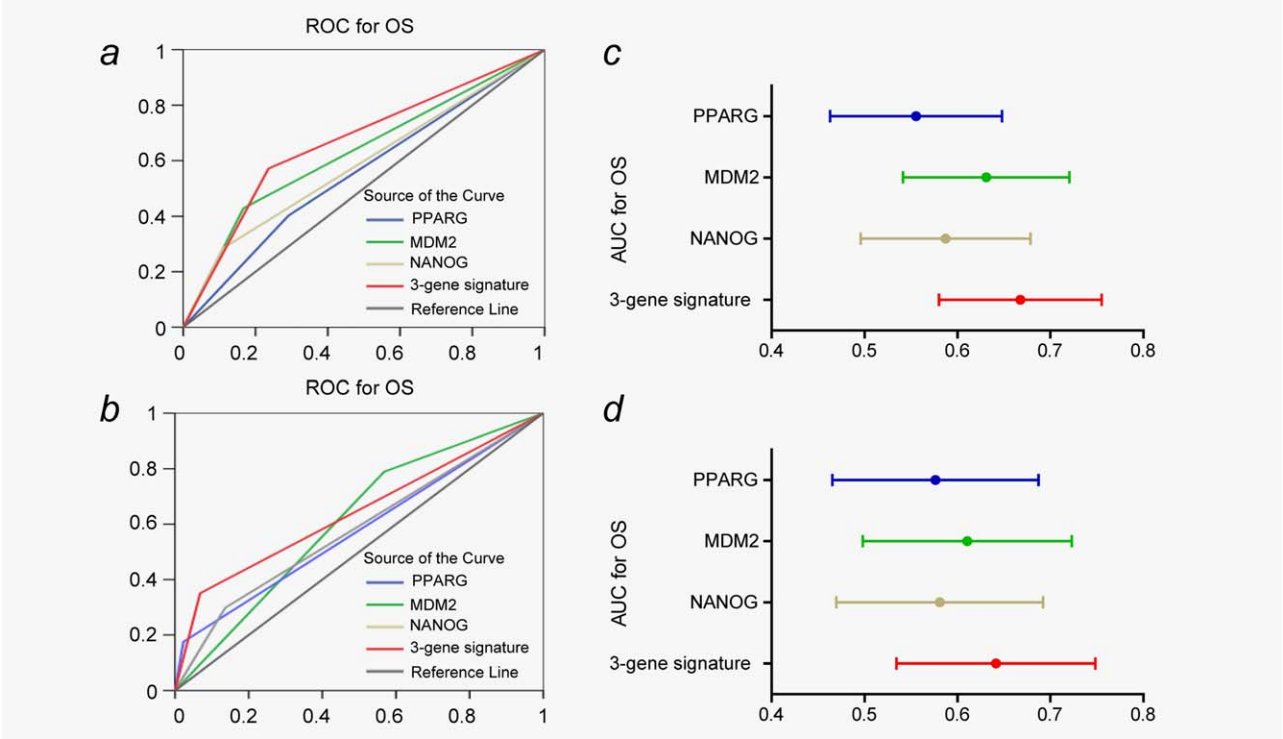

Figure 3. The predictive ability of the three-gene signature compared with single markers by receiver operating characteristic (ROC) curves (*a* in cohort 1 and *b* in cohort 2) and areas under the curve (AUC) with 95% CI (*c* in cohort 1 and *d* in cohort 2). The results show that the predictive ability of the three-gene model was more robust than that of a single biomarker.

These results indicate that the three-gene signature may also be useful in the future for planning treatment strategies for the clinical management of patients with G2 disease. In addition, as we can see in Supporting Information Tables 2 and 3, the number of ESCC patients with T1/T2 or G1/G3 is small in cohort 1 and cohort 2, which is the reason that they are not included in analysis. In China, Diagnosis of ESCC at its early stage or G1 still remains difficult and patients in the late phase of ESCC or G3 frequently displays local invasion and lymph node metastasis resulting in conservative treatment, which is one of the important reasons for the low number of ESCC patients with T1/T2 or G1/G3. It could be well worth studying the patients of T1/T2 and G1/G3, and we will gather such more specimens to study in the future and in our subsequent study.

The identification of three genes that can predict the clinical outcome in patients with ESCC may reveal targets for the development of therapy for esophageal cancer. PPARG, a member of the PPAR family (a subfamily of the nuclear receptor superfamily), has been reported to be significantly correlated with both tumor progression and patient prognosis in several types of carcinomas, including breast cancer, colon cancer, tongue squamous cell carninoma, prostate cancer, and pancreatic cancer.[24–28] NANOG, a cell-fate regulatory molecule known to be important for the self-renewal of embryonic stem cells, was found not only in germ cell tumors, but also in breast, cervical, oral cavity, kidney, and ovarian tumors.[29–34] This embryonic stem cell self-renewal molecule NANOG may conceptually contribute to tumorigenesis by a mechanism related to its role in embryonic stem cells. Alternatively, NANOG may enhance proliferation of cancer cells. This concept was supported by reports of exogenous overexpression of NANOG in mesenchymal stem cells and NIH3T3 cells that promoted cell proliferation and enhanced colony formation.[35,36] In addition, NANOG upregulated the expression of ezrin, which is involved in tumor progression and regulates cellular activities including survival, adhesion, and migration/invasion by organizing membrane-cytoskeleton-associated complexes.[37,38] MDM2 is overexpressed in most types of cancer from various tissues. Despite the sometimes conflicting results from studies, the overall trend is that MDM2 expression is associated with decreased OS, increased recurrence, increased metastasis, and decreased response to therapeutic intervention in a wide variety of tumors.[23] In this study, we first identified three prognostic factors (PPARG, MDM2, and NANOG) from seven biomarkers (including GASC1) through immunohistochemistry. Then, we built a predictive model based on the genes (PPARG, MDM2, and NANOG) using risk scores based on immunohistochemical analyses, which is correlated with the clinical outcome of patients with ESCC. We also speculate that other genes (SOX2, MYC and KLF4) without being prognostic may work in other ways. Recently, Neumann *et al.* demonstrated that SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon

cancer.[39] Additionally, in 2009, Rapp UR et al. discovered that MYC is a metastasis gene for non-small-cell lung cancer.[40] What's more, KLF4 is a transcriptional regulator of genes critical for EMT, including jnk1 (mapk8).[41] Considering all above, we speculate that these genes (SOX2, MYC and KLF4) have no direct prognostic significance but correlate with tumor invasion and migration by affecting the downstream genes of them or the genes in passageways. In addition, GASC1 interacts with hypoxia-inducible factor 1 (HIF1) and stimulates transcription of HIF-1 target genes, thereby promoting breast cancer growth and lung metastasis.[42] Studies also demonstrated that overexpression of GASC1 enhances sphere formation, a characteristic property of stem/progenitor cells, in breast and colonic cancer cells by mediating expression of Wnt and Notch pathway genes.[43] Therefore, we also speculate that GASC1 may be not the executor of the function in direct but the one which is involved in the development and progression of many kinds of tumors by regulating the expression of GASC1 target genes.

In conclusion, the three-gene prognostic signature we identified is closely associated with the clinical outcome in patients with surgically resected ESCC. This signature could be useful in stratifying patients according to risk in trials of adjuvant treatment of the disease.

References

1. Pisani P, Parkin DM, Bray F, et al. Estimates of the worldwide mortality from 25 cancers in 1990. *Int J Cancer* 1999;83:18–29.
2. Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol* 2006;24:2137–50.
3. Katlic MR, Wilkins EW, Jr, Grillo HC. Three decades of treatment of esophageal squamous carcinoma at the Massachusetts General Hospital. *J Thorac Cardiovasc Surg* 1990;99:929–38.
4. Enzinger PC, Mayer RJ. Esophageal cancer. *N Engl J Med* 2003;349:2241–52.
5. Tachibana M, Kinugasa S, Hirahara N, et al. Lymph node classification of esophageal squamous cell carcinoma and adenocarcinoma. *Eur J Cardiothorac Surg* 2008;34:427–31.
6. Liu G, Bollig-Fischer A, Kreike B, et al. Genomic amplification and oncogenic properties of the GASC1 histone demethylase gene in breast cancer. *Oncogene* 2009;28:4491–500.
7. Zhou W, Chen H, Zhang L. The PcG protein hPc2 interacts with the N-terminus of histone demethylase JARID1B and acts as a transcriptional co-repressor. *BMB Rep* 2009;42:154–59.
8. Cloos PA, Christensen J, Agger K, et al. The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature* 2006;442:307–11.
9. Suikki HE, Kujala PM, Tammela TL, et al. Genetic alterations and changes in expression of histone demethylases in prostate cancer. *Prostate* 2010;70:889–98.
10. Italiano A, Attias R, Aurias A, et al. Molecular cytogenetic characterization of a metastatic lung sarcomatoid carcinoma: 9p23 neocentromere and 9p23-p24 amplification including JAK2 and JMJD2C. *Cancer Genet Cytogenet* 2006;167:122–30.
11. Hélias C, Struski S, Gervais C, et al. Polycythemia vera transforming to acute myeloid leukemia and complex abnormalities including 9p homogeneously staining region with amplification of MLLT3, JMJD2C, JAK2, and SMARCA2. *Cancer Genet Cytogenet* 2008;180:51–5.
12. Wang J, Zhang M, Zhang Y, et al. The histone demethylase JMJD2C is stage-specifically expressed in preimplantation mouse embryos and is required for embryonic development. *Biol Reprod* 2010;82:105–11.
13. Lizcano F, Romero C, Vargas D. Regulation of adipogenesis by nuclear receptor PPARγ is modulated by the histone demethylase JMJD2C. *Genet Mol Biol* 2011;34:19–24.
14. Ishimura A, Terashima M, Kimura H, et al. Jmjd2c histone demethylase enhances the expression of Mdm2 oncogene. *Biochem Biophys Res Commun* 2009;389:366–71.
15. Sun LL, Holowatyj A, Xu XE, et al. Histone demethylase GASC1, a potential prognostic and predictive marker in esophageal squamous cell carcinoma. *Am J Cancer Res* 2013;3:509–17.
16. Yang ZQ, Imoto I, Fukuda Y, et al. Identification of a novel gene, GASC1, within an amplicon at 9p23–24 frequently detected in esophageal cancer cell lines. *Cancer Res* 2000;60:4735–9.
17. Yang ZQ, Imoto I, Pimkhaokham A, et al. A novel amplicon at 9p23 - 24 in squamous cell carcinoma of the esophagus that lies proximal to GASC1 and harbors NFIB. *Jpn J Cancer Res* 2001;92:423–8.
18. Hu N, Wang C, Hu Y, et al. Genome-wide association study in esophageal cancer using GeneChip mapping 10K array. *Cancer Res* 2005;65:2542–6.
19. Ragvin A, Valvatne H, Erdal S, et al. Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300. *J Mol Biol* 2004;337:773–88.
20. Zhang FR, Tao LH, Shen ZY, et al. Fascin expression in human embryonic, fetal, and normal adult tissue. *J Histochem Cytochem* 2008;56:193–9.
21. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* 2004;10:7252–9.
22. 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.
23. Rayburn E, Zhang R, He J, et al. MDM2 and human malignancies: expression, clinical pathology, prognostic markers, and implications for chemotherapy. *Curr Cancer Drug Targets* 2005;5:27–41.
24. Jiang Y, Zou L, Zhang C, et al. PPARgamma and Wnt/beta-Catenin pathway in human breast cancer: expression pattern, molecular interaction and clinical/prognostic correlations. *J Cancer Res Clin Oncol* 2009;135:1551–9.
25. Ogino S, Shima K, Baba Y, et al. Colorectal cancer expression of peroxisome proliferator-activated receptor gamma (PPARG, PPARgamma) is associated with good prognosis. *Gastroenterology* 2009;136:1242–50.
26. Theocharis S, Klijanienko J, Giaginis C, et al. Peroxisome proliferator-activated receptor-γ in mobile tongue squamous cell carcinoma: associations with clinicopathological parameters and patients survival. *J Cancer Res Clin Oncol* 2011;137:251–9.
27. Nakamura Y, Suzuki T, Sugawara A, et al. Peroxisome proliferator-activated receptor gamma in human prostate carcinoma. *Pathol Int* 2009;59:288–93.
28. Giaginis C, Katsamangou E, Tsourouflis G, et al. Peroxisome proliferator-activated receptor-gamma and retinoid X receptor-alpha expression in pancreatic ductal adenocarcinoma: association with clinicopathological parameters, tumor proliferative capacity, and patients' survival. *Med Sci Monit* 2009;15:BR148–56.
29. Ezeh UI, Turek PJ, Reijo RA, et al. Human embryonic stem cell genes OCT4, NANOG, STELLAR, and GDF3 are expressed in both seminoma and breast carcinoma. *Cancer* 2005;104:2255–65.
30. Alldridge L, Metodieva G, Greenwood C, et al. Proteome profiling of breast tumors by gel electrophoresis and nanoscale electrospray ionization mass spectrometry. *J Proteome Res* 2008;7:1458–69.
31. Ye F, Zhou C, Cheng Q, et al. Stem-cell-abundant proteins Nanog, Nucleostemin and Musashi1 are highly expressed in malignant cervical epithelial cells. *BMC Cancer* 2008;8:108.
32. Chiou SH, Yu CC, Huang CY, et al. Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. *Clin Cancer Res* 2008;14:4085–95.
33. Bussolati B, Bruno S, Grange C, et al. Identification of a tumor-initiating stem cell population in human renal carcinomas. *FASEB J* 2008;22:3696–705.
34. Zhang S, Balch C, Chan MW, et al. Identification and characterization of ovarian cancer-initiating cells from primary human tumors. *Cancer Res* 2008;68:4311–20.
35. Liu TM, Wu YN, Guo XM, et al. Effects of ectopic Nanog and Oct4 overexpression on mesenchymal stem cells. *Stem Cells Dev* 2009;18:1013–22.
36. Zhang J, Wang X, Chen B, et al. Expression of Nanog gene promotes NIH3T3 cell proliferation. *Biochem Biophys Res Commun* 2005;338:1098–102.

37. Lin YL, Han ZB, Xiong FY, et al. Malignant transformation of 293 cells induced by ectopic expression of human Nanog. *Mol Cell Biochem* 2011;351:109–16.
38. McClatchey AI. Merlin and ERM proteins: unappreciated roles in cancer development? *Nat Rev Cancer* 2003;3:877–83.
39. Neumann J, Bahr F, Horst D, et al. SOX2 expression correlates with lymph-node metastases and distant spread in right-sided colon cancer. *BMC Cancer* 2011;11:518.
40. Rapp UR, Korn C, Ceteci F, et al. Myc is a metastasis gene for non-small-cell lung cancer. *Plos One* 2009;4:e6029.
41. Tiwari N, Meyer-Schaller N, Arnold P, et al. Correction: Klf4 Is a Transcriptional Regulator of Genes Critical for EMT, Including Jnk1 (Mapk8). *PLoS One* 2013;8:e57329.
42. Luo W, Chang R, Zhong J, et al. Histone demethylase JMJD2C is a coactivator for hypoxia-inducible factor 1 that is required for breast cancer progression. *Proc Natl Acad Sci USA* 2012;109:E3367–76.
43. Yamamoto S, Tateishi K, Kudo Y, et al. Histone demethylase KDM4C regulates sphere formation by mediating the crosstalk between Wnt and Notch pathways in colonic cancer cells. *Carcinogenesis* 2013;34:2380–8.

