LOXL2 IS HIGHLY EXPRESSED IN CANCER-ASSOCIATED FIBROBLASTS AND ASSOCIATES TO POOR COLON CANCER SURVIVAL

Sofía Torres1*, Irene García-Palmero1, Mercedes Herrera2, Rubén A. Bartolomé1, Cristina Peña2, Mª Jesús Fernandez-Aceñero3, Guillermo Padilla4, Alberto Peláez-García1, María Lopez-Lucendo4, Rufo Rodriguez-Merlo5, Antonio García de Herreros6, Félix Bonilla2 and J. Ignacio Casal1*

2. Department of Medical Oncology. Hospital Puerta de Hierro Majadahonda. Madrid. Spain
4. Core facilities. CIB-CSIC. Madrid. Spain
5. Department of Pathology. Hospital Virgen de la Salud. Toledo. Spain
6. IMIM-Hospital del Mar. Barcelona. Spain

RUNNING TITLE: LOXL2 in colon cancer prognosis

KEYWORDS: LOXL2; prognostic biomarker; cancer-associated fibroblasts; colon cancer

Word count: 5107 Figures: 5 Tables: 1

CONFLICT OF INTEREST: The authors declare no conflict of interest.

*Corresponding author:
J. Ignacio Casal
Department of Cellular and Molecular Medicine
Centro de Investigaciones Biológicas (CIB-CSIC)
Ramiro de Maeztu, 9
28040 Madrid. Spain
Phone: +34 918373112
Fax: +34 91 5360432
e-mail: icasal@cib.csic.es
LOXL2 in colon cancer prognosis

**Translational relevance**

There is an urgent need to find adequate biomarkers for correct risk stratification of colon cancer patients, particularly at stage II and III. Cancer-associated fibroblasts (CAFs) have been associated to recurrence and survival. After proteomic characterization of purified colon cancer-associated fibroblasts we identified a number of stromal-associated proteins that can be used as biomarkers. The validity of this model was supported by the similar protein profile obtained after TGFβ activation of normal fibroblasts. Among the selected stromal biomarkers, LOXL2 and TAGLN exhibited strong predictive value as molecular classifiers. High expression of LOXL2 showed a clear association with poor prognosis, either disease-free or overall survival. In addition, it displayed a significant prognostic value in the identification of high-risk stage II and III patients. LOXL2 has the potential to be used in clinical practice for risk classification, which could lead to a more aggressive therapy and better outcome of the patients.
LOXL2 in colon cancer prognosis

ABSTRACT

Purpose: Cancer associated-fibroblasts (CAFs) are major mediators in tumor microenvironment. We investigated the changes in protein expression in colon cancer associated-fibroblasts compared to normal fibroblasts (NFs) in the context of searching for prognostic biomarkers, particularly for stage II patients.

Experimental design: CAFs and NFs isolated from colon cancer patients were used to identify differentially-expressed proteins using quantitative proteomics. Stromal expression of deregulated proteins was analyzed by immunohistochemistry. Prognostic impact was studied using external gene expression datasets for training, then quantitative PCR and immunohistochemistry for validation in different cohorts of patients. Combined datasets were used for prediction of risk assessment at stage II and III.

Results: A desmoplastic signature composed of 32 proteins, highly specific for stromal components in colon cancer, was identified. These proteins were enriched for extracellular matrix organization components, TGFβ signaling pathway, fibrosis and wound healing proteins. The expression in CAFs of 11 up-regulated proteins and 4 down-regulated proteins, selected for biomarker validation, was verified by orthogonal techniques. LOXL2 displayed a high prognostic impact by using external independent datasets and further validation in two different cohorts of patients. High expression of LOXL2 was associated with higher recurrence $P = 0.001$ HR: 5.38 (1.70-17.01) and overall survival $P = 0.001$ HR: 8.52 (1.90-38.29). Immunohistochemistry analysis revealed a prognostic value for LOXL2 in stage II patients.
Conclusions: We identified LOXL2 to be associated with the outcome of colon cancer patients. Furthermore, it can be used to stratify patients at stages II and III for further therapeutic decisions.
INTRODUCTION

Cancer-associated fibroblasts (CAFs) play an important role in the development and progression of tumors (1, 2). In many solid tumors, cancer progression is accompanied by the activation of stromal fibroblasts into myofibroblasts or cancer-associated fibroblasts (CAFs), characterized by “de novo” expression of α-smooth muscle actin (α-SMA) (3). Fibroblasts are involved in tissue remodeling and repair. CAFs are also responsible for the increased deposition of extracellular matrix (ECM) components in the tumor invasion front, known as desmoplastic reaction (4). CAFs acquire a phenotype similar to wound healing-activated myofibroblasts. While this activation is temporal in a normal wound; in cancer, the fibroblasts remain endlessly activated. They synthesize multiple factors that stimulate and accelerate the healing process, angiogenesis and immune cell infiltration, all contributing to create an appropriate microenvironment for tumor growth and invasion.

A tumor cannot develop without the parallel expansion of a tumor stroma, as described for breast (5), lung (6), prostate (7) and colon cancer (8), among others. TGFβ regulates fibrosis and is the most important factor affecting CAF activation (9), promoting the expression of factors involved in paracrine signaling, ECM production and remodeling (10). Stromal changes have been associated with tumor progression and prognostic impact. Unfortunately, only a few selective biomarkers of myofibroblasts are well established: α-SMA, fibroblast activation protein (FAP) and the fibroblast specific protein (FSP1), also known as S100A4, are the most commonly and widely used CAF markers in solid cancers. However, these proteins are not only present in CAFs but also in pericytes or endothelial cells (11). Moreover, CAFs show a variable expression of these markers in different tumors.
LOXL2 in colon cancer prognosis

The epithelial-stromal ratio of colon carcinomas has been used as a predictor of survival, independent from lymph node status and tumor stage (8, 12). Some studies have correlated the number of stromal myofibroblasts (13), vimentin expression (14) and FAP expression (15) with colon cancer prognosis. In other exploratory studies for stromal biomarkers, Nakagawa et al. (16) compared metastatic colon cancer fibroblasts with liver and skin fibroblasts to get a gene expression profile that included many adhesion molecules and ECM-remodeling genes. Recently, a comparative transcriptomic analysis of paired CAFs and normal fibroblasts (NFs) identified 108 genes, mostly down-regulated (17). In a proteomic study, colon cancer CAFs were compared with bone-marrow precursors to identify tenascin C, ED-A fibronectin and SDF-1 as deregulated in CAFs (18). Recently, we used a murine model of sporadic colon cancer based on azoxymethane/ dextran sodium sulfate to optimize fibroblast purification and characterize alterations in murine CAFs (19).

An “in depth” proteomic analysis of human CAFs should give us a more complete picture of stromal biomarkers for prognostic classification in order to help us in the stratification of high- and low-risk colon cancer patients. This is particularly relevant for stage II and III patients, where chemotherapy choices have to be made. Proteomics displays several advantages over transcriptomics for a rapid clinical translation. Quantified proteins are easier to adapt to standard laboratory techniques. Direct protein identification avoids problems of translational stability and poor correlation between mRNA and protein abundance (20).

Here, we investigated the differences in the protein expression profile between colon cancer CAFs and NFs by using iTRAQ (isobaric Tags for Relative and Absolute Quantification), which allows the concurrent identification and relative quantification of the proteome from different biological samples in a single experiment (21). A problem
LOXL2 in colon cancer prognosis

associated to human CAFs is the heterogeneity and relatively large variability between patients. To overcome this problem, we used pools of samples for the proteomic study. Pool results were then confirmed in individual patient analysis. A large collection of proteins deregulated in colon cancer stroma as compared to healthy tissue were identified and validated by different approaches, including TGFβ-activation of NFs. We obtained a panel of stromal markers with a prognostic impact using different datasets. LOXL2 exhibits a great promise in colon cancer prognosis and survival prediction, including stage II-only patients.
LOXL2 in colon cancer prognosis

MATERIALS AND METHODS

Patients and tissue samples

For proteomic analysis, surgically resected biopsies of colon cancer and paired non-cancerous tissues (collected 10 cm apart from the tumor) were collected from 12 patients at stages II and III from Hospital Virgen de la Salud (Supplementary Table S1A). For LOXL2 validation, two sample collections from patients followed for more than five years were obtained from the Hospital Fundación Jimenez Diaz, 70 for quantitative PCR and 121 for immunohistochemistry analysis in tissue microarray format (Supplementary Table S1B, C, respectively). Informed written consent was obtained from all participants, as required and approved by the Research Ethics Committees of Virgen de la Salud Hospital (Toledo) and the Hospital Fundación Jimenez Diaz (Madrid), respectively. The histologic diagnosis for each sample was reconfirmed using microscopic examination of hematoxylin/eosin-stained sections of each tissue block.

External cohorts of validation

We used three independent external cohorts in the prognostic study. For biomarker training, we used the GSE17538 superserie, which contains a cohort of 232 patients of colon cancer (Moffitt samples and VMC samples). For risk stratification at stage II and III, we used the Amsterdam cohort GSE33113 with 90 patients at stage II and the Berlin cohort GSE12945, which contains data for 21 patients at stage III (22).

Fibroblast isolation and cell culture

CAFs and NFs were obtained from colonic tissues by the explant technique as described (19, 23, 24). See Supplementary Methods for additional information. Colon cancer cell
LOXL2 in colon cancer prognosis

lines SW480, SW620, KM12C and KM12SM were cultured as BJ-hTERT cells. KM12C and KM12SM human colon cancer cells were obtained from I. Fidler’s lab (MD Anderson Cancer Center. Houston, TX, USA). Other cell lines were obtained directly from the American Type Culture Collection (ATCC). All these cell lines were authenticated by short tandem repeat analysis or characterized by karyotype analysis. These cell lines were passaged fewer than 6 months after purchase for all the experiments

**ITRAQ proteomic analysis**

A detailed description of sample preparation and mass spectrometry analysis is given in Supplementary Methods.

**Immunohistochemistry analysis**

Each sample was deparaffinized for antigen retrieval using sodium citrate (pH 6.0) for 20 min and subsequent incubation with the respective primary antibody: ACAN (1:100; MAB1220; R&D system), αSMA (Clone 1A4; DAKO), COL14A1 (1:100; HPA023781; Sigma), DKK3 (1:100; sc-25518; Santa Cruz Biotechnology), CDH13 (1:100; HPA001380; Sigma), TAGLN (1:200; AF7886; R&D system), TGM2 (1:250; ab73170; Abcam) or LOXL2 (1:200; NBP1-32954; Novus Biologicals). The reaction was developed using diaminobenzidine (DAB) as chromogen and hematoxylin for counterstaining. In all cases, sections from normal colonic mucosa distant from the tumor site were used as negative controls. In some cases, Human Protein Atlas (HPA) was used for meta-analysis of expression of deregulated proteins in colorectal tissues (25).

**Quantitative PCR**
LOXL2 in colon cancer prognosis

Total RNA was isolated from cellular cultures using the mirVana isolation kit (Ambion). RNA from formalin-fixed paraffin-embedded (FFPE) tissue was isolated from 3x10µm-sections using the NucleoSpin totalRNA FFPE XS kit (Macherey-Nagel). cDNA was subsequently obtained using SuperScript II First Strand Synthesis System with random hexamers (Invitrogen). Real-time quantitative PCR (QPCR) was performed using the FastStart Master Mix (Roche) with probes from the Universal Probe Library Set (Roche). Amplifications were run in a 7900 HT-Fast Real-Time PCR System (Applied Biosystems). Each value was adjusted using GAPDH, β-actin and 18S RNA levels as reference. Primer sequences and probes used in this study can be found in Supplementary Methods. Expression of the selected genes was considered positive when the tumor/normal ratio showed a fold-change ≥2.

Prognostic analyses using public datasets

To remove technical variations between datasets, expression levels for all probes within each sample (patient) were transformed to z-score by subtracting the overall average probe intensity from each expression value and dividing the difference by the standard deviation of all the intensities. Subsequently, z-score ratios were obtained by subtracting the average z-score for the cohort from each individual z-score and dividing the result by the standard deviation of the z-scores for the cohort (26). Positive (z score > 0) and negative (z score < 0) ratios indicate a significantly higher or lower expression level of genes in the module, respectively. We derived Kaplan-Meier survival curves for patients, dividing the cohort into two risk groups, for each selected markers. The log-rank test $p$ value was determined for differences in time to recurrence (stage I-III patients) and death (stage I-IV patients) between the two risk groups. Univariant Cox proportional hazard regression models were fitted for estimation of independent hazard ratios (HR) for each marker. Kaplan-Meier survival curves were plotted with IBM
LOXL2 in colon cancer prognosis

SPSS Statistics 20. Log-rank tests and Cox regressions were performed with the SAS® software version 9.3 (SAS). A multivariable Cox regression analysis, adjusted for age, gender, AJCC stage and differentiation grade was used to calculate hazard ratios (HR).
RESULTS

Differentially-expressed proteins in colon CAFs using iTRAQ quantification

Biopsies corresponding to 12 colon cancer patients (Supplementary Table S1A) were cut in small fragments and fibroblasts were isolated using the explant technique. To study differences in protein expression between CAFs and NFs, we prepared pools from complete cell extracts and conditioned medium collected from 6 paired fibroblast cultures at 95-100% confluence. Protein extracts were trypsin-digested. Then, peptides were iTRAQ-labeled and fractionated with OFFGEL (pH 3-10). To avoid biases in peptide labeling, we performed biological replicates using the 114 and 116 iTRAQ tags for NFs and 115 and 117 for CAFs (Supplementary Fig. S1). Each fraction was analyzed by duplicate on a LTQ-Orbitrap Velos. In total, 1780 and 305 proteins were identified in the whole cellular extract and conditioned medium, respectively, using Proteome Discoverer v1.4 and MASCOT (Supplementary Fig. S2). Total quantified proteins are listed in Supplementary Table S2 (whole cell extract and conditioned medium). After applying a fold-change ≥ 1.5, we found 57 and 43 proteins deregulated in complete cell lysates and conditioned medium, respectively (Supplementary Table S3). Sixty percent of the proteins were up-regulated, including extracellular matrix (ECM) proteoglycans like aggrecan (ACAN), biglycan (BGN) or chondroitin sulfate proteoglycan 4 (CSPG4); proteins involved in organization and modification of ECM such as lysyl oxidase-like 2 protein (LOXL2) and procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2 (PLOD2); proteins implicated in cytoskeleton regulation and organization such as tropomyosin 1 (TPM1), palladin (PALLD), myosin regulatory light polypeptide 9 (MYL9) or transgelin (TAGLN); proteins related with wound healing like transglutaminase 2 (TGM2), myosin light chain kinase (MYLK) or thrombospondin 1 (THBS1) and other proteins like CDH13 or DKK3. Among the
LOXL2 in colon cancer prognosis

down-regulated proteins there were also ECM proteins like decorin (DCN), collagen type XIV (COL14A1) or periostin (POSTN) and cytokines like colony-stimulating factor 1 (CSF1).

After literature and data-mining search for novel proteins, not previously associated to colon cancer, we selected 15 proteins (9 from whole cellular extract and 6 from secretome) for validation by Western Blot and semi-quantitative PCR. Western blot results, either in pools or individual samples of CAFs and NFs, were consistent with the iTRAQ quantification data. CDH13, TAGLN and TGM2 proteins were up-regulated, whereas COL14A1 was down-regulated in the cell lysate. ACAN, DKK3 and LOXL2 were up-regulated and POSTN down-regulated in the secretome (Fig. 1A). In addition, we used semi-quantitative PCR to verify the changes in expression of these 15 proteins (Fig. 1B). In general, mRNA ratios of expression followed a similar trend to those observed by iTRAQ, confirming the alterations in expression for these proteins.

**Functional annotation and protein-protein interaction networks in colon CAFs**

We carried out functional annotation of colon CAFs-deregulated proteins based on GO analysis using Genomatix software (**Supplementary Table S4**). Regarding “biological processes” the top rank was for “ECM organization” (27 genes, p-value 3.09E-23) (**Supplementary Fig. S3A**), followed by “platelet activation” (17 genes, p-value 7.43E-14), “cell adhesion” (28 genes, p-value 8.79E-12) or “wound healing” (23 proteins, p-value 7.77E-12) (**Supplementary Fig. S3B**). Regarding signal transduction pathways, the TGFβ/SMAD signaling pathway (p-value 7.58E-12) with 26 deregulated proteins (16 up-regulated) was the most represented (**Supplementary Fig. S3C**), followed by “Focal adhesion kinase” or “Integrin-linked kinase”. Finally, regarding “Overrepresented diseases”, “Fibrosis” (26 proteins, p-value 1.54E-18)
LOXL2 in colon cancer prognosis

(Supplementary Fig. S3D) and “Wounds and Injuries” (27 proteins, p-value 6.97E-18) were the top alterations. “Neoplasm metastasis” was also highly represented with 27 proteins but a lower p-value 4.89E-10. In summary, GO analysis confirmed ECM organization, TGFβ/SMAD signaling pathway and fibrosis and wound healing-related proteins as the most represented functionalities exhibited by the deregulated proteins in CAFs.

**TGFβ activation of fibroblasts displays a protein profile similar to colon CAFs**

The activation of the TGF-β pathway plays a major role in fibroblast activation during wound healing, organ fibrosis and cancer as confirmed by GO analysis of our data (Supplementary Fig. S3C) (27). Consequently, we expected that the activation of normal fibroblasts with exogenous TGFβ would yield an expression pattern similar to that observed in CAFs. Therefore, BJh-TERT fibroblasts were treated with TGFβ for 24 h and candidate CAFs biomarkers were analyzed by quantitative PCR and western blot. Except for POSTN and ACAN, which did not get amplified after RT-PCR, most of the up-regulated proteins in CAFs (BGN, CNN3, CSPG4, DKK3, LOXL2, MYL9, PLOD2, TAGLN and TGM2) and down-regulated (COL14A1 and CSF1) were also differentially expressed in TGFβ-activated BJh-TERT fibroblasts by qPCR (Fig. 1C) and western blot (Fig. 1D). These findings confirmed that activated colon CAFs display an expression profile similar to TGFβ-activated fibroblasts.

**Stromal specificity of deregulated proteins in colon CAFs**

To analyze cellular specificity of altered CAF proteins, we used two strategies. First, we used the GSE39396 dataset corresponding to the expression profile of four sorters-isolated cell types (endothelial cells: CD31+, epithelial cells: EPCAM+, inflammatory cells: CD45+ and CAFs: FAP+) from six colorectal cancer patients (28). Data were
LOXL2 in colon cancer prognosis

transformed in z-score values and analyzed using color-coded heat map. Positive and negative z-scores indicated a significantly higher or lower expression level, respectively. 28 out of 60 up-regulated proteins (Fig. 2A) and 24 out of 40 down-regulated proteins (Fig. 2B) corresponded to genes highly-expressed in the fibroblast compartment. Second, we analyzed the gene expression for these proteins in KM12C, KM12SM, SW480 and SW620 epithelial colorectal cancer cells using QPCR. Ten proteins (ACAN, BGN, CDH13, CSPG4, DKK3, LOXL2, MYL9, PLOD2, TAGLN, and TGM2) were specific or overexpressed in fibroblasts (Fig. 2C). Together, these results confirm the value of these proteins as specific colon cancer stromal markers.

**Desmoplastic signature associated to CAFs biomarkers**

Activated fibroblasts contribute greatly to desmoplastic activity. Desmoplastic changes are associated to fibrosis and correlate with highly-expressed smooth muscle proteins. Therefore, we investigated for each protein whether or not smooth muscle expression was among the five top expression sites according to Bio-GPS and HPA. Many up-regulated (45 out of 60, 75%) and down-regulated (24 out of 40, 60%) CAFs proteins displayed preferential smooth muscle expression (Supplementary Table S5). Some representative up-regulated proteins were tested for stromal staining in colon cancer using immunohistochemical analysis (Fig. 3) or the HPA database (Supplementary Fig. S4). The HPA showed a preferential expression in cancer stroma of 36 up-regulated CAFs proteins. We found that ACAN, αSMA, CDH13, DKK3, TAGLN, TGM2 and LOXL2 were preferentially expressed in cancer stroma, with variable or no expression in epithelium or in normal colon mucosa stroma (Fig. 3). In agreement with expression data (Fig. 2A), LOXL2 also exhibited some expression in endothelial cells, as previously reported (29). ACAN and CDH13 were only present in the leading edge
LOXL2 in colon cancer prognosis

of the tumor. In contrast, COL14A1 was highly expressed in normal crypt epithelial cells, with very weak expression in normal stroma and no expression in tumoral tissues.

To study if desmoplastic changes found in the primary colon cancer lesion would take place also in the metastatic lesions, we performed an expression analysis in liver metastasis (Fig. 3). αSMA, TAGLN, TGM2 and LOXL2 were highly expressed at the metastatic site, whereas ACAN, CDH13 and DKK3 showed reduced levels respect to the primary tumors. This result suggests minor differences between the desmoplastic site at the primary tumor and the metastatic site. In summary, 32 out of the 36 tested proteins showed myofibroblast-like staining, with preferential stromal expression. These 32 proteins would constitute a desmoplastic signature for colon CAFs.

Prognostic value of the individual genes according to external datasets

We selected 15 confirmed stromal markers to identify a molecular signature in order to predict disease-specific survival and recurrence. First, we used the pooled cohort GSE17538 dataset (232 patients with colon cancer). Patients were divided into “low expression” with negative score and “high expression” with positive score. The individual prognostic value for each gene was analyzed using Kaplan-Meier survival curves and its significance was tested with log rank tests and univariate Cox proportional hazards models. In the training set consisting of 232 patients, five genes, LOXL2 (log-rank \( P < 0.0001 \)) HR: 2.63 (CI, 1.68-4.12), TAGLN (log-rank \( P = 0.025 \)) HR:1.61 (CI, 1.06-2.45), MYL9 (log-rank \( P = 0.022 \)) HR: 1.63 (1.07-2.49), POSTN (log-rank \( P=0.008 \)) HR:1.80 (CI, 1.16-2.81) and CSF1 (log-rank \( P=0.015 \)) HR:0.60 (CI, 0.39-0.91) showed the strongest prognostic relevance (Supplementary Fig. S5).

We carried out a similar analysis for the prediction of disease-free survival in the same cohort using 173 patients at stages I-III. The six genes showing better disease-free
LOXL2 in colon cancer prognosis

Prediction were LOXL2 log-rank $P = 0.0002$ HR: 4.18 (1.82-9.59), BGN log-rank $P < 0.0001$ HR: 5.13 (2.13-12.38), MYL9 log-rank $P = 0.002$ HR: 3.22 (1.46-7.10), POSTN log-rank $P = 0.0001$ HR: 5.97 (2.11-16.94), PLOD2 log-rank $P = 0.008$ HR: 2.81 (1.27-6.21) and TAGLN log-rank $P < 0.0001$ HR: 4.86 (2.02-11.73) (Supplementary Fig. S6). From these results, we selected LOXL2 and TAGLN for further validation, as they were consistently reliable in overall and disease-free survival prognosis.

Validation of LOXL2 and TAGLN prognostic value in a different patient cohort

The prognostic value of LOXL2 and TAGLN was initially validated in a different cohort of 70 human colon cancer samples followed for more than five years. We did not include rectal adenocarcinomas as they are managed with neoadjuvant therapy prior to surgery. We purified mRNA from the 70 paired FFPE tissues for real-time quantitative PCR. We observed an increased expression for LOXL2 in 44% of the tumoral samples respect to normal tissue. High expression of LOXL2 was associated with overall survival $P = 0.001$ HR: 8.52 (1.90-38.29) (Fig. 4A) and higher recurrence $P = 0.001$ HR: 5.38 (1.70-17.01) (Fig. 4B). Whereas high expression of TAGLN correlated with overall survival log-rank $P = 0.004$ HR: 4.10 (1.45-11.62), it did not show association with disease-free survival by QPCR (Fig. 4A, B). After adjustment for age, sex, stage and grade, multivariate analysis confirmed that LOXL2 expression was associated with poor prognosis and an independent prognostic factor for disease-free and overall survival using GSE17538 ($P = 0.0037$, $P < 0.0001$, respectively) and internal mRNA data set ($P = 0.0006$, $P = 0.0035$, respectively) (Table 1). We did not find association between LOXL2 expression and other clinicopathological characteristics (Supplementary Table S6). Subsequently, we used another sample cohort of 121 patients for immunohistochemistry analysis. About 32% of the patients exhibited strong LOXL2 expression (>50% of the sample stained positively) and 23% showed moderate
LOXL2 in colon cancer prognosis

eexpression (20-50%) (Supplementary Table S1C). LOXL2 staining was mainly nuclear in the stroma (Fig. 3 and Supplementary Fig S7). Stromal (nuclear) positive staining correlated with overall survival \( P = 0.033 \) HR: 4.19 (1.22-14.41) and disease-free survival \( P = 0.015 \) HR: 2.20 (1.14-4.25). Therefore, after testing and validation using different cohorts, the expression of LOXL2 showed a significant association to poor clinical outcome and higher recurrence in human colon cancer. Combinations of LOXL2 with other markers did not improve its prognostic power.

**Predictive prognostic value of LOXL2 for stage II and stage III patients**

To explore the value of LOXL2 for risk assessment in stage II and III patients, we analyzed recurrence of stage II patients from combined datasets GSE17538 (70 patients at stage II) and GSE33113 (89 patients at stage II, with disease-free information) or survival of patients with stage III from combined GSE17538 (76 patients at stage III) and GSE12945 (21 patients at stage III). The prognostic value was validated by QPCR and immunohistochemistry. Stage II patients showed shorter time to recurrence when LOXL2 was highly expressed \( P = 0.011 \) HR: 2.74 (1.22-6.20) (Fig. 5A). The 3-year DFS was 73.9% (64.4-84.9%) for LOXL2-high versus 89.6% (82.6-97.2%) for LOXL2-low patients. In line with these results, high LOXL2 expression was associated with shorter disease-free using QPCR in 47 patients at stages II and III \( P = 0.001 \) HR: 5.69 (1.79-18.07). Using the immunohistochemistry values of the 55 patients at stage II from the previous cohort (Supplementary Table S1C), the disease-free survival was \( P = 0.005 \) HR: 4.35 (1.43-13.25) (Fig. 5A). Regarding overall survival for patients with stage III, we found association between high expression of LOXL2 and overall survival \( P = 0.004 \) HR: 3.40 (1.40-8.24) (Fig. 5B), confirmed by QPCR in patients at stage II and III \( P = 0.001 \) (Fig. 5B). In summary, these results suggest that LOXL2 is a suitable marker for poor prognosis, recurrence and patient stratification at stage II. This
LOXL2 in colon cancer prognosis

promising value for risk-stratification will require further validation in a larger independent cohort.
DISCUSSION

We used primary cultures of purified fibroblast populations isolated from colon cancer patients to identify stromally-expressed LOXL2 as a protein displaying a significant prognostic value. We demonstrated the prognostic power of LOXL2 using combined external colon cancer datasets for biomarker training and, then, validation by Q-PCR and immunohistochemistry on different cohorts of patients. Since twenty-five percent of stage II colon cancer patients will have recurrence of disease within 5 years (8), we were interested in testing the prognostic value in this group of patients. Using public databases, LOXL2 identified a high-risk group within stage II patients that could benefit from adjuvant chemotherapy. This prognostic power was confirmed by immunohistochemistry. Therefore, LOXL2 could complement clinical guidelines for high-risk definition (number of resected specimens, tumor differentiation, vascular invasion, etc) and correct patient classification. A similar prognostic ability of LOXL2 was found for stage III patients. Although these patients are usually given adjuvant treatment, some groups (elderly and frail patients) do not always receive treatment (30). In this case, LOXL2 might help in clinical decisions for treating patients older than 75 years.

In addition, we identified a desmoplastic signature for colon cancer that contained 32 CAFs-derived proteins showing cancer stromal staining. Proteins contributing to desmoplastic lesions are characteristic of tumor type. We observed a limited coincidence with mouse fibroblasts coming from a chemically-induced colon cancer (19), among other coincidences we found: FN1, TPM2, TNC, COL1A2, IGFBP7, FSTL1, THBS1, COL4A2, BGN and CSF1. We also observed CDH13, a mesenchymal cadherin similar to CDH11, previously reported in mouse cancer fibroblasts (19). However, affected biological functions were equivalent between mice
and human: TGFβ activation, ECM, cell adhesion and wound-healing. In any case, the relatively low level of coincidence with mouse fibroblasts could be explained because the mouse model does not recapitulate metastasis.

A protein profile similar to CAFs was obtained after activation of NFs with TGFβ. Remarkably, many stromal proteins (BGN, TAGLN, TGM2, LOXL2) exhibit a strong prognostic impact alone or in combination with other markers and deserve further investigation. TAGLN contributes to the cross-linking and polymerization of actin (31) and phenotype differentiation (32), showing a good predictive power that would require larger cohorts for validation. BGN and TGM2 are associated to the extracellular matrix.

Recently, LOXL2 has been shown to be critical in metastatic niche formation in hepatocellular carcinoma (33) and a marker of poor prognosis in breast, gastric cancer and squamous cell carcinomas (34-36), mostly associated to metastasis. In colon cancer, LOXL2 was associated to less differentiated tumors (37), but no effect on prognosis was reported. Here, LOXL2 alone showed to be an excellent molecular classifier by two different approaches, QPCR and immunohistochemistry analysis. Prognostic value was associated to the nuclear staining of stromal cells, although some increase in cytoplasmic expression was observed in epithelial cells in the metastasis. Therefore, the nuclear activity of LOXL2 seems essential for the association with poor prognosis. However, it is remarkable that LOXL2 expression in complete tumors leads to prognostic value, even if the differential expression was identified in fibroblasts of the stromal compartment. This finding should facilitate the incorporation of LOXL2 expression analysis to clinical routine. The identification of potential molecular targets that define prognosis can become even more important if these molecules can be inhibited with some of the drugs in development. For LOXL2, a specific inhibitory
LOXL2 in colon cancer prognosis

monoclonal antibody has been developed (38), which is currently in clinical trials for a number of fibrotic conditions and cancers.

LOXL2 is a member of the lysyl oxidase gene family that catalyzes the crosslinking of collagens and elastin in the extracellular matrix. This provokes an increase of ECM stiffness and leads to increased activation of the PI3K pathway (39). The cross-linking facilitates the recruitment of bone marrow-derived cells that release cytokines and growth factors, which facilitate the cancer cell colonization in distant organs (40). Moreover, LOXL2 interacts with Snail1 transcriptional repressor and represses E-cadherin and cell-polarity genes by Snail-dependent and independent mechanism (34, 41). In addition, a nuclear role for LOXL2 has been described (42). LOXL2 works as a transcriptional co-repressor through its activity in histone H3 deamination (43). Furthermore, LOXL2 could oxidize H3 to remove the trimethylated amino group located in the ε-position of Lysine 4 (44), which is usually related with transcriptional control (45). The association of preferential nuclear staining in the stroma with poor survival could be related with the role of LOXL2 as transcriptional regulator in stromal cells.

In summary, our results support the relevance of CAFs-derived proteins, such as BGN, TAGLN or LOXL2, in colon cancer desmoplastic analysis, prognosis and survival. We identified a large number of colon cancer stroma-specific proteins and LOXL2 as a single and highly predictive prognostic factor for overall survival and relapse in colon cancer patients using different cohorts for training and validation. Currently, lymph node involvement is admitted as the main prognostic factor and determines the administration of adjuvant systemic therapy. However, in stage II other tumour characteristics influence the indication of therapy. In this context, a biomarker
LOXL2 in colon cancer prognosis

as LOXL2 could be useful in the decision to treat patients more aggressively, especially the subgroup of stage IIA patients with LOXL2 overexpressed in tumour tissue.
LOXL2 in colon cancer prognosis

GRANT SUPPORT

S. Torres was a recipient of a Juan de la Cierva programme. R.A. Bartolomé was supported by a grant to established research groups of the Asociación Española Contra el Cáncer (AECC). María Lopez-Lucendo was a recipient of a ProteoRed contract. A. Peláez was a FPI fellow from the MINECO. This research was supported by grants to established research groups of the “Asociación Española Contra el Cancer (AECC)”, BIO2012-31023 from the Spanish Ministry of Economy and Competitiveness, S2010/BMD-2344/ Colomics2 from the Comunidad de Madrid and the Grant PRB2 (IPT13/0001 - ISCIII-SGEFI / FEDER).
LOXL2 in colon cancer prognosis

REFERENCES

LOXL2 in colon cancer prognosis


LOXL2 in colon cancer prognosis


### Table 1: Univariate and multivariable analysis for LOXL2 in disease-free and overall survival in colon cancer patients

<table>
<thead>
<tr>
<th></th>
<th>Disease-free survival</th>
<th>Overall survival</th>
<th>Disease-free survival</th>
<th>Overall survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P-value</td>
<td>n (events)</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td><strong>Univariate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOXL2 negative</td>
<td>1</td>
<td>0.0002</td>
<td>86(7)</td>
<td>1</td>
</tr>
<tr>
<td>LOXL2 positive</td>
<td>4.18 (1.82-9.59)</td>
<td>0.0002</td>
<td>87(28)</td>
<td>2.63 (1.68-4.12)</td>
</tr>
<tr>
<td><strong>Multivariable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOXL2 positive</td>
<td>3.49 (1.50-8.09)</td>
<td>0.0037</td>
<td></td>
<td>3.19 (2.00-5.07)</td>
</tr>
<tr>
<td>Sex</td>
<td>0.71 (0.34-1.48)</td>
<td>0.3656</td>
<td></td>
<td>1.09 (0.69-1.71)</td>
</tr>
<tr>
<td>Age</td>
<td>0.99 (0.97-1.07)</td>
<td>0.4951</td>
<td></td>
<td>1.02 (1.00-1.04)</td>
</tr>
<tr>
<td><strong>Stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III vs I/II</td>
<td>2.41 (1.18-4.93)</td>
<td>0.0162</td>
<td></td>
<td>1.78 (0.99-3.21)</td>
</tr>
<tr>
<td>IV vs I/II</td>
<td>9.21 (5.23-16.2)</td>
<td>&lt;0.0001</td>
<td></td>
<td>9.21 (5.23-16.2)</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mod. dif. vs well dif</td>
<td>0.96 (0.29-3.21)</td>
<td>0.9497</td>
<td>0.94 (0.37-2.37)</td>
<td>0.8938</td>
</tr>
<tr>
<td>poorly. dif. vs well dif</td>
<td>1.31 (0.32-5.40)</td>
<td>0.7059</td>
<td>1.46 (0.53-4.08)</td>
<td>0.4670</td>
</tr>
<tr>
<td>undiff. vs well dif</td>
<td>1.48 (0.48-5.42)</td>
<td>0.4908</td>
<td></td>
<td>1.48 (0.48-5.42)</td>
</tr>
</tbody>
</table>

Multivariable analysis adjusted for age, gender, AJCC stage (I, II, III, IV) and grade of differentiation (well differentiated, moderately differentiated, poorly differentiated and undifferentiated). Abbreviations: HR = Hazard ratio; CI = confidence interval.
Legend to Figures

Figure 1. Validation of CAFs deregulated proteins. (A) Whole cellular extracts or concentrated protein samples from conditioned medium of CAFs and NFs were separated on SDS-PAGE gels, transferred to nitrocellulose membranes and probed with the indicated antibodies. Tubulin was used as a loading control. Protein abundance was quantified by densitometry to compare the expression with the iTRAQ ratios. (B) cDNAs synthesized from total RNA obtained from NFs and CAFs were subjected to QRT-PCR using specific primers for the selected genes and 18S rRNA for normalization. BJh-TERT cells were starved in serum-free medium, treated with TGFβ (5 ng/mL) for 24 h and lysed. The extracts were analyzed by QPCR (C) and Western blot (D) using specific primers and the indicated antibodies. Data represent the mean ± SD of three experiments. (*, p < 0.05, **, p < 0.01; ***, p < 0.001).

Figure 2. Stromal specificity of CAFs proteins. Differentially-expressed genes in CAFs vs NFs (60 up-regulated (A) and 40 down-regulated (B)) are represented by heat-maps of scaled gene-expression levels according to a dataset for: epithelial cells: EPCAM+, leukocytes: CD45+, fibroblasts: FAP+ and endothelial cells: CD31+ (28). Color scale represents the value of the z-score calculated in the dataset. Red indicates a higher z-score value, whereas blue indicates a lower value. (C) cDNAs synthesized from total RNA from NFs, CAFs and epithelial colon cancer cell lines (KM12SM, KM12C, SW480 and SW620) were subjected to QPCR as in Figure 1. Data represent the mean ± SD of three experiments.

Figure 3. Expression of stromal biomarkers in paired tissues of colon cancer patients. Immunohistochemical staining for ACAN, αSMA, COL14A1, CDH13,
LOXL2 in colon cancer prognosis

DKK3, TAGLN, TGM2 and LOXL2 in normal mucosa, primary colon cancer and liver metastasis. Pictures were taken at x200 magnification.

**Figure 4. Prognostic value of LOXL2 and TAGLN in colon cancer.** Biomarker validation was carried out in two steps: 1. By QPCR in a different cohort of 70 colon cancer patients: (A) overall survival and (B) disease-free. 2. By immunohistochemistry analysis using 121 samples in a tissue microarray slide. Positive expression was associated to nuclear staining of the stromal fibroblasts. (C) Overall survival and (D) disease-free survival. Kaplan-Meier curves were performed using the log-rank test. The Hazard Ratio was based on the Cox model.

**Figure 5. Risk stratification of colon cancer patients at stages II and III using LOXL2 expression.** (A) Tumor recurrence-free survival was determined using: GSE17538+GSE31331 datasets of colon patients at stage II, QPCR with a different cohort of 48 patients at stage II-III or immunohistochemistry (IHC) with 55 samples from patients at stage II. (B) Overall survival was determined using GSE17538+GSE12945 datasets of patients at stage III or QPCR with samples from 48 patients at stage II-III. Kaplan-Meier curves were performed using the log-rank test. The Hazard Ratio was based on the Cox model.
Figure 1

(A) Immunoblot analysis of protein expression in fibroblasts (NFs) and cancer-associated fibroblasts (CAFs) with varying molecular weights (KDa).

(B) mRNA expression levels of TGFβ1 in Conditioned medium and Cellular extract. * indicates p < 0.05, ** indicates p < 0.01, *** indicates p < 0.001.

(C) TGFβ1 mRNA expression levels in Individual samples. *** indicates p < 0.001.

(D) Cellular extract and Conditioned medium protein analysis showing the expression of various proteins. ** indicates p < 0.01, *** indicates p < 0.001.
Figure 2

(A) Heatmap showing the expression levels of various genes in different cancer cell lines.

(B) Heatmap showing the down- and up-regulated proteins in different cancer cell lines.

(C) Bar graph showing the mRNA expression levels of different genes in various cancer cell lines (a.u.).
Figure 3
**Figure 4**

**mRNA from paired FFPE tissues**

**LOXL2**

- **A**
  - Overall survival
  - Time (months)
  - P = 0.001
  - HR = 8.52 (1.90-38.29)
  - n = 39

- **B**
  - Disease free survival
  - Time (months)
  - P = 0.001
  - HR = 5.38 (1.70-17.01)
  - n = 31

**TAGLN**

- **C**
  - Overall survival
  - Time (months)
  - P = 0.004
  - HR = 4.10 (1.45-11.62)
  - n = 20

- **D**
  - Disease free survival
  - Time (months)
  - P = 0.428
  - HR = 1.57 (0.53-4.58)
  - n = 14

**LOXL2**

- **C**
  - Overall survival
  - Time (months)
  - P = 0.033
  - HR = 4.19 (1.22-14.41)
  - n = 54

- **D**
  - Disease free survival
  - Time (months)
  - P = 0.015
  - HR = 2.20 (1.14-4.25)
  - n = 45

**IHC analysis**

- **B**
  - Negative
  - Positive
  - Time (months)
  - n = 33

- **C**
  - Negative
  - Positive
  - Time (months)
  - n = 50

- **D**
  - Negative
  - Positive
  - Time (months)
  - n = 67

---

Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Research.

Downloaded from clincancerres.aacrjournals.org on July 16, 2017. © 2015 American Association for Cancer Research.
Figure 5

**A**

**GSE33113/GSE17538**

*Stage II*

- **Low**
  - Disease free survival: \(89.6 (82.6-97.2)\)%
  - Time (months): 120
  - \(n = 77\)
- **High**
  - Disease free survival: \(73.9 (64.4-84.9)\)%
  - Time (months): 120
  - \(n = 82\)

\(P = 0.011\)

\(HR = 2.74 (1.22-6.20)\)

**mRNA from paired FFPE tissues**

*Stage II + III*

- **Negative**
  - Disease free survival: \(89.6 (82.6-97.2)\)%
  - Time (months): 120
  - \(n = 28\)
- **Positive**
  - Disease free survival: \(73.9 (64.4-84.9)\)%
  - Time (months): 120
  - \(n = 19\)

\(P = 0.001\)

\(HR = 5.69 (1.79-18.07)\)

**IHC analysis**

*Stage II*

- **Negative**
  - Disease free survival: \(89.6 (82.6-97.2)\)%
  - Time (months): 120
  - \(n = 28\)
- **Positive**
  - Disease free survival: \(73.9 (64.4-84.9)\)%
  - Time (months): 120
  - \(n = 27\)

\(P = 0.005\)

\(HR = 4.35 (1.43-13.25)\)

**B**

**GSE12945/GSE17538**

*Stage III*

- **Low**
  - Disease free survival: \(81.7 (69.1-96.4)\)%
  - Time (months): 120
  - \(n = 40\)
- **High**
  - Disease free survival: \(51.6 (38.9-68.4)\)%
  - Time (months): 120
  - \(n = 57\)

\(P = 0.004\)

\(HR = 3.40 (1.40-8.24)\)

**mRNA from paired FFPE tissues**

*Stage II + III*

- **Negative**
  - Disease free survival: \(81.7 (69.1-96.4)\)%
  - Time (months): 120
  - \(n = 28\)
- **Positive**
  - Disease free survival: \(51.6 (38.9-68.4)\)%
  - Time (months): 120
  - \(n = 20\)

\(P = 0.001\)

\(HR = 3.40 (1.40-8.24)\)
Clinical Cancer Research

LOXL2 IS HIGHLY EXPRESSED IN CANCER-ASSOCIATED FIBROBLASTS AND ASSOCIATES TO POOR COLON CANCER SURVIVAL

Sofia Torres, Irene Garcia-Palmero, Mercedes Herrera, et al.

*Clin Cancer Res* Published OnlineFirst July 23, 2015.

**Updated version**
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-14-3096

**Supplementary Material**
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2015/07/27/1078-0432.CCR-14-3096.DC1

**Author Manuscript**
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

**E-mail alerts**
Sign up to receive free email-alerts related to this article or journal.

**Reprints and Subscriptions**
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

**Permissions**
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.