A three-gene signature from protein–protein interaction network of LOXL2- and actin-related proteins for esophageal squamous cell carcinoma prognosis

Xiuhui Zhan1,2,a, Jiwei Jiao1,2,a, Hai-feng Zhang2, Chun-quan Li2,3, Jian-mei Zhao1,2, Lian-di Liao2,4, Jian-ji Wu1,2, Bing-ji Wu1,2, Zhi-yong Wu5, Shao-hong Wang6, Ze-peng Du6, Jin-hui Shen6, Haiying Zou1,2, Gena Neufeld7, Liyan Xu2,4 & Enmin Li3,2

1Department of Biochemistry and Molecular Biology, Shantou University Medical College, Shantou, Guangdong, China
2The Key Laboratory of Molecular Biology for High Cancer Incidence Coastal Chaoshan Area, Shantou University Medical College, Shantou, Guangdong, China
3College of Medical Informatics, Daqing Campus, Harbin Medical University, Daqing, Heilongjiang, China
4Institute of Oncologic Pathology, Shantou University Medical College, Shantou, Guangdong, China
5Department of Tumor Surgery, Shantou Central Hospital, Affiliated Shantou Hospital of Sun Yat-sen University, Shantou, Guangdong, China
6Department of Pathology, Shantou Central Hospital, Affiliated Shantou Hospital of Sun Yat-sen University, Shantou, Guangdong, China
7Cancer Research and Vascular Biology Center, The Bruce Rappaport Faculty of Medicine, Technion, Israel Institute of Technology, Haifa, Israel

Keywords
Actin-related proteins, esophageal squamous cell carcinoma, lysyl oxidase-like 2, prognosis, protein–protein interaction network, three-gene signature

Abstract
Current staging is inadequate for predicting clinical outcome of esophageal squamous cell carcinoma (ESCC). Aberrant expression of LOXL2 and actin-related proteins plays important roles in ESCC. Here, we aimed to develop a novel molecular signature that exceeds the power of the current staging system in predicting ESCC prognosis. We found that LOXL2 colocalized with filamentous actin in ESCC cells, and gene set enrichment analysis (GSEA) showed that LOXL2 is related to the actin cytoskeleton. An ESCC-specific protein–protein interaction (PPI) network involving LOXL2 and actin-related proteins was generated based on genome-wide RNA-seq in 15 paired ESCC samples, and the prognostic significance of 14 core genes was analyzed. Using risk score calculation, a three-gene signature comprising LOXL2, CDH1, and FN1 was derived from transcriptome data of patients with ESCC. The high-risk three-gene signature strongly correlated with poor prognosis in a training cohort of 60 patients (P = 0.003). In mRNA and protein levels, the prognostic values of this signature were further validated in 243 patients from a testing cohort (P = 0.001) and two validation cohorts (P = 0.021, P = 0.007). Furthermore, Cox regression analysis revealed that the signature was an independent prognostic factor. Compared with using the signature or TNM stage alone, the combined model significantly enhanced the accuracy in predicting ESCC prognosis. In conclusion, our data reveal that the tumor-promoting role of LOXL2 in ESCC is mediated by perturbing the architecture of actin cytoskeleton through its PPIs. We generated a novel three-gene signature (PPI interfaces) that robustly predicts poor clinical outcome in ESCC patients.
Introduction

Esophageal cancer is the sixth most common cause of death from cancer worldwide, and consists of two primary cancer types, esophageal adenocarcinoma and esophageal squamous cell carcinoma (ESCC). Cancer of the esophagus is a highly aggressive malignancy with very poor survival, and where the overall ratio of mortality to incidence is as high as 0.88 [1]. The current staging system is inadequate for predicting outcome of treatment, and the addition of cancer biology will make great sense to ESCC therapeutic discovery [2].

Lysyl oxidase-like 2 (LOXL2), a member of lysyl oxidase family, is involved in extracellular matrix (ECM) remodeling, transcription regulation, tumor progression, and metastasis in various types of carcinomas [3–6]. We previously demonstrated that LOXL2, with reduced nuclear and ectopic cytoplasmic expression, plays a pivotal role in lymph node metastasis and poor prognosis in ESCC [7]. There is also current evidence that intracellular, but not secreted LOXL2, promotes migration and invasion of clear cell renal cell carcinoma [8]. Thus, it is worth considering that the intracellular (cytoplasmic) distribution of LOXL2 in ESCC may have a novel and complex biological function independent of its established secreted role.

On the other hand, a large amount of protein–protein interactions (PPIs) participates in driving tumorigenesis through the regulation of oncogenic networks [9]. Some cancer-associated PPI network (PPIN) maps have been established by large-scale genomic experiments [10]. Actins are highly conserved proteins that are involved in maintenance of the cytoskeleton, in which Actin beta (encoded by ACTB) and Actin gamma 1 (encoded by ACTG1) are found as the two nonmuscle cytoskeletal actin in vertebrates [11]. Moreover, aberrant expression of actin-related proteins plays an important role in tumor progression as the actin cytoskeletal architecture impacts many cancer-related cellular processes [12, 13]. Technological advances in the field of bioinformatics, including microarray analysis of gene expression, now enable the identification of molecules and important prognostic PPIs in clinical settings.

In our current study, we investigated the prognostic value of a signature, based on biological experiments and PPIN, among different clinical cohorts in ESCC. We found that LOXL2 is related to actin cytoskeleton function in ESCC cells. Subsequently, we examined the prognostic significance of 14 candidate proteins from LOXL2- and actin-related proteins within ESCC-specific PPIN.

Moreover, a developed risk score calculating strategy was adopted and a three-gene signature was derived to accurately predict overall survival in ESCC.

Material and Methods

Cell lines and cell culture

Human ESCC cell lines KYSE150 [14] and SHEEC [15] were used in the study. KYSE150 cells were cultured in RPMI-1640 medium with 10% fetal bovine serum, whereas SHEEC cells were maintained in high-glucose DMEM medium with 4.0 mmol/L L-Glutamine and Ham’s Nutrient Mixture F12 with 10% newborn bovine serum.

Patients and tissue specimens

Human ESCC tumors and adjacent nontumorous esophageal epithelial tissues were collected directly after surgical resection between January 2000 and January 2013, at Shantou Central Hospital (Sun Yat-sen University, Shantou, China). Fifteen paired tissues between 2012 and 2013 were used for genome-wide transcriptome sequencing (RNA-seq, SPA accession number: SRP064894, Shanghai Biotechnology Corporation, Shanghai, China) [16]. We validated the three-gene signature using a validation tissue microarray cohort of 124 randomly selected patients with clear pathologic diagnosis and follow-up information (Table S1). All tumors were confirmed as ESCC by pathologists in the Clinical Pathology Department of the hospital, and the tissue specimens were classified according to the tumor-node metastasis (TNM) classification of the International Union against Cancer. Written informed consent was obtained from all patients or their appropriate surrogates. The study was approved by the ethical committee of Shantou Central Hospital and Shantou University Medical College.

Gene expression data

Published gene expression datasets, comprising 316 patients of ESCC in total, were from different centers and platforms in the study. RNA-seq data and clinical information of 84 patients with ESCC were downloaded from The Cancer Genome Atlas (TCGA) database [17]. The gene microarray expression profiles and corresponding clinical information of patients with ESCC were obtained from the publicly available Gene Expression Omnibus (GEO) database, including 53 patients from GSE23400 (Affymetrix HG-U133A...
from clinical and biologic data and sample identification. The analyses were done blinded as a validation cohort for assessing the prediction performance. The freely available gene expression data of ESCC specimens. When the RPKM of one gene in 15 tumor samples is higher than 10, the gene will be regarded as an effective detecting candidate for the following analysis. Self-loops and duplicated edges of proteins in the networks were removed. Gene Ontology (GO) and KEGG pathway assignments were performed using Gene Functional Classification Tool from the DAVID Bioinformatics Resources database (Release 6.7). The GO terms in the Biological Process (GOTERM-BP-FAT) and KEGG pathways with $P < 0.05$ and enrichment score > 2 were considered as significantly enriched function annotations.

Tissue Microarrays and Immunohistochemistry

Tissue microarray (TMA) construction of esophageal carcinoma tissue has been described in our previous study [20]. The SuperPicTure Polymer Detector Kit and the Liquid DAB Substrate Kit (Zymed/Invitrogen, Carlsbad, CA) were used to conduct immunohistochemistry according to the manufacturer’s instructions. Primary antibodies used for immunohistochemistry staining in this study are listed below: anti-LOXL2 antibody (1:100; Abnova, Taipei, China), anti-LOXL2 C-terminal antibody (1:100; Abcam, Cambridge, UK), or anti-FlagM2 (1:200; Sigma-Aldrich Guangzhou, China) in blocking buffer, and then probed with Alexa Fluor 488-conjugated Affinipure donkey antirabbit secondary antibody (1:200; Jackson ImmunoResearch, West Grove, PA) and Actin-stain™ 555 phalloidin (1:200; Cytoskeleton, Denver, CO), followed by counterstaining with 0.1 $\mu$g/mL 4′,6-diamidino-2-phenylindole (Sigma Aldrich, Guangzhou, China). Confocal imaging was performed on a laser-scanning confocal microscope (LSM 880; Carl Zeiss, Oberkochen, Germany).

Bioinformatics analysis

Genes involved in regulation of the actin cytoskeleton were obtained from KEGG database (Release 76.0, KEGG database/map04810, gene numbers = 181). Gene set enrichment analysis (GSEA2.2.0, Cambridge, MA), a powerful analytical method, was performed using Pearson’s correlation to determine the degree of linear relationship between LOXL2 and actin cytoskeleton genes.

A PPI human reference set of 224,522 PPIs between 18,130 proteins was obtained from the Biological General Repository for Interaction Datasets (BioGRID, Release 3.4), and was mapped into Cytoscape software (Version 3.2.1, http://www.cytoscape.org) as a visualizing parent PPI network. LOXL2 PPIs from variable resources and PPIs of ACTB or ACTG1 were extracted from the parent PPI network to generate a LOXL2-ACTB/ACTG1 PPIN. The network then created a LOXL2- ACTB/ACTG1 PPIN in ESCC with a filtering method according to reads per kilobase of exon per million mapped reads (RPKM) of the RNA-seq of ESCC specimens. When the RPKM of one gene in 15 tumor samples is higher than 10, the gene will be regarded as an effective detecting candidate for the following analysis. Self-loops and duplicated edges of proteins in the networks were removed. Gene Ontology (GO) and KEGG pathway assignments were performed using Gene Functional Classification Tool from the DAVID Bioinformatics Resources database (Release 6.7). The GO terms in the Biological Process (GOTERM-BP-FAT) and KEGG pathways with $P < 0.05$ and enrichment score > 2 were considered as significantly enriched function annotations.

Calculation of risk scores

The freely available gene expression data of ESCC specimens in GSE53624 and GSE53622 were already log2 scale transformed [19]. High expression or low expression of a given gene was defined on the basis of the level of mRNA log2 expressions by the automatic optimum selection method in X-tile software (Release 3.6.1), and the
same was applied to immunohistochemistry scores of the protein. In all cohorts in the study, each gene of each patient was assigned a numerical value in which high expression equaled 1 and low expression equaled 0. To estimate overall survival of a signature formed by LOXL2, CDH1, and FN1 genes, we adopted a developed strategy using a univariate Cox proportional hazards regression analysis coefficient for the three genes [21]. A given patient’s prognosis risk score was then calculated as the sum of each gene’s score, which was derived by multiplying the gene expression level by its corresponding regression coefficient (prognostic risk score = ΣCox coefficient of gene Gi × expression value of gene Gi). Finally, the patients were separated into two groups (high-risk signature and low-risk signature) using the 50th percentile cut-off of the risk score of the three-gene signature as a threshold. To reduce the influence of heterogeneity among different patients or cohorts, the risk scores and the cut-off value derived from the training cohort were reestimated and then applied to the other three cohorts.

Statistical analysis

Unless otherwise stated, statistical analyses were performed with SPSS 22.0 software (IBM, Chicago, IL) or GraphPad Prism 5 software (La Jolla, CA). The differences between groups were analyzed by the Student t test or paired-samples t test. The chi-squared test or Fisher’s exact test was used to assess the correlations between category variables. Overall survival (OS) was estimated by the Kaplan–Meier method. Multivariate Cox proportional hazards regression (HR) analysis was applied to evaluate independent prognostic factors associated with overall survival, and the three-gene signature, age, gender, and tumor stage, was used as covariates. To estimate the sensitivity and specificity of the signature and area under the curve (AUC) value, receiver operating characteristic (ROC) analysis was adopted in three cohorts. The differences between the AUC of the three-gene signature, TNM stage, and a combination of both were assessed by MedCalc software (Vision 15.8, Ostend, Belgium). A P < 0.05 was considered to indicate statistical significance, and all statistical tests were two-sided.

Results

LOXL2 is related to the actin cytoskeleton in ESCC

To explore the subcellular location of LOXL2 in ESCC, we studied the human ESCC cell lines KYSE150 and SHEEC. LOXL2 in SHEEC cells was mainly localized at the median and basal cytoplasm and nucleus, but weak staining was also observed at the apical region (Fig. 1A upper panel).

Development and analyses of a LOXL2-ACTB/ACTG1 PPI network in ESCC

Based on the work above, we assumed that the relevance of LOXL2 and actin cytoskeleton in ESCC is through a number of their interacting proteins. We gathered certain and potential LOXL2 protein–protein interactors (PPIs) from variable protein databases and published reports (Fig. 2A). All these and acknowledged human PPI data from the BioGRID database were combined to provide the original data for subsequent analyses. A LOXL2-ACTB/ACTG1 PPI, from all human tissues, containing 504 nodes and 3568 edges was generated by mapping total LOXL2, ACTB, and ACTG1 PPIs to the parental PPI network (Fig. 2C left panel). To create an ESCC-specific subnetwork, we examined 21,157 gene expressions in RNA-seq of 15 patients with ESCC. The low-expression ones out of 504 proteins in ESCC tissues were excluded through an RPKM-filtering method, and the remaining 2580 PPIs between 362 proteins were simultaneously selected for the following survey (Fig. 2B, C right panel).

We found that the genes in the ESCC-specific network were intimately connected to regulation and reorganization of the cytoskeleton, as well as multiple typical carcinoma-related cell behaviors, such as cell adhesion, cell invasion, and cell migration (Fig. 2D and Table S3). To better know the network composition of the ESCC-specific network, we grouped the proteins in three subcategories upon the network topological properties: central proteins consisting of LOXL2, ACTB, and ACTG1, core proteins comprising 14 interacting proteins, and other 345 common proteins (Fig. 2C).

Derivation of a three-gene signature predictive of survival in ESCC

The 14 LOXL2-ACTB/ACTG1 PPI core proteins are listed as below: A2M, ANGPTL4, CAP1, CDH1, CTGF, CTNND1, DAB2, EZH2, FN1, HSP90AA1, PFN1, TAF11, TMSB4X, and TWHAZ. Due to its absence in microarray expression
profiles in our study (GSE53624 and GSE53622), the core protein PFN1 was left out at first. Central gene LOXL2 and the other 13 core genes were considered as potential ESCC prognostic biomarkers and were further studied.

For prognostic gene signature analysis, we obtained microarray data from 119 patients with ESCC (GEO accession number: GSE53624) and randomly assigned it to either the training cohort or the testing cohort. Among the 60 cases in the training cohort, potentially crucial genes, excluding CTGF, DAB2, and TMSB4X, were significantly differentially expressed between ESCC tumors and paired normal specimens (Fig. 3A and Fig. S2). Five from the differentially expressed genes were upregulated in tumor tissues, whereas the others were downregulated in contrast. Moreover, high-level expression of five genes was found to be significantly associated ($P < 0.05$ by the log-rank test) with outcome of patients with ESCC in the training cohort: four genes (FN1, LOXL2, TAF11, and TMSB4X) were associated with shorter OS and high-level expression of one gene, CDH1, was associated with a better clinical outcome (Fig. 3B and S3).

Summarizing the analytic results in the training cohort, LOXL2, CDH1, FN1, and TAF11 were in accordance with the inclusive criteria. On account of the biological functions of these genes, we ultimately selected LOXL2 (lysyl oxidase-like 2), CDH1 (E-cadherin 1, type 1), and FN1 (fibronectin 1) to integrate a three-gene signature in ESCC.

Afterward, the three-gene signature in the patients among the training cohort was evaluated with a risk score algorithm provided by univariate Cox regression. The Cox coefficients for LOXL2, CDH1, and FN1 were 1.089, 0.980, and −0.827, respectively. The 50th percentile of the prognostic risk scores was 0 (range from −0.827 to 2.069), therefore, all individuals were classified into two groups:
high-risk signature (risk score > 0) and low-risk signature (risk score ≤ 0). Patients with a low-risk gene signature (n = 45) had a longer median overall survival than patients with a high-risk gene signature (n = 15) (42.0 months vs. 23.3 months, P = 0.003) (Fig. 3B). Multivariate Cox proportional hazard regression analyses revealed that mortality risk assessed by the three-gene signature was, along with tumor stage, an independent predictor of survival (P < 0.001, HR = 4.48, 95% CI = 2.03–9.90) among ESCC patients in the training cohort (Table 1). In conclusion, high risk, predicted by the three-gene signature, predicts poor survival in ESCC patients.

Validation of the three-gene signature and survival
To test the three-gene signature for its prognostic value, we applied it to the microarray data of the testing cohort. The same risk score algorithm and criteria, as which were derived from the training cohort, classified 19 and 40 ESCC patients of the testing cohort into the high-risk and low-risk groups, respectively. The high-risk gene signature was associated with a median OS of 23.8 months, whereas the low-risk signature was associated with a median OS of 43.0 months (P = 0.001 by the log-rank test) (Fig. 4A and S4). The three-gene signature was then validated in another cohort comprising 60 patients with ESCC (GEO accession number: GSE53622). Similar to the above, patients in the low-risk group had significantly longer survival time than those in the high-risk group (median survival months: 37.7 vs. 26.2, P = 0.021) (Fig. 4B and S5A). The high-risk three-gene signature remained significantly associated with shorter overall survival in the testing cohort (P = 0.002, HR = 2.87, 95% CI = 1.47–5.61) and the validation cohort (P = 0.032, HR = 2.21, 95% CI = 1.07–4.58) (Table 1).

Not only in the mRNA level, but the three-gene signature was further validated in the protein level with a
tissue microarray cohort of 124 patients with ESCC. Each protein displayed two immunostaining phenotypes; that is, low staining and high diffuse staining (Fig. 4C). The OS in the high-risk signature group was significantly shorter than that in the low-risk signature group (median survival months: 46.5 vs. 73.6, \( P = 0.007 \)) (Fig. 4D and Fig. S5B). Most importantly, the high-risk group predicted by three-gene signature also represented a higher mortality risk than the low-risk group (\( P = 0.016, \text{HR} = 1.88, 95\% \text{CI} = 1.13-3.14 \)) (Table 1). Therefore, the three-gene signature is a robust high-risk factor for ESCC prognosis.
Combination of the three-gene signature and TNM stage among the four cohorts

In current clinical analysis, TNM stage classification is deemed to be the optimal prognostic indicator for cancer. As demonstrated in Table S4 and S5, there was no statistically evident relationship between expression level of the three-gene signature and any ESCC clinical characteristic in not only the testing cohort but also in the three validation cohorts. A prognostic model combining TNM stage and the three-gene signature was conducted and receiver operating characteristic (ROC) analysis was performed. The combined model had a higher area under curve (AUC) than both the signature and the TNM stage alone in the four cohorts (Fig. 5). In applying this criteria to ESCC patients, we found that the predictive ability of

Table 1. Hazard ratios for death from ESCC among four cohorts, according to multivariate Cox regression analysis.

| Variable                                      | Hazard Ratio (95%CI) | P-value |
|-----------------------------------------------|----------------------|---------|
| Training cohort                               |                      |         |
| High-risk three-gene signature                | 4.48 (2.03-9.90)     | < 0.001 |
| Tumor stage III                              | 2.76 (1.32-5.81)     | 0.007   |
| Testing cohort                                |                      |         |
| High-risk three-gene signature                | 2.87 (1.47-5.61)     | 0.002   |
| Validation cohort                             |                      |         |
| High-risk three-gene signature                | 2.21 (1.07-4.58)     | 0.032   |
| Older age                                     | 2.12 (1.03-4.36)     | 0.042   |
| Validation tissue microarray cohort           |                      |         |
| High-risk three-gene signature                | 1.88 (1.13-3.14)     | 0.016   |
| Tumor stage III                              | 2.75 (1.62-4.68)     | < 0.001 |

1Variables were selected with a stepwise selection method. CI denotes confidence interval.

Figure 4. Validation of the three-gene signature. Kaplan–Meier estimates of overall survival of ESCC patients according to the three-gene signature in the testing cohort (A) and the validation cohort (B). (C) Representative low/high expression of LOXL2, CDH1, and FN1 by immunohistological study in tissue microarrays. The bar indicates 50 μm. (D) Kaplan–Meier curves of the three-gene signature for overall survival in the validation tissue microarray cohort.
Figure 5. The predictive ability of the three-gene signature. The four receiver operating characteristics (ROC) curves in the training cohort (A), the testing cohort (B), the validation cohort (C), and the validation tissue microarray cohort (D). Comparison of sensitivity and specificity for survival prediction by the three-gene signature, TNM stage, and combination of the two factors. AUC shows the area under the ROC. CI, confidence interval; TNM, pathologic tumor, node, and metastasis.
the combined model was significantly better than the signature or the TNM stage alone. As an example, the predictive ability between the combined model and TNM stage alone was 0.696 versus 0.616 ($P < 0.05$, Fig. 5B).

**Discussion**

China has high rates of cancer mortality, of which esophageal cancer is the fourth most prevalent cancer, with a 5-year relative survival below 30% [22, 23]. In our previous study, we discovered that LOXL2 is mainly expressed in the cell cytoplasm in ESCC tissues and cells [7, 24]. In this study, we reconfirmed the intracellular location of LOXL2 protein in ESCC, which is separated from its conventional location in extracellular space [25]. We found that LOXL2 and filamentous actin (F-actin) colocalized in the cytoplasm near the cell membrane and inside ESCC cells near sites of cell–cell contact. As one of the three main cytoskeletal polymers, F-actin carries out central functions in cell structure, intracellular organization, and cell movement [26]. While cancer cells regulate their cytoskeletal structures, they stimulate cellular motility to metastasize away from the initial niche and migrate into adjacent tissues, and even invade distant tissues [27]. As we expected, LOXL2 expression positively correlates with the regulation of the actin cytoskeleton in ESCC patients. Recent studies using loss- or gain-of-functions approaches also indicate that LOXL2 regulates actin cytoskeletal reorganization in several cancers. LOXL2 cross-links ECM components and increases tissue stiffness, thereby enhancing the reorganization of cytoskeleton in hepatocellular carcinoma cells [28]. Whereas, LOXL2 knockdown markedly reduces stress fiber and focal adhesion formation in human clear cell renal cell carcinoma cells [8].

From another perspective, the number of high-quality binary PPIs in the literature has grown roughly linearly, and publications about cancer-related PPIs has also risen rapidly since the 1990s [9, 29]. As proteins almost never perform their function alone, analyzing their network-based protein function can give valuable insight into their roles in a cell [30]. Here, we developed an ESCC-specific PPIN of LOXL2- and actin-related proteins through RPKM filtering based on the RNA-seq data of 15 patients with ESCC. Genes in the ESCC-specific network are primarily involved in regulation and reorganization of the actin cytoskeleton, as well as common cancer-related cell behaviors. Therefore, we suggest that the intracellular (cytoplasmic) distribution of LOXL2 in ESCC probably has a novel biological function, independent of its classical secreted role as extracellular LOXL2, whereby LOXL2 indirectly causes perturbation of actin cytoskeleton through its PPIs in ESCC, and contributes to hyperactivation or loss-of-function of some key proteins in carcinogenic pathways to enhance oncogenicity. Nevertheless, further research is needed for the underlying molecular mechanism.

There is no doubt that immunohistochemistry, tissue microarrays, and RT-PCR are in general use to provide molecular signatures for cancer diagnosis and prognosis. However, the emergence and convergence of high-throughput techniques and bioinformatics predictions will expand the landscape of molecular signatures in cancer for therapeutic and prognostic discovery [31–33]. In this study, through the sum of experimental and bioinformatics results, we derived a three-gene signature (involving LOXL2, CDH1, and FN1), using a risk score calculation based on published microarray data of ESCC specimens. The presence of a high-risk three-gene signature in the ESCC patients of the training cohort was strongly associated with a decreased overall survival, and the prognostic capacity was validated in both mRNA and protein levels within a testing and two validation cohorts. Furthermore, the prognostic signature compensates for current ESCC staging and improves precision for prognosis of ESCC patients. Meanwhile, the signature more effectively predicted survival when combined with TNM stage. Thus, we believe that the prognostic capacity of the three-gene signature is amenable to patients with ESCC.

Lysyl oxidase-like 2 (LOXL2), a copper-dependent amine oxidase known to initiate cross-linking of collagen and elastin, induces desmoplastic stroma, fibrosis, dedifferentiation, and tumor progression in several solid cancers, including in ESCC [6, 34, 35]. E-cadherin (epithelial cadherin, also known as CDH1, E-cad) is a calcium-dependent cell–cell adhesion glycoprotein, and its transassociation in homophilic complexes, supported by actin filaments (F-actin) through catenin, is required for tissue stability [36–38]. LOXL2 is necessary and sufficient to repress E-cadherin under hypoxia condition in cancer [39]. Loss of E-cadherin, considered as a hallmark of EMT, contributes to cancer progression in by increasing proliferation, invasion, and metastasis [40, 41]. LOXL2 interacts with E47 to functionally collaborate in the repression of E-cadherin promoter, and fibronectin is essential for this LOXL2/E47-mediated modulation of early metastasis colonization [42]. Fibronectin 1 (FN1) mediates a wide variety of cellular interactions with the ECM, and is involved in cell adhesion, migration processes, and metastasis [43, 44]. Moreover, it is reported that altered expression of fibronectin 1 exists in several squamous cell carcinomas, including ESCC, head and neck squamous cell carcinoma, and oral squamous cell carcinoma [45–47]. The derivation of these three genes predictive of clinical outcome in patients with ESCC may reveal PPI anticancer targets for the development of therapy for esophageal cancer. Recently, a LOXL2-specific monoclonal antibody (AB0023/4) was shown to
be effective against primary tumor growth and reduce metastatic burden in mouse models [48]. Similarly, AB0024 synthesized by Gilead Sciences has begun to be used in clinical trials in humans with advanced solid tumors (ClinicalTrials.gov Identifier: NCT01323933). Since PPI interfaces as an anticancer strategy has become a reality [49], the three-gene signature, acting as a cancer-enabling PPI interface, may become a new promising therapeutic target in ESCC.

In conclusion, this is a pioneering study of a molecular signature for cancer prognosis based on both biological experiments and bioinformatics data. For the first time, our data reveal that the tumor-promoting role of LOXL2 in ESCC probably perturbs the architecture of the actin cytoskeleton through its PPIs and induces additional oncogenic effects by hyperactivation or loss-of-function of key proteins. The three-gene signature (PPI interfaces), derived by us, is robustly associated with the clinical outcome in patients with ESCC. Whether the development mode of molecular signatures can be universally used requires a study with a sufficient number of other cancers.

Acknowledgements

We thank Professor Stanley Li Lin for manuscript revision.

Conflict of Interest

None declared.

References

1. Torre, L. A., F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal. 2015. Global cancer statistics, 2012. CA Cancer J. Clin. 65:87–108.
2. Ohashi, S., S. I. Miyamoto, O. Kikuchi, T. Goto, Y. Amanuma, and M. Muto. 2015. Recent advances from basic and clinical studies of esophageal squamous cell carcinoma. Gastroenterology 149:1700–1715.
3. Barker, H. E., T. R. Cox, and J. T. Erler. 2012. The rationale for targeting the LOX family in cancer. Nat. Rev. Cancer 12:540–552.
4. Moreno-Bueno, G., F. Salvador, A. Martin, A. Floristán, E. P. Cuevas, V. Santos, et al. 2011. Lysyl oxidase-like 2 (LOXL2), a new regulator of cell polarity required for metastatic dissemination of basal-like breast carcinomas. EMBO Mol. Med. 3:528–544.
5. Herranz, N., N. Dave, A. Millanes-Romero, L. Morey, V. M. Diaz, V. Lorenz-Fonfría, et al. 2012. Lysyl oxidase-like 2 deaminates lysine 4 in histone H3. Mol. Cell 46:369–376.
6. Cano, A., P. G. Santamaria, and G. Moreno-Bueno. 2012. LOXL2 in epithelial cell plasticity and tumor progression. Future Oncol. 8:1095–1108.
7. Li, T., L. Xu, Z. Wu, L. Liao, J. Shen, X. Xu, et al. 2012. Reduced nuclear and ectopic cytoplasmic expression of lysyl oxidase-like 2 is associated with lymph node metastasis and poor prognosis in esophageal squamous cell carcinoma. Hum. Pathol. 43:1068–1076.
8. Hase, H., K. Jingushi, Y. Ueda, K. Kitae, H. Egawa, I. Oshbio, et al. 2014. LOXL2 status correlates with tumor stage and regulates integrin levels to promote tumor progression in ccRCC. Mol. Cancer Res. 12:1807–1817.
9. Ivanov, A. A., F. R. Khuri, and H. Fu. 2013. Targeting protein-protein interactions as an anticancer strategy. Trends Pharmacol. Sci. 34:393–400.
10. McLendon, R., A. Friedman, D. Bigner, E. G. Van Meir, D. J. Brat, G. M. Mastrogiannakis, et al. 2008. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455:1061–1068.
11. Sonnemann, K. J., D. P. Fitzsimons, J. R. Patel, Y. Liu, M. F. Schneider, R. L. Moss, et al. 2006. Cytoplasmic gamma-actin is not required for skeletal muscle development but its absence leads to a progressive myopathy. Dev. Cell 11:387–397.
12. Gaspar, P., and N. Tapon. 2014. Sensing the local environment: actin architecture and Hippo signalling. Curr. Opin. Cell Biol. 31:74–83.
13. Mouneimne, G., S. D. Hansen, L. M. Selfors, L. Petrak, M. M. Hickey, L. L. Gallegos, et al. 2012. Differential remodeling of actin cytoskeleton architecture by profilin isoforms leads to distinct effects on cell migration and invasion. Cancer Cell 22:615–630.
14. Shimada, Y., M. Imamura, T. Wagata, N. Yamaguchi, and T. Tobe. 1992. Characterization of 21 newly established esophageal cancer cell lines. Cancer 69:277–284.
15. Shen, Z., S. Cen, J. Shen, W. Cai, J. Xu, Z. Teng, et al. 2000. Study of immortalization and malignant transformation of human embryonic esophageal epithelial cells induced by HPV18 E6E7. J. Cancer Res. Clin. Oncol. 126:589–594.
16. Li, C., G. Huang, Z. Wu, Y. Xu, X. Li, Y. Xue, et al. 2017. Integrative analyses of transcriptome sequencing identify novel functional lncRNAs in esophageal squamous cell carcinoma. Oncogenesis 6:e297.
17. Kim, J., R. Bowlby, A. J. Mungall, A. G. Robertson, R. D. Odze, A. D. Cherniack, et al. 2017. Integrated genomic characterization of oesophageal carcinoma. Nature 541:169–175.
18. Su, H., N. Hu, H. H. Yang, C. Wang, M. Takikita, Q. H. Wang, et al. 2011. Global gene expression profiling and validation in esophageal squamous cell carcinoma and its association with clinical phenotypes. Clin. Cancer Res. 17:2955–2966.
19. Li, J., Z. Chen, L. Tian, C. Zhou, M. Y. He, Y. Gao, et al. 2014. LncRNA profile study reveals a
Three- gene signature associated with the survival of patients with oesophageal squamous cell carcinoma. Gut 63:1700–1710.

Zhang, F. R., L. H. Tao, Z. Y. Shen, Z. Lv, L. Y. Xu, and E. M. Li. 2007. Fascin expression in human embryonic fetal and normal adult tissue. J. Histochem. Cytocem. 56:193–199.

Chen, H., S. Yu, C. Chen, G. Chang, C. Chen, A. Yuan, et al. 2007. A five- gene signature and clinical outcome in non-small-cell lung cancer. N. Engl. J. Med. 356:11–20.

Zhou, M., H. Wang, J. Zhu, W. Chen, L. Wang, S. Liu, et al. 2016. Cause- specific mortality for 240 causes in China during 1990-2013: a systematic subnational analysis for the Global Burden of Disease Study 2013. Lancet 387:251–272.

Zheng, R., H. Zeng, S. Zhang, T. Chen, and W. Chen. 2016. National estimates of cancer prevalence in China, 2011. Cancer Lett. 370:33–38.

Lv, G., H. Zhou, L. Liao, H. Cao, F. Zeng, B. Wu, et al. 2014. Identification of a novel lysyl oxidase-like 2 alternative splicing isoform, LOXL2Δ13, in esophageal squamous cell carcinoma. Biochem. Cell Biol. 92:379–389.

Lucero, H. A., and H. M. Kagan. 2006. Lysyl oxidase: an oxidative enzyme and effector of cell function. Cell. Mol. Life Sci. 63:2304–2316.

Fletcher, D. A., and R. D. Mullins. 2010. Cell mechanics and the cytoskeleton. Nature 463:485–492.

Gross, S. R. 2014. Actin binding proteins. Cell Adh. Migr. 7:199–213.

Wong, C. C., A. P. Tse, Y. Huang, Y. Zhu, D. K. Chiu, R. K. Lai, et al. 2014. Lysyl oxidase-like 2 is critical to tumor microenvironment and metastatic niche formation in hepatocellular carcinoma. Hepatology 60:1645–1658.

Rolland, T., M. Tasxan, B. Charloiteaux, S. J. Pevzner, and Q. Zhong. 2014. A proteome- scale map of the human interactome network. Cell 159:1212–1226.

Sharan, R., I. Ulitsky, and R. Shamir. 2007. Network- based prediction of protein function. Mol. Syst. Biol. 3:88.

Landemaine, T., A. Jackson, A. Bellahcene, N. Rucci, S. Sin, B. M. Abad, et al. 2008. A six- gene signature predicting breast cancer lung metastasis. Cancer Res. 68:6092–6099.

Garcia, I., G. Mayol, J. Rios, G. Domenech, N. K. V. Cheung, A. Oberthuer, et al. 2012. A three- gene expression signature model for risk stratification of patients with neuroblastoma. Clin. Cancer Res. 18:2012–2023.

Chang, W., X. Gao, Y. Han, Y. Du, Q. Liu, L. Wang, et al. 2014. Gene expression profiling- derived immunohistochemical signature with high prognostic value in colorectal carcinoma. Gut 63:1457–1467.

Martin, A., F. Salvador, G. Moreno-Bueno, A. Floristan, C. Ruiz-Herguido, E. P. Cuevas, et al. 2015. Lysyl oxidase-like 2 represses Notch1 expression in the skin to promote squamous cell carcinoma progression. EMBO J. 34:1090–1109.

Kasashima, H., M. Yashiro, H. Kinoshita, T. Fukuoka, T. Morisaki, G. Masuda, et al. 2014. Lysyl oxidase-like 2 (LOXL2) from stromal fibroblasts stimulates the progression of gastric cancer. Cancer Lett. 354:438–446.

Lecuit, T., and A. S. Yap. 2015. E- cadherin junctions as active mechanical integrators in tissue dynamics. Nat. Cell Biol. 17:533–539.

van Roy, F., and G. Berx. 2008. The cell- cell adhesion molecule E- cadherin. Cell. Mol. Life Sci. 65:3756–3788.

Cavey, M., M. Rauzi, P. O. Lenne, and T. Lecuit. 2008. A two- tiered mechanism for stabilization and immobilization of E- cadherin. Nature 453:751–756.

Schietke, R., C. Warnecke, I. Wacker, J. Schodel, D. R. Mole, V. Campean, et al. 2010. The lysyl oxidases LOX and LOXL2 are necessary and sufficient to repress E- cadherin in hypoxia: insights into cellular transformation processes mediated by HIF- 1. J. Biol. Chem. 285:6658–6669.

Thiery, J. P. 2002. Epithelial- mesenchymal transitions in tumour progression. Nat. Rev. Cancer 2:442–454.

Canel, M., A. Serrels, M. C. Frame, and V. G. Brunton. 2013. E- cadherin- integrin crosstalk in cancer invasion and metastasis. J. Cell Sci. 126:393–401.

Canesin, G., E. P. Cuevas, V. Santos, C. López- Menéndez, G. Moreno-Bueno, Y. Huang, et al. 2015. Lysyl oxidase-like 2 (LOXL2) and E47 EMT factor: novel partners in E- cadherin repression and early metastasis colonization. Oncogene 34:951–964.

Pankov, R. 2002. Fibronectin at a glance. J. Cell Sci. 115:3861–3863.

Ruiz-Garcia, E., V. Scott, C. Machavoine, J. M. Bidart, L. Lacroix, S. Delaloge, et al. 2010. Gene expression profiling identifies Fibronectin 1 and CXCL9 as candidate biomarkers for breast cancer screening. Br. J. Cancer 102:462–468.

Zhang, J., H. Zhi, C. Zhou, F. Ding, A. Luo, X. Zhang, et al. 2005. Up- regulation of fibronectin in oesophageal squamous cell carcinoma is associated with activation of the Erk pathway. J. Pathol. 207:402–409.

Jennings, F., R. Ceder, S. Garvin, R. Grenman, R. C. Grafstrom, and K. Roberg. 2010. Fibronectin 1 is a potential biomarker for radioresistance in head and neck squamous cell carcinoma. Cancer Biol. Ther. 10:1244–1251.

Morita, Y., K. Hata, M. Nakamichi, T. Omata, N. Morita, Y. Yura, et al. 2015. Cellular fibronectin 1 promotes VEGF-C expression, lymphangiogenesis and lymph node metastasis associated with human oral
squamous cell carcinoma. Clin. Exp. Metastas 32:739–753.
48. Barry-Hamilton, V., R. Spangler, D. Marshall, S. McCauley, H. M. Rodriguez, M. Oyasu, et al. 2010. Allosteric inhibition of lysyl oxidase-like-2 impedes the development of a pathologic microenvironment. Nat. Med. 16:1009–1017.
49. Wells, J. A., and C. L. McClendon. 2007. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450:1001–1009.

Supporting Information

Additional supporting information may be found in the online version of this article:

Figure S1. Colocalization of LOXL2 and F-actin in confocal microscopy of ESCC cell. (a) Co-localization of LOXL2 (C-terminal) and F-actin in confocal microscopy of SHEEC cells with 2D section images. (b) Co-localization of LOXL2-Flag and F-actin in confocal microscopy of KYSE150 cells with 2D section images.

Figure S2. Analysis of 11 other LOXL2-ACTB/ACTG1 PPIN core genes expression between tumor and paired normal samples in the training cohort.

Figure S3. Kaplan–Meier curves and log-rank tests of 11 other LOXL2-ACTB/ACTG1 PPIN core genes for overall survival in the training cohort.

Figure S4. Three genes of the signature in a testing cohort of ESCC patients.

Figure S5. Kaplan–Meier curves and log-rank tests of three genes of the signature for overall survival in the validation cohort (a) and the validation tissue microarray cohort (b).

Table S1. Patient demographics and clinical characteristics for validation tissue microarray cohort in ESCC.

Table S2. Patient demographics and clinical characteristics for the training cohort, testing cohort, and validation cohort.

Table S3. Gene Ontology (GO) biological process and KEGG pathway analysis for genes of LOXL2-ACTB/ACTG1 PPIN in ESCC.

Table S4. Clinical and pathological characteristics of patients with ESCC with high- or low-risk three-gene signature in the training cohort, testing cohort, and validation cohort.

Table S5. Clinicopathological characteristics of ESCC patients with high- or low-risk three-gene signature in the validation tissue microarray cohort (n = 124).