MOLECULAR ONCOLOGY 8 (2014) 1290–1305

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/molonc

Differences between CAFs and their paired NCF from adjacent colonic mucosa reveal functional heterogeneity of CAFs, providing prognostic information

Mireia Berdiel-Acer[a], *Rebeca Sanz-Pamplona*[b], *Alexandre Calon*[c], *Daniel Cuadras*[b], *Antoni Berenguer*[b], *Xavier Sanjuan*[d], *Maria José Paules*[d], *Ramon Salazar*[e], *Victor Moreno*[b], *Eduard Batlle*[c], *Alberto Villanueva*[a], *David G. Molleví*[a,*]

[a]*Translational Research Laboratory, Catalan Institute of Oncology, IDIBELL, L'Hospitalet de Llobregat, 08908 Catalonia, Spain*
[b]*Biomarkers and Susceptibility Unit, Cancer Prevention and Monitoring Programme, Catalan Institute of Oncology, IDIBELL, L'Hospitalet de Llobregat, Catalonia, Spain*
[c]*Oncology Programme, Institute for Research in Biomedicine IRB, Barcelona, Catalonia, Spain*
[d]*Pathology Department, Hospital Universitari de Bellvitge, IDIBELL, L'Hospitalet de Llobregat, Catalonia, Spain*
[e]*Medical Oncology Department, Catalan Institute of Oncology, IDIBELL, L'Hospitalet de Llobregat, Catalonia, Spain*

ARTICLE INFO

Article history:
Received 10 February 2014
Received in revised form
14 April 2014
Accepted 15 April 2014
Available online 6 May 2014

Keywords:
Carcinoma-associated fibroblasts
Microenvironment
Stroma
Biomarker
Colorectal

ABSTRACT

Little is known about the difference in gene expression between carcinoma-associated fibroblasts (CAFs) and paired normal colonic fibroblasts (NCFs) in colorectal cancer. Paired CAFs and NCFs were isolated from eight primary human colorectal carcinoma specimens. In culture conditions, soluble factors secreted by CAFs in the conditioned media increased clonogenicity and migration of epithelial cancer cells lines to a greater extent than did NCF. *In vivo*, CAFs were more competent as tumour growth enhancers than paired NCFs when co-inoculated with colorectal cell lines. Gene expression analysis of microarrays of CAF and paired NCF populations enabled us to identify 108 deregulated genes (38 upregulated and 70 downregulated genes). Most of those genes are fibroblast-specific. This has been validated *in silico* in dataset GSE39396 and by qPCR in selected genes. GSEA analysis revealed a differential transcriptomic profile of CAFs, mainly involving the Wnt signallingsignalling pathway, focal adhesion and cell cycle. Both deregulated genes and biological processes involved depicted a considerable degree of overlap with deregulated genes reported in breast, lung, oesophagus and prostate CAFs. These observations suggest that similar transcriptomic programs may be active in the transition from normal fibroblast in adjacent tissues to CAFs, independently of their anatomic demarcation. Additionally NCF already depicted an activated pattern associated with inflammation.

The deregulated genes signature score seemed to correlate with CAF tumour promoter abilities *in vitro*, suggesting a high degree of heterogeneity between CAFs, and it has also prognostic value in two independent datasets.

Abbreviations: DEG, differentially expressed gene; CAF, carcinoma-associated fibroblast; NCF, normal colonic fibroblasts; ECM, extracellular matrix; CM, conditioned medium.

* *Corresponding author*. Tel.: +34 93 2607370; fax: +34 93 2607466.
E-mail addresses: dgmollevi@iconcologia.net, davidg.mollevi@gmail.com (D.G. Molleví).

http://dx.doi.org/10.1016/j.molonc.2014.04.006

Further characterization of the roles these biomarkers play in cancer will reveal how CAFs provide cancer cells with a suitable microenvironment and may help in the development of new therapeutic targets for cancer treatment.

1. Introduction

The tumour microenvironment is a place where tumour cells co-exist and co-evolve with "non-tumoral" cells such as adipocytes, fibroblasts, endothelial and immune cells, all of which are embedded in an extracellular matrix (ECM). When changes occur in this compartment, the altered stroma can influence cancer development and progression (Bissell et al., 1982; Bissell and Hines, 2011). Carcinoma-associated fibroblasts (CAFs), the main constituents of tumour stroma, actively drive tumourigenesis and cancer progression (Elenbaas and Weinberg, 2001; Hwang et al., 2008; Olumi et al., 1999; Orimo et al., 2005; Zhang et al., 2009). They participate throughout tumour development by establishing cell–cell interactions with tumour cells or through secretion of cytokines, chemokines and growth factors. CAFs are most often denoted by the expression of α-smooth muscle actin (αSMA), but other markers such as vimentin and fibroblast-activating protein (FAPα) are also used to identify them. Nevertheless, these are also markers of myofibroblasts or activated fibroblasts, and not properly and exclusively of CAFs. An exclusive marker for CAFs that can clearly distinguish them from normal fibroblasts from adjacent mucosa or other closely related cell types is yet to be identified. Chang et al. (2002) reported that transcriptional patterns displayed by fibroblasts from different anatomic sites were distinct and characteristic, and suggested that fibroblasts from different organs could be considered as distinct differentiated cell types. However, in a subsequent study, the same group defined a common transcriptomic profile in fibroblasts stimulated with serum and with a pattern similar to that observed in a wound healing process and associated with tumour progression (Chang et al., 2004). It is known that the involvement of CAFs in solid tumours, in which desmoplasia is a characteristic feature, is of huge importance. Gene expression profiles and their association with cancer aggressiveness, has been reported in several tumours types (Peng et al., 2013). The current work is focused on colorectal cancer, in which, to our knowledge, the genetic and molecular profile of CAFs has been less thoroughly explored. The purpose of this study is to identify, for the first time in paired samples, genes that are differentially expressed in two fibroblast populations derived from the same colorectal cancer (CRC) patients (NCFs and their paired CAFs from primary tumour) and to examine their associations with phenotypic differences. Using microarray technology, we identified 109 deregulated probes (corresponding to 108 genes) differentially expressed between the two populations and functionally involved in cancer progression. Knowledge of alterations in the stroma surrounding a tumour might provide tools and biomarkers of use in designing new ways to attack tumour-supportive CAFs and could add valuable information for future and prognostic treatments.

2. Materials and methods

2.1. *Culture of primary fibroblasts and preparation of conditioned medium*

Fresh surgical specimens from colorectal cancer patients were obtained with the approval of the Ethics Committee of the Hospital Universitari de Bellvitge (IDIBELL). Tissue samples from morphologically normal colonic mucosa (at least 5 cm from the surgical margin), and from colorectal primary tumour were minced and incubated with collagenase and dispase for 2 h at 37 °C. Cells were resuspended and plated with Dulbecco's modified Eagle's medium-F12 (DMEM F12, Gibco) containing 10% fetal bovine serum (FBS) and penicillin/streptomycin antibiotics. Primary fibroblast cultures were established and routinely maintained at 37 °C in a humidified atmosphere containing 5% CO_2. After a maximum of 5 passages, RNA and protein were obtained to check for fibroblast purity. 10^6 fibroblasts were incubated for 48 h in 10-cm diameter dishes in DMEM-F12 with or without FBS. Conditioned medium was collected, centrifuged for 5 min at 3000 rpm to remove cell debris, sterile-filtered through 0.22-μm filter units (Millex® GS, Millipore) and stored at −80 °C until use.

2.2. In vitro *cellular assays*

Migration of cancer cells and CAFs was measured by wound healing assay. Cells were seeded in 6-cm diameter plates and cultured until confluent. The cell monolayer was scratched with a yellow 200-μl pipette tip to create a wound. After several PBS (phosphate-buffered saline 1×) washes to remove floating cells, in an epithelial cell migration assay, conditioned medium from NCF or CAF was added. Pictures were taken at different times. Distances between cell margins were measured with Leica software (Wetzal, Germany) on three occasions and each assay was performed in duplicate. Clonogenic capacity was assessed by cloning assay. We plated 100 cells for each epithelial colon cancer cell line (DLD1, SW620, SW480 and SW1116) in 12-well plates and incubated them for 9 days in DMEM F12 10% (control) or the appropriate conditioned medium. The number of colonies was counted after crystal violet staining. A WST-1 cell proliferation assay was conducted in CAFs alone and in DLD1 cells stimulated with CAF conditioned medium (CM) (24 h without FBS being collected, as mentioned above). Briefly, 1000 cells were seeded in a 96-well plate and cultured at several times (0, 24, 72 and

144 h), taking time 0 as the first measurement once cells were attached. After performing the time assay, the culture medium was removed and replaced with 10 μl of WST-1 reagent (Cell Proliferation Reagent WST-1, Roche) in 100 μl serum-free DMEM/F12. Absorbance at 450 nm was measured after 2 h incubation (37 °C, 5% CO_2, in darkness).

2.3. Western blot analysis

To extract total protein, monocultured fibroblasts were homogenized with RIPA lysis buffer (PBS 1×, 1% SDS, 1% nonidet-40, 0.5% sodium deoxycholate), supplemented with complete EDTA-free Protease Inhibitor Cocktail Tablets (Roche), orthovanadate, PMSF, β-glycerol and leupeptin. Lysates were cleared by centrifugation and protein samples were loaded onto SDS/polyacrylamide gels and transferred to polyvinylidene fluoride membranes. Membranes were blocked for 1 h at room temperature in 10% nonfat milk and 0.1% Tween in TBS 1×, washed, and incubated overnight at 4 °C with the corresponding dilution of primary antibodies. A second incubation was performed using ECL™ horseradish peroxidase-linked secondary murine/rabbit antibody (GE Healthcare), and enhanced chemiluminescence was detected by Novex® ECL chemiluminescent substrate reagent kit (Invitrogen).

2.4. Antibodies and reagents

Epithelial colon cancer cell lines DLD1, SW620, SW480, SW1116, HCT116, HCT-15, CaCO2, LoVo, Colo205, RKO, KM12C, HT-29, Co115 were purchased from the American Type Culture Collection (ATCC). All cell lines were maintained in DMEM-F12 10% FBS with added antibiotics. Primary antibodies used in western blot were: pre-diluted anti-α smooth muscle actin (Abcam) at 1/3 dilution, anti-vimentin (Invitrogen) at 1/1000, anti-vinculin (Invitrogen) at 1/400, anti-E-cadherin and N-cadherin (BD Biosciences) at 1/500, VE-cadherin (Abcam) and Tubulin (Sigma) at 1/3000 and 1/500 dilutions respectively.

2.5. In vivo tumorogenicity assay

Hsd: Athymic Nude-Foxn1^{nu} mice were co-injected subcutaneously in the right and left lateral flanks, with 1.2×10^6 DLD1 colon cancer cells alone ($n = 4$ injections) or in combination with 5×10^5 NCF ($n = 5$ different NCFs, $n = 10$ coinjections) or CAF ($n = 5$ different CAFs, $n = 10$ coinjections). Each experiment was repeated twice. Only two pairs NCF/CAF from the same patient were tested. After only 21 days, the volume of the generated tumours using CAFs was >2 cm^3 in some cases and the affected mice had to be sacrificed. Tumour volume and weight were determined. Additionally, we evaluated Ki67 to assess the proliferative status of tumour cells (Dako antibody clone MIB-1) and human specific Vimentin immunohistochemical staining (Dako antibody clone V9), for evidencing the presence of the coinjected human fibroblasts.

2.6. RNA isolation and quantitative real-time PCR (qRT-PCR)

Total RNA from different isolated fibroblasts and colorectal cancer cell lines was extracted using TRIzol® reagent method and column purification using PureLink™ RNA Mini Kit (Invitrogen). RNA quantity was determined by NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Rockland, DE) and 100 ng of total RNA was reverse-transcribed using M-MLV reverse transcriptase (Invitrogen) following the manufacturer's instructions. A 0.1-μg equivalent of the corresponding cDNA was used for each quantitative PCR assay performed with the LightCycler® 480, SYBR Green I Master (Roche Applied Science, Mannheim, Germany). Primers were designed using Primer3 Input (http://primer3.wi.mit.edu) and predicted PCR product sequences were verified using BLAST (http://www.ncbi.nlm.nih.gov/blast). All primer sequences can be provided upon request.

2.7. Microarray analysis

RNA from 8 paired NCF/CAF cultures was extracted as described above. Total mRNA was hybridized in an Affymetrix GeneChip Human Gene 1.0 ST Array. All computations and statistical analyses were performed using the R language and environment (R Development Core Team, 2008). The microarray data were read, normalized, and transformed to numerical expression values using the *justRMA* function in the *simpleaffy* package (Wilson and Miller, 2005). The resulting data were used to look for genes that were differentially expressed between groups (NCF *vs.* CAF) using the *Significance Analysis of Microarrays* (SAM) technique, available in the *samr* package (Tusher et al., 2001). To obtain a reduced list of genes, we considered those with a false discovery rate (FDR) < 0.05 and a >2-fold change.

2.8. Gene set enrichment analysis and gene ontology

Gene set enrichment analysis (GSEA) (Subramanian et al., 2005) was applied to the pre-ranked list of differentially expressed genes (by value of SAM *d* statistic). We wanted to determine whether there was enrichment (ES) in our deregulated genes list for particular pathways described in previously known gene sets. This bioinformatics tool analyzes the complete list of deregulated genes and allows small but coordinated changes in expression to be taken into consideration. The statistical significance of the ES was estimated from 1000 gene permutations. We used gene sets C2.CP.KEGG.v4.0 and C2.CP.ALL.v4.0.

Functional gene ontology (GO) annotation of genes of interest (FDR *q*-value < 0.1) was performed using the DAVID database. Gene functional classification and functional annotation clustering were performed to identify functional gene groups and ontology terms that were significantly overrepresented among the genes of interest.

2.9. Cocultures of NCF with colorectal cancer cell lines

The same NCF isolated to obtain transcriptomic profiles were used to establish five days transwell cocultures with colorectal cancer cell lines (DLD1 cells). Total mRNA was hybridized in an Affymetrix GeneChip Human Gene 1.0 ST Array. All procedures were performed as described in 2.7.

2.10. Signature score

The signature score for each of the fibroblasts included in the array was obtained by computing the standardized values of all 109 probes (108 genes) comprising the signature. Score values were plotted for each NCF/CAF pair.

The relevance of the signature score in an *in vitro* model was assessed by evaluating phenotypic differences in terms of proliferation in the two isolated fibroblasts with the greatest difference in score (according to microarray expression data). High and low score CAFs were selected for various functional assays for the purpose of elucidating phenotypic differences.

2.11. Validation of proinflammatory genes in GSE44076

A set of 98 paired adjacent-normal and tumour tissues from CRC patients and 50 colon mucosa from healthy donors (246 samples in total) were included in this dataset. Patients were selected to form a homogeneous clinical group of stage II, microsatellite stable (MSS) colorectal tumours. All had been treated with radical surgery, had not received adjuvant therapy and had a minimum follow up of three years. Adjacent normal tissue from patients was dissected from the resection margins, with a minimum distance of 10 cm from the tumour. Healthy donors were invited to participate in this study when they received a colonoscopy indicated for screening or symptoms, but with a result of no lesions in the colon. All samples were selected from Bellvitge University Hospital-IDIBELL (Spain) and the protocol was approved by the Ethics Committee. Written and informed consent from patients and healthy donors was required for inclusion in this study. RNA extracted from each sample was hybridized in Affymetrix chips Human Genome U219. After a quality control assessment following Affymetrix standards, data was normalized using the RMA algorithm. Both raw and normalized data are available in the Gene Expression Omnibus (GEO) database with accession number GSE44076.

Additionally, we used 10 healthy colonic mucosas, 10 normal adjacent mucosas and 10 paired tumours to validate by means of quantitative RT-PCR some of the differentially expressed genes between NCF and CAFs.

2.12. Prognostic analysis

Available annotated clinical data for GSE17537 (Smith et al., 2010) and GSE14333 (Jorissen et al., 2009) datasets included AJCC staging, age, gender and disease-free survival intervals. We analysed time to recurrence in stage I–III patients using the pooled GSE17537-GSE14333 dataset. To remove systematic biases between datasets, expression levels for all genes were transformed to z-scores before pooling. We reinforced the prognostic value in an independent dataset of stage II patients (GSE33113) (de Sousa et al., 2011). These datasets were chosen since no epithelial cell enrichment was performed on the sample (whole tumour samples).

To assess the predictive power of the gene signatures on recurrence we computed the mean signature expression of overexpressed genes and tested its significance with a univariate Cox proportional hazards model likelihood ratio test. The mean signature expression was computed from z-scores obtained for each of the genes and from summation z-scores across all genes in the signature.

We derived Kaplan–Meier survival curves for patients, dividing the cohort into two risk groups, using the mean as the cut-off (zero). We obtained a smooth estimate of the relationship between mean signature expression and recurrence hazard ratio (smoothCoxph function in the phenoTest package). Low and high signature expression (low and high risk, respectively) were defined as values below and above the mean, respectively. We also segregated the cohort into three risk groups (low, medium and high risk) with respect to the smooth estimate of the recurrence hazard ratio *vs.* the gene signature score. Statistical significance of the average signature score, introduced as a continuous covariate in the Cox model, was determined. Statistically significant variables from the univariate model were included in a multivariate Cox model, which also included age, gender, AJCC stage and study/hospital of data collection as adjustment covariates. Statistical significance was concluded for values of $p < 0.05$.

Survival analyses were done in the Biostatistics and Bioinformatics Unit of the Institute for Research in Biomedicine, Barcelona, Spain.

Finally we validate by means of quantitative RT-PCR some of the differentially expressed genes between NCF and CAFs in 60 samples of stage II/III colorectal cancer patients from our institution (30 non-recurrent and 30 recurrent).

3. Theory

Given the relevance that the tumour stroma is taking from the data provided by recent molecular classifications, the DEG between normal colonic fibroblasts obtained from adjacent mucosa and carcinoma-associated fibroblasts might be a tool to develop a gene classifier with a limited number of genes in order to predict outcome in colorectal cancer patients. The identification of a signature with prognostic value is a filter to remove those genes least likely to provide useful information to categorize patients according to their risk of recurrence. Our group is currently working to develop a 5-gene classifier based on a more amenable technique such as quantitative PCR.

4. Results

4.1. Isolation and characterization of CAFs and NCF

Cultures of carcinoma-associated fibroblasts (CAFs) and their paired normal colonic fibroblasts (NCFs) from adjacent colon mucosa from 8 patients were successfully established. Additionally, 5 NCFs and 5 CAFs (unpaired samples) were also isolated. The fibroblast cell population was first verified by cell morphology under the microscope. Figure 1A shows the typical spindle-like shape of the NCFs and CAFs. The purity of isolated fibroblasts was checked by detection of specific fibroblast biomarkers. These cells strongly expressed vimentin, vinculin, and α-smooth muscle actin (αSMA), whereas they were mostly negative for cytokeratin 18, E-cadherin (epithelial markers), and VE-cadherin (endothelial marker). These expression changes were checked by qRT-PCR or western blot in the 8

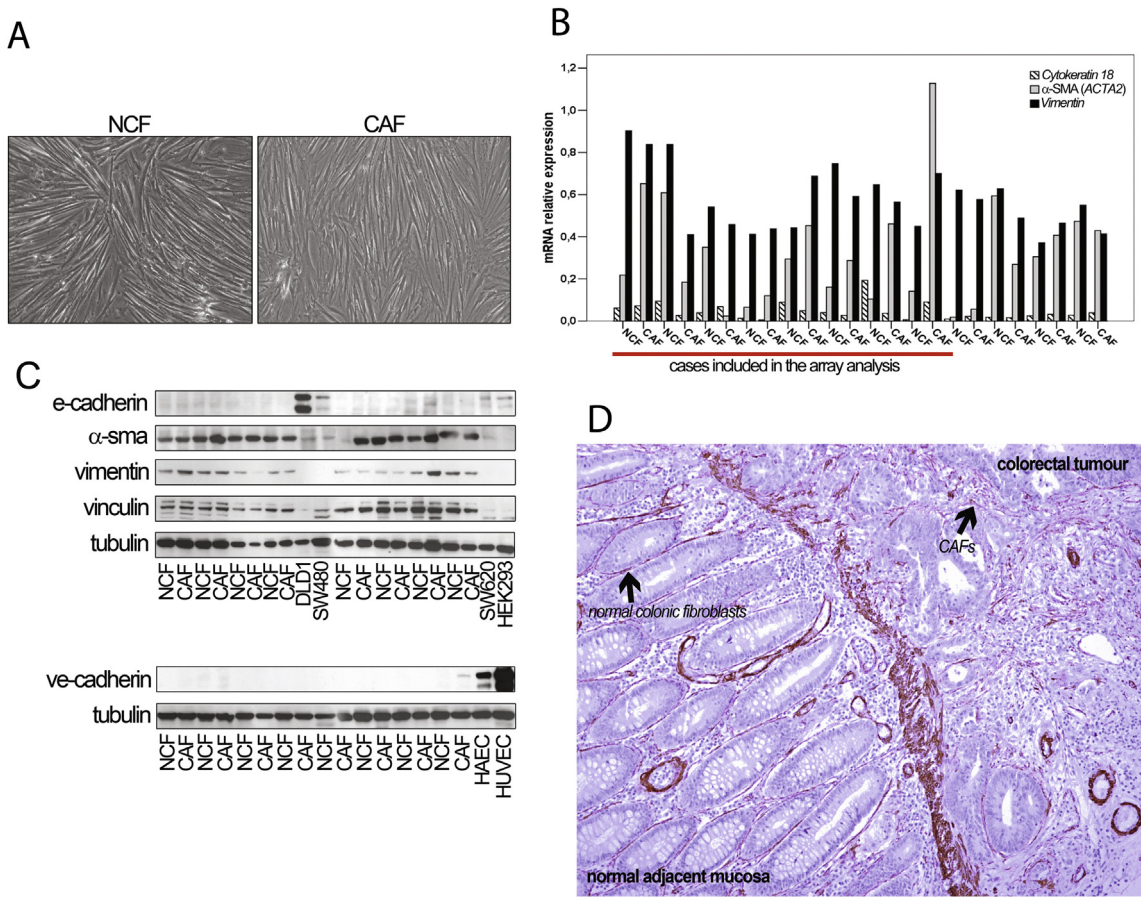

Figure 1 – Characterization of fibroblasts isolated from human colon cancer tissue. A) Cultured CAFs and their normal NCF counterparts visualized under the microscope, showing their characteristic cell morphology. B) Relative expression of specific mesenchymal markers such as vimentin and α-SMA in the 8 paired NCF/CAF cell lines (samples included in the array analysis) and in four more pairs NCF/CAF. None of them showed a specific pattern between the two types of fibroblasts, indicating that although they were good mesenchymal biomarkers, they did not discard between NCFs and CAFs. Low levels of the epithelial-specific marker cytokeratin 18 confirmed the purity of the fibroblast cultures. C) Protein levels of additional specific fibroblastic markers were tested by western blot, confirming the purity of the fibroblast cultures (samples are ordered in pairs). D) Immunohistochemical staining of α-SMA in a specimen from a colorectal cancer patient, illustrating staining of myofibroblasts in both compartments, normal adjacent mucosa and adenocarcinoma. All the fibroblasts surrounding the non-tumoral crypts are positive for α-SMA.

paired CAFs and NCFs and additional unpaired fibroblasts (Figure 1B and C). The results indicated that fibroblast cells were correctly isolated from tumour tissue and adjacent mucosa. Astonishingly, levels of αSMA in NCF were in some cases comparable to those in CAFs. To exclude any interference because of *in vitro* culturing we checked αSMA staining in tumour specimens. As displayed in Figure 1D, normal adjacent mucosa also stained for αSMA, confirming that some fibroblasts in adjacent normal mucosa are also myofibroblasts an observation also recently noticed by Hawinkels et al. (2014).

4.2. *CAFs enhance aggressiveness of colorectal cancer cell lines* in vitro

In a wound-healing assay, CAF CM increased the migratory potential of DLD1 cancer cells compared with NCF CM (Figure 2A). We also observed that when different colorectal cancer cell lines (DLD1, SW620, SW480 and SW1116) were exposed to CM from CAFs, their clonogenic capability was enhanced after 9 days of culture (Figure 2B). We previously described the differential proliferative capacity of CMs depending on the fibroblast origin in different colorectal cell lines (Berdiel-Acer et al., 2011).

4.3. In vivo *tumorogenic assay*

In an *in vivo* assay, when DLD1 cells were co-injected in combination with NCF, there were no differences weight compared with the DLD1 cells injected alone ($P = 0.460$, U Mann–Whitney test). However, co-injection of DLD1 with CAFs significantly accelerated tumour growth (Figure 2C). Additionally we observed an intense Ki67 staining in tumours generated with the coinjection of DLD1 cells + CAF while the staining was very mild when coinjection were made with NCF (Figure 2D). Interestingly, after 21 days we still observed the presence of CAFs but not of NCFs through human-specific vimentin staining, demonstrating the replicative capabilities of CAFs in a xenogeneic environment (Figure 2D).

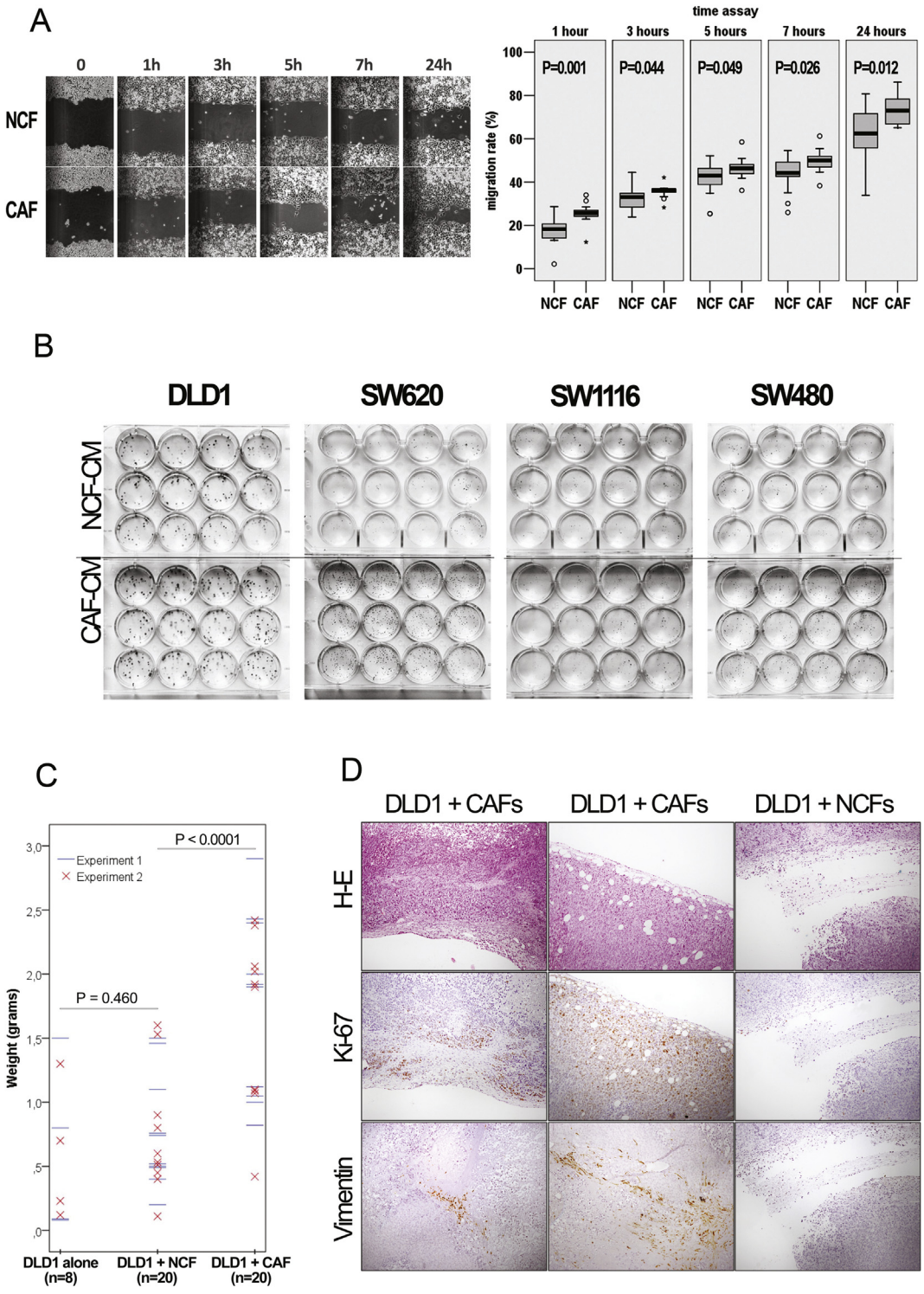

Figure 2 – Functional assays of CAFs *vs.* NCFs. A) In a wound-healing assay, DLD1 cells significantly increased their migratory potential when cultured with CAF conditioned medium (CM), compared with their normal paired CM (left and right panels). B) The clonogenic capacity of four cell lines (DLD1, SW620, SW480 and SW116) was greater in the presence of 6 CMs from CAFs compared with 6 CMs from NCFs (two replicates per CM). C) Co-injection of CAFs with DLD1 cancer cells into athymic mice significantly enhanced tumour growth compared with NCFs ($P < 0.0001$). D) Hematoxylin-eosin and immunohistochemical staining. 21 days after coinjection, human specific vimentin were detected in almost all slides of tumours generated with DLD1 cells + CAFs. No staining was detected in tumours DLD1 cells + NCF. In addition, DLD1 cells in tumours generated with CAFs displayed high levels of Ki67 staining at the periphery of the tumours, whereas the staining was very mild in case of coinjection with NCF.

4.4. Genes differentially expressed between CAFs and paired NCFs

To explain the biological heterogeneity between CAFs and NCFs in the primary colorectal cancer microenvironment, the gene expression profile of CAFs and their normal counterparts from 8 colorectal cancer patients was obtained using the Affymetrix GeneChip Human Gene 1.0 ST Array. Microarray data were normalized with RMA and differentially expressed genes (DEG) were identified using *Significance Analysis in Microarrays* (SAM). The gene list obtained was ranked according to the FDR q-value <0.05 and a >2-fold change (in genes overexpressed in CAFs) and a <0.5-fold change (genes underexpressed in CAFs). We identified 109 probes that were differentially expressed in CAFs *vs.* NCFs, 39 upregulated probes corresponding to 38 known genes, and 70 downregulated probes corresponding to 70 known genes. The ranked list of the 108 genes deregulated between NCF and CAF is summarized in Supplementary Table 1. Expression profiles of statistically significant genes overexpressed in CAFs and NCFs were explored by unsupervised sample-based cluster analysis. Hierarchical clustering clearly separates CAFs and NCFs and allowed us to classify matched samples depending on whether they came from tumour or from adjacent mucosa (Figure 3A). We checked the expression of some selected genes (*CDH2*, *ENC1*, *TNFSF4*, *ST6GALNAC5*, *SEMA5A*, *SLC7A2* and *TGFB2*) by western blot, and confirmed that in the paired fibroblasts tested, overexpression was also translated at protein level to almost all pairs (Figure 3B). Quantitative RT-PCR validation of some of these differentially expressed genes in samples of colonic mucosa from healthy patients ($n = 10$), adjacent colonic mucosas ($n = 10$) and paired colorectal tumours ($n = 10$) confirmed the expression values obtained in cultured fibroblasts (Figure 3C).

4.5. Normal colonic fibroblasts adjacent to the tumour expressed pro-inflammatory genes

Surprisingly, many proinflammatory cytokines, FGFR ligands and other secreted proteins of the TNF family were overexpressed in NCF compared to counterpart CAFs. This fact was validated comparing mRNA levels in 50 healthy colonic mucosa and 98 paired normal adjacent mucosa and colorectal carcinoma (GSE44076). As shown in Supplementary Figure 1A, the expression was confirmed, demonstrating that normal adjacent mucosa has already proinflammatory changes compared to normal healthy mucosa (*PROS1*, *FGF13*, *CCL8*, *CXCL12*, *FGF13*, *SLIT3*, *FGF7*, *FGF10*, *SFRP1*) suggesting certain degree of activation of fibroblasts from the adjacent mucosa. Such activated status was also confirmed for classical fibroblast activation markers like endosialin (CD248), αSMA (ACTA2) and vimentin (VIM) in samples from healthy mucosas, adjacent mucosas and paired tumours (Supplementary Figure 1B). Other cytokines like *CCL11*, *CCL13* or *TNFSF10* are higher in both healthy mucosa and adjacent mucosa compared to tumour.

4.6. Fibroblast specificity of the differentially expressed genes

Additionally, the CAF-specificity of the 108 deregulated genes was checked in GSE39396, a dataset corresponding to six different colorectal cancer patients and the expression profile of four sorter-isolated cell types (endothelial cells: CD31+, epithelial cells: EPCAM+, inflammatory cells: CD45+ and CAFs: FAPα+) (Calon et al., 2012). According to GSE39396 data, many of the genes could be considered to be solely expressed in fibroblast cells (those framed in green in Figure 4A and B; i.e: *TGFB2*, *AMIGO2*, *PDLIM3*, *ENC1* or *RARB*). A small proportion of genes are basically expressed at a higher level in epithelial cells (framed in red). Other genes are more characteristic of endothelial cells, and finally, there is a group of genes that are expressed ubiquitously in various stromal cells.

To validate the microarray data, the expression levels were analysed by qRT-PCR in isolated fibroblasts (8 pairs included in the array and 5 more pairs), and in a set of 13 colorectal cancer cells (CRCs). Validated genes were chosen according to their gene product, considering some of those coding for membrane proteins and soluble factors such as *TNFSF4*, *SEMA5A*, *ST6GALNAC5*, *CLDN1*, *EFNB2*, *PKP2* (Rickelt et al., 2009), *NTF3* (Louie et al., 2013), *TGFB2*, *ULBP2* and *CXCL12* (Figure 4C). Additionally we validated some genes with a >2-fold change and a q-value between 0.05 and 0.10, like *WNT2*, *NRXN3*, *INHBA*, *FGF13* and *COL11A1* (Figure 4D). The expression pattern of the genes differentially expressed between NCFs and CAFs detected by qRT-PCR was consistent with the microarray results and the absence or low level of expression of most of the selected genes in a set of 13 epithelial colon cancer cell lines confirmed their stromal specificity.

4.7. Functional annotation and GSEA

To evaluate the biological significance of coordinated variation in expression changes observed between NCF and CAF transcriptomic programs, we performed GSEA for the complete list of ranked genes. We determined that the expression profile of overexpressed genes in CAFs correlated with fibroblast serum response genes, whereas underexpressed genes were associated with serum or core serum response repressed genes (Supplementary Figure 2). These results were consistent with those of Chang et al. (2004), who suggested that in prostate and hepatocellular carcinoma, normal tissue samples had the serum-repressed signature, while tumours had the serum-induced signature or "wound-healing phenotype". GSEA also revealed that genes deregulated in CAFs were related to different KEGG pathways, the most representative overexpressed pathways with FDR <0.01 are depicted in Table 1. DAVID (Database Annotation, Visualization and Integrated Discovery) was used to classify the gene function of deregulated genes in CAFs and their paired NCFs applying a q-value of <0.1 as a cut-off (305 underexpressed probes, $p = 0.003$; 97 overexpressed probes, $p = 0.001$). Gene ontology analysis (Supplementary Table 2) revealed that genes overexpressed in CAF were associated with biological processes related to development (*TGFB2*, *PDGFC*, *cMET*, *CADM1*, *WNT2*) and cell–cell signalling (*TFAP2C*, *NTF-3*, *SEMA5A*, *EFNB2*, and *INHBA*). On the other hand, underexpressed genes were clustered in categories involved in homeostasis (*SNCA*, *AGT*, *ABCA1*, *PLA2G4A*, and *BCL-2*), response to external stimulus (*CCL8*, *CCL7*, *CCL11*, *CCL13*, *proS1*) and cell–cell signalling (*FGF10*, *FGF13*, *TNFSF10*, *NOVA1*).

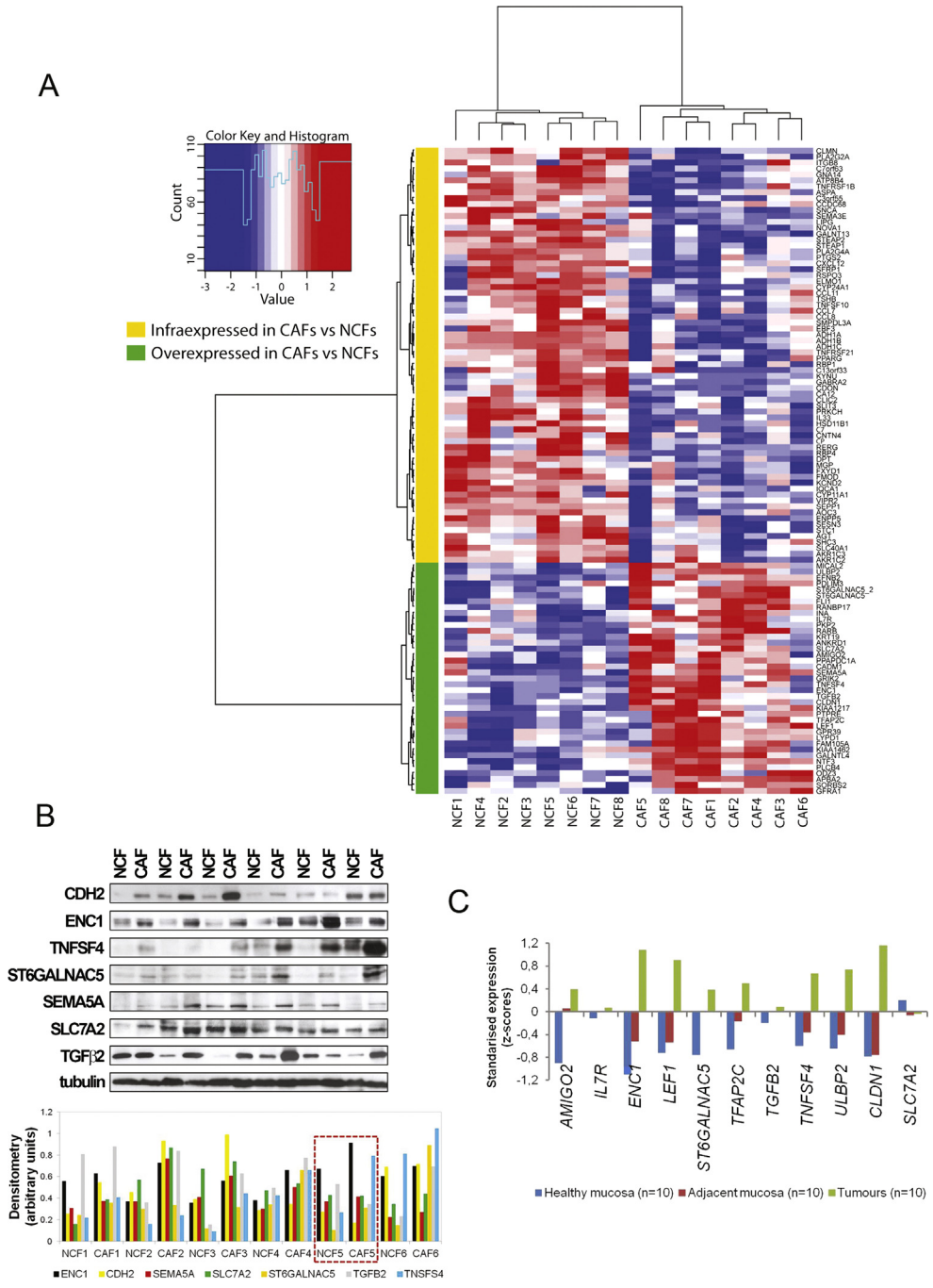

Figure 3 – Unsupervised heatmap of deregulated genes in CAF *vs.* NCF. A) Genes differentially expressed in 8 paired CAFs *vs.* NCFs. A total of 108 differentially expressed genes, 70 downregulated (yellow side bar) and 38 upregulated (green side bar), were identified by SAM analysis. Fibroblasts were clustered independently of their patient origin. B) Expression of some upregulated genes in CAFs was tested and confirmed at the protein level by western blot, except for the pair highlighted with the red dashed line in the densitometry analysis (CDH2, SLC7A2 and TGFB2). C) Quantitative RT-PCR validation of 11 of the differentially expressed genes in 10 healthy colonic mucosas (blue bars), 10 adjacent normal mucosas (red bars) and 10 paired tumours (green bars). Expression values are expressed as *z*-scores obtained from the relative expression values normalised for two housekeeping genes (*PMM1* and *ACTB*).

4.8. *Interaction between NCF and tumour cells*

Transcriptomic profile of NCFs change after 5 days coculture in transwell inserts with tumour cells (DLD1 cells), acquiring expression values for many of the differentially expressed genes similar to those observed in paired CAFs (Supplementary Figure 3). The paracrine interaction between both cell types induces gene expression changes in NCF even for inflammatory cytokines like CCL11, CCL8 or the apoptotic inducer TNFSF10 (Supplementary Figure 3B, low panel).

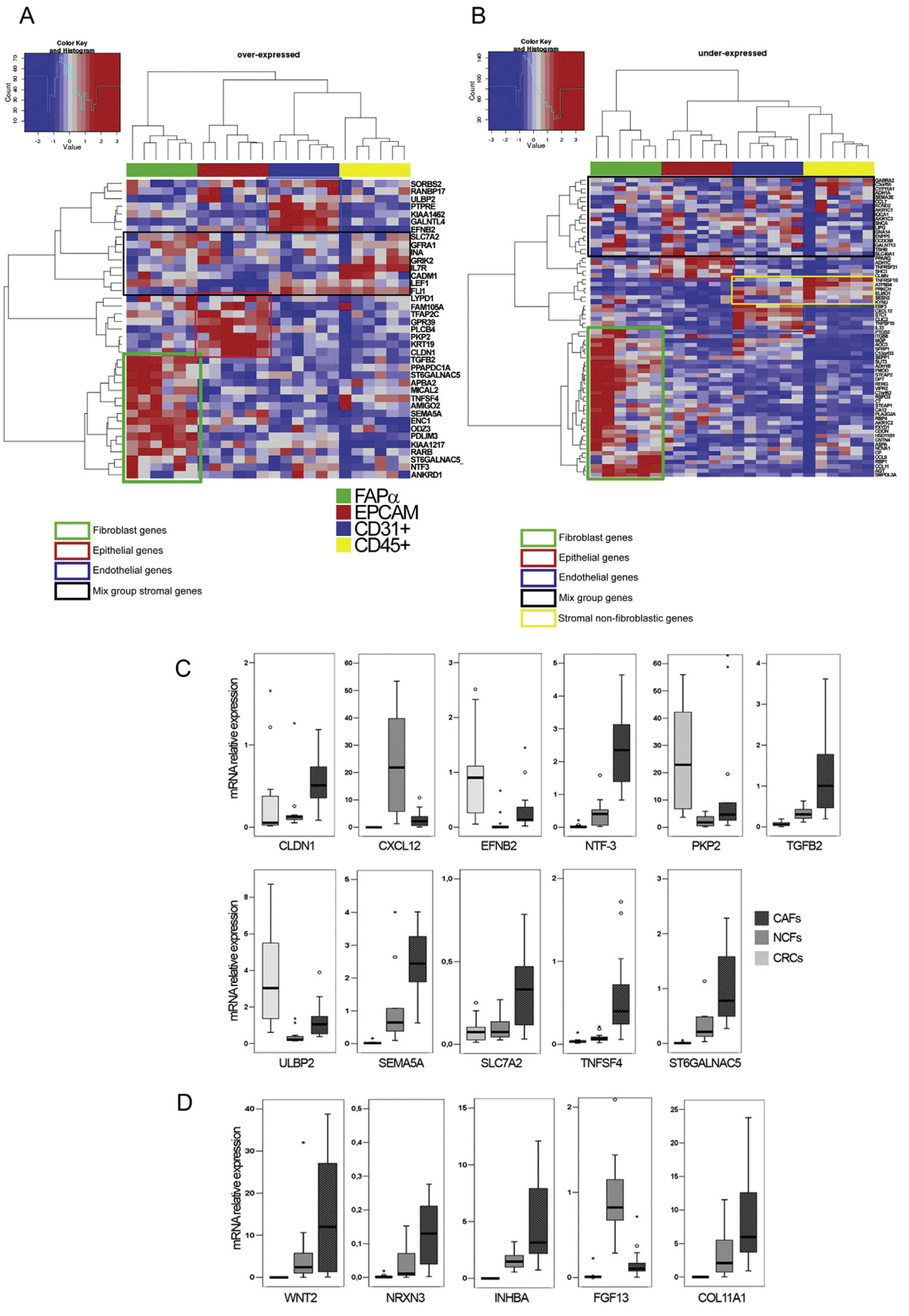

Figure 4 – CAF specificity of deregulated genes. A-B) Unsupervised heatmaps of deregulated genes (overexpressed and underexpressed, respectively) in the GSE39396 dataset. The expression profile of 108 deregulated genes was checked in 4 CRC cell populations: FAPα+ (fibroblasts), EPCAM+ (epithelial cells), CD45+ (leukocytes) and CD31+ (endothelial cells). Gene expression levels clearly classified cells into the correct class. Of the complete list of genes, almost half are specific to the fibroblast cell type population (green box). A set of stromal genes heterogeneously expressed in all cell types but not in epithelia are depicted in black as a mixed group of stromal genes. A group of genes is mostly

4.9. *Relevance of signature score*

Signature scores for the isolated fibroblasts used in the array were obtained by adding up the expression values for all 108 deregulated genes. The score confirmed that the expression of those selected genes allowed the two fibroblast populations to be discriminated (Figure 5A) and emphasized the relevance of genetic disparity between the two types of fibroblasts even though they had the same origin (same patient). To identify any phenotypic differences associated with score values, we chose the CAFs that differed most in expression (high and low scores) and evaluated differences in migration and proliferation. In a wound-healing assay, when we assessed the migratory capacity of fibroblasts, we found higher scores were associated with more migration (Figure 5B). Concerning proliferation ability, after a 144 h MTT assay, high-score CAFs showed greater proliferation compared with their low-score counterparts (Mann–Whitney test, $p = 0.05$; Figure 5C). When epithelial tumour cells (DLD1) were cultured with CAF CM (mentioned above), an increase in cell proliferation was evident if cultured with high score signature CM compared with the one from low-score signature CAFs (Mann–Whitney test, $p = 0.034$; Figure 5D).

4.10. *Prognostic value of the signature score*

We wanted to assess the association of the score of the differentially expressed genes with prognosis in colorectal cancer. Using two different independent datasets, the pooled cohort GSE14333-17537 (Stages I-to-III) and GSE33113 (Stage II), using the mean value, zero, as a cut-off, patients with a low signature score displayed a significantly different disease-free survival time than high-scoring patients (Cox regression, $p = 0.0038$ for pooled cohort GSE14333-GSE17537, $p = 0.026$ for GSE33113; Figure 5E upper left panel and lower panel). Treating the signature score as a continuous variable, the hazard ratio was HR = 1.48 (95% CI 1.13–1.94, $p = 0.004$ per unit increase in the pooled cohort GSE14333-GSE17537 and HR = 1.89 (95% CI 1.14–3.12, $p = 0.011$) in GSE33113.

We validated the expression of some of the differentially expressed genes by means of quantitative RT-PCR in 30 samples from non-recurrent stage II/III tumours and 30 samples from recurrent stage II/III tumours. As illustrated in Figure 5G, genes that we found overexpressed in CAFs vs NCFs were higher in recurrent tumours and genes overexpressed in NCFs vs CAFs were higher in non-recurrent tumours.

5. Discussion

Using microarray technologies, this study identifies for the first time transcriptomic differences between carcinoma-associated fibroblasts (CAFs) and their paired normal colonic fibroblasts (NCFs) in patients with primary colorectal carcinomas. We defined a 108-gene signature that better characterises each fibroblast subpopulation and is also correlated with the proliferative and migratory properties of the CAFs. Thus, the higher the score signature, the more competent the fibroblasts are in terms of proliferation and migration.

We have previously reported differential effects at the functional level of CM from unpaired fibroblasts from locations where a colorectal cancer is situated (i.e., normal mucosa, primary tumour or a liver metastasis) (Berdiel-Acer et al., 2011). Other authors compare transcriptomic profiles of CAFs *vs* normal fibroblasts from healthy tissues (Torres et al., 2013). In the present work, to go further and relate transcriptomic differences with the phenotype, we ascertained that CAFs clearly enhanced the aggressiveness of colorectal cancer cells with respect to their NCF counterparts. We demonstrated that CM obtained from CAFs confers a higher migratory and clonogenic capacity than in paired NCF and is dependent on their signature score. We also confirmed that CAFs are more proficient *in vivo* as tumour promoters than NCFs from the same patient, although being both cell types myofibroblasts.

Differences between the two cell subpopulations have also been described in terms of gene expression profiles in several solid tumours (Bauer et al., 2010; Costea et al., 2013; Hawsawi et al., 2008; Navab et al., 2011; Peng et al., 2013); but, as far as we are aware, not for colorectal cancer using paired samples.

In addition, all these aforementioned published works are focused on the assumption that CAFs present in malignant tissues are derived from normal resident fibroblasts on healthy tissues (Bauer et al., 2010; Costea et al., 2013; Hawsawi et al., 2008; Navab et al., 2011; Peng et al., 2013). Other authors suggested that different cells could be precursors of myofibroblasts, like bone marrow mesenchymal stem cells (MSCs) (Quante et al., 2011; Guo et al., 2008) or from epithelial normal or transformed cells via epithelial to mesenchymal transition (EMT), or finally from endothelial cells via endothelial to mesenchymal transition (EndMT) (Cirri and Chiarugi, 2012). Although we cannot exclude any of the previous hypothesis and even many of them might be happening simultaneously (Worthley et al., 2010), in our opinion the activation and recruitment of normal resident pericryptal fibroblasts is the most plausible theory, since the coculturing of NCF with tumour cells induces transcriptional changes in fibroblasts acquiring a transcriptomic profile more similar to their paired CAFs (Supplementary Figure 3).

We have assessed the gene expression profile of CAFs and their normal paired counterparts in colon cancer samples in order to identify: (i) relevant pathways responsible for pro-tumorogenic effects; (ii) new biomarkers for cancer treatment focused on the stroma and (iii) we evaluate the prognostic value of differentially expressed genes. Using 8

expressed in endothelial cells (blue box) but not in the other cell types. C) Expression levels of selected deregulated genes was validated by qRT-PCR in paired ($n = 12$), as well as in a set of 13 colorectal cancer cells lines. Results confirmed the microarray data and demonstrated stromal specificity in the vast majority of selected genes. D) We expanded PCR validation to some genes with a $>$2-fold change in expression but with a false discovery rate q-value < 0.1.

Table 1 – GSEA analysis (Kegg pathways) of genes deregulated in CAFs.

Name	Size	ES	NES	NOM *p*-val	FDR *q*-val
GSEA KEGG pathways overexpressed in CAFs					
HSA04510_FOCAL_ADHESION	189	0.235	3441	0	0.00E+00
HSA04110_CELL_CYCLE	111	0.290	3297	0	0.00E+00
HSA04310_WNT_SIGNALING_PATHWAY	146	0.219	2878	0	0.00E+00
HSA04810_REGULATION_OF_ACTIN_CYTOSKELETON	197	0.186	2766	0	6.28E−04
HSA05222_SMALL_CELL_LUNG_CANCER	85	0.274	2756	0	5.02E−04
HSA05213_ENDOMETRIAL_CANCER	52	0.337	2754	0	4.19E−04
HSA04360_AXON_GUIDANCE	126	0.207	2590	0	9.67E−04
HSA05215_PROSTATE_CANCER	86	0.244	2528	0	1.11E−03
HSA05223_NON_SMALL_CELL_LUNG_CANCER	53	0.308	2523	0	9.88E−04
HSA05214_GLIOMA	63	0.287	2518	0	8.89E−04
HSA04520_ADHERENS_JUNCTION	75	0.248	2429	0	1.81E−03
HSA04514_CELL_ADHESION_MOLECULES	129	0.202	2420	0	1.75E−03
HSA05220_CHRONIC_MYELOID_LEUKEMIA	76	0.250	2405	0	1.78E−03
HSA04670_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION	107	0.209	2346	2.08E−03	2.38E−03
HSA05211_RENAL_CELL_CARCINOMA	68	0.250	2299	2.10E−03	3.36E−03
HSA04115_P53_SIGNALING_PATHWAY	65	0.248	2249	4.12E−03	4.43E−03
HSA04512_ECM_RECEPTOR_INTERACTION	86	0.217	2214	0.00E+00	5.32E−03
HSA04530_TIGHT_JUNCTION	129	0.175	2180	2.11E−03	6.61E−03
HSA05212_PANCREATIC_CANCER	73	0.229	2177	2.07E−03	6.33E−03
HSA00562_INOSITOL_PHOSPHATE_METABOLISM	47	0.275	2168	2.20E−03	6.18E−03
HSA05210_COLORECTAL_CANCER	84	0.212	2164	2.11E−03	6.08E−03
HSA05130_PATHOGENIC_ESCHERICHIA_COLI_INFECTION_EHEC	50	0.266	2163	2.11E−03	5.80E−03
HSA01430_CELL_COMMUNICATION	135	0.172	2150	2.06E−03	5.99E−03
HSA05131_PATHOGENIC_ESCHERICHIA_COLI_INFECTION_EPEC	50	0.266	2140	2.02E−03	6.05E−03
HSA03050_PROTEASOME	22	0.388	2113	4.17E−03	7.16E−03
HSA04012_ERBB_SIGNALING_PATHWAY	83	0.207	2101	2.24E−03	7.70E−03
HSA05221_ACUTE_MYELOID_LEUKEMIA	53	0.255	2098	4.36E−03	7.50E−03
HSA04070_PHOSPHATIDYLINOSITOL_SIGNALING_SYSTEM	73	0.218	2041	8.44E−03	1.13E−02
HSA04540_GAP_JUNCTION	93	0.186	1981	8.32E−03	1.57E−02
HSA04120_UBIQUITIN_MEDIATED_PROTEOLYSIS	39	0.273	1944	4.19E−03	1.91E−02
HSA00190_OXIDATIVE_PHOSPHORYLATION	125	0.159	1940	6.52E−03	1.90E−02
HSA00970_AMINOACYL_TRNA_BIOSYNTHESIS	37	0.283	1929	1.43E−02	1.98E−02
HSA04210_APOPTOSIS	81	0.191	1919	2.11E−03	2.01E−02
HSA04916_MELANOGENESIS	99	0.176	1903	2.03E−03	2.19E−02
HSA00240_PYRIMIDINE_METABOLISM	85	0.183	1897	8.46E−03	2.18E−02
HSA04320_DORSO_VENTRAL_AXIS_FORMATION	27	0.311	1860	1.58E−02	2.58E−02
HSA05218_MELANOMA	71	0.185	1754	1.30E−02	4.45E−02
GSEA KEGG pathways infraexpressed in CAFs					
HSA00980_METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450	61	−0.438	−3.869	0.00E+00	0.00E+00
HSA00641_3_CHLOROACRYLIC_ACID_DEGRADATION	15	−0.728	−3.249	0.00E+00	0.00E+00
HSA00350_TYROSINE_METABOLISM	58	−0.382	−3.213	0.00E+00	0.00E+00
HSA00071_FATTY_ACID_METABOLISM	45	−0.423	−3.201	0.00E+00	0.00E+00
HSA00561_GLYCEROLIPID_METABOLISM	57	−0.364	−3.100	0.00E+00	0.00E+00
HSA00480_GLUTATHIONE_METABOLISM	35	−0.456	−2.993	0.00E+00	0.00E+00
HSA00120_BILE_ACID_BIOSYNTHESIS	38	−0.423	−2.931	0.00E+00	1.32E−04
HSA03320_PPAR_SIGNALING_PATHWAY	65	−0.306	−2.764	0.00E+00	3.39E−04
HSA00624_1_AND_2_METHYLNAPHTHALENE_DEGRADATION	23	−0.496	−2.748	0.00E+00	3.02E−04
HSA00010_GLYCOLYSIS_AND_GLUCONEOGENESIS	63	−0.298	−2.601	0.00E+00	4.44E−04
HSA00590_ARACHIDONIC_ACID_METABOLISM	50	−0.324	−2.581	0.00E+00	4.04E−04
HSA04060_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	245	−0.149	−2.324	0.00E+00	3.41E−03
HSA00380_TRYPTOPHAN_METABOLISM	58	−0.272	−2.278	0.00E+00	4.76E−03
HSA00220_UREA_CYCLE_AND_METABOLISM_OF_AMINO_GROUPS	30	−0.343	−2.181	1.89E−03	7.75E−03
HSA00280_VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION	44	−0.283	−2.167	5.66E−03	7.91E−03
HSA00512_O_GLYCAN_BIOSYNTHESIS	29	−0.339	−2.156	0.00E+00	8.27E−03
HSA03010_RIBOSOME	62	−0.241	−2.107	4.04E-03	9.76E−03
HSA04610_COMPLEMENT_AND_COAGULATION_CASCADES	68	−0.222	−2.060	1.97E−03	1.34E−02
HSA00340_HISTIDINE_METABOLISM	41	−0.271	−1.946	2.01E−03	2.68E−02
HSA00410_BETA_ALANINE_METABOLISM	25	−0.336	−1.897	5.92E−03	3.46E−02
HSA00640_PROPANOATE_METABOLISM	34	−0.289	−1.875	7.81E−03	3.73E−02
HSA02010_ABC_TRANSPORTERS_GENERAL	44	−0.249	−1.830	1.17E−02	4.63E−02

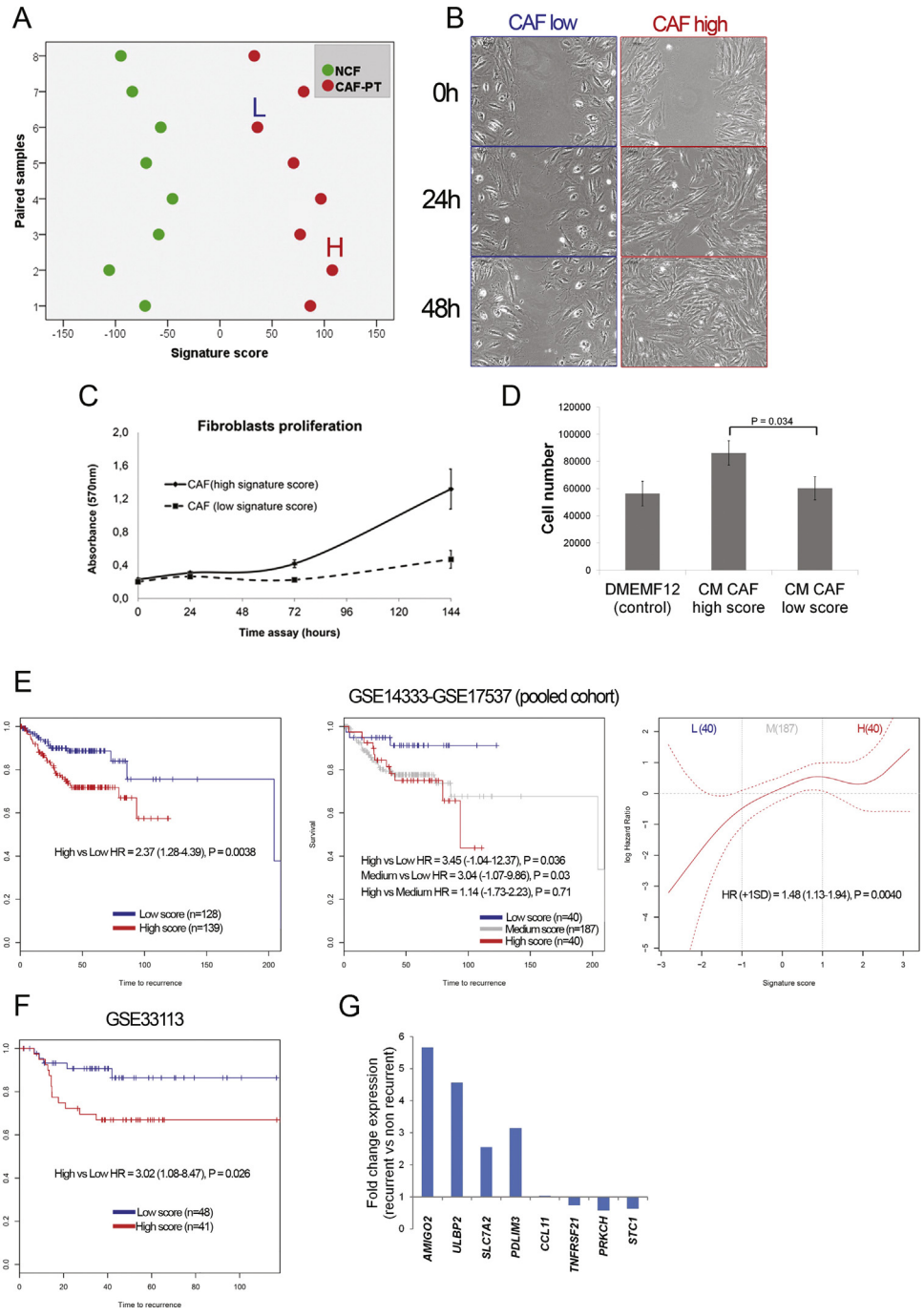

Figure 5 – Signature score. A) NCFs and CAFs used in the microarray were plotted according to their signature score (expression level of all 108 deregulated genes). Expression levels of those genes clearly separated the two fibroblast populations even though they came from the same patient. Fibroblasts used in subsequent functional assays are highlighted with a blue "L" (low score), and a red "H" (high score). B) Migration rates of CAFs were proportional to signature scores, the high score CAFs being the fastest to close the wound. C) CAFs with higher 108-signature scores had a higher proliferation rate than those with lower expression levels. D) When conditioned media from the same fibroblasts were added to a colorectal cancer cell line (DLD1), significant differences were observed between CM from a high score CAF vs CM from low score CAF (Mann–Whitney test, $p = 0.034$). E) Prognostic value of the differentially expressed genes in a pooled cohort (GSE14333-GSE17537) of stages I-to-III colorectal patients. Left and middle panels show Kaplan–Meier plots for disease-free survival, by the two risk groups (left panel; using the mean value as a cut off) or stratifying into three risk groups, using a smooth function of risk of recurrence based on the signature score (middle panel). The smooth function correlates the gene signature score with the relative risk of recurrence. Red dashed lines: 95% confidence interval (CI). Grey dashed lines: thresholds for patient selection into groups depicted in the middle panel: low (L; blue, 40 patients), medium (M; gray, 187 patients), and high (H; red, 40 patients); p values and increases in HR per standard deviation increase in expression (+1 SD) are shown (right panel). F) The prognostic value was corroborated using 90 colorectal patients stage II (GSE33113). G) Quantitative RT-PCR validation of some differentially expressed genes (four upregulated and four downregulated in CAFs vs NCFs) in 30 non-recurrent stage II/III colorectal tumours and 30 recurrent stage II/III colorectal tumours maintaining stages proportionality between the two groups.

paired NCF-CAF samples, we identified a subset of 108 genes that were differentially expressed between the two fibroblast populations. Unsupervised hierarchical cluster analysis of the expression data from the 108 genes clearly differentiated two patterns on the basis of the fibroblasts' anatomic location. Furthermore, the knowledge of genetic changes that occur when normal colonic fibroblasts from adjacent tissues are transformed into CAFs could help us elucidate pathways involved in the education of fibroblasts by tumour cells and to understand how to overcome such crosstalk between the cell types. Such information may provide relevant biomarkers for future treatments (target discovery) or tumour classification. Regarding canonical pathways responsible for pro-tumorogenic effects of CAFs, it is remarkable that GSEA correlates overexpressed genes in CAFs with a higher proliferative status (cell cycle) and signatures associated with fibroblasts response to serum (Chang et al., 2004) (wound healing phenotype). By contrast, integrating GSEA and DAVID (http://david.abcc.ncifcrf.gov) overexpressed genes in NCF from adjacent mucosa clustered into categories involved in homeostasis and metabolism. This fact might explain the dual role of the microenvironment as a tumour suppressor, while homeostasis is maintained, or as a tumour promoter when the homeostatic balance is lost (Bissell and Hines, 2011).

Initially astonishing, we observed higher levels of several proinflammatory cytokines, TNF family molecules and FGFR ligands in NCF. These values were validated using 50 colonic mucosa samples of healthy volunteers, 98 adjacent normal mucosa and their paired colorectal carcinomas, confirming higher levels of *CCL11*, *CCL8*, *CCL13*, *FGF10* or *TNFSF10*, among others, in normal adjacent mucosa compared to colorectal carcinomas. Even classical genes associated with an activated status of fibroblasts (*ACTA2*, *VIM*, *CD248*) were higher in normal adjacent mucosas compared with normal mucosas from healthy patients (Supplementary Figure 1B). Such validation confirms that adjacent mucosa already has changes, confirming also the values of αSMA detected in cultured NCF. In addition, someone might speculate why adjacent tissue responds with such activated state. One possibility might be a consequence of paracrine crosstalk with the tumour to create an adequate niche for the tumour to increase their size and invade. Another possibility could be a defensive response of the host. But further experiments are needed to solve this observation.

These results differ significantly compared to recent publications (Torres et al., 2013). In this manuscript, authors compare CAFs from mice with normal fibroblasts obtained from healthy mice. Additionally, chemically induced colitis-associated colorectal models in mice did not recapitulates the desmoplastic reaction and lack the invasive and metastatic phenotypes of human tumours (De Robertis et al., 2011). Moreover, stromal factors responsible for tumour promotion in colitis-associated colorectal cancer differ from the stromal factors, mainly fibroblasts, which promote sporadic colorectal cancer in humans. Thus, in order to find relevant molecules for target discovery, it would be more adequate to use specific CAF markers not expressed by their counterpart resident fibroblasts in adjacent tissues. Otherwise, undesired side effects may affect normal fibroblasts in adjacent tissue. In relation to this, some of the DEG in CAFs could be interesting druggable targets for treating cancer (especially those secreted or membrane receptors), in order to selectively remove those fibroblasts which exerted a protective action against the damage induced by chemotherapy on malignant cells (http://dgidb.genome.wustl.edu/). Retinoic acid receptor β (*RARB*) is consistently overexpressed in CAFs *versus* counterpart NCF from adjacent mucosa. Stromal RARβ has already been shown to promote mammary gland tumorogenesis, in large part, mediated through the CXCL12/CXCR4/ErbB2 signalling pathway. Ablation of *RARB* leads to a remodelling of the stroma during tumour progression that includes a decrease in angiogenesis, in the recruitment of inflammatory cells, and in the number myofibroblasts (Liu et al., 2011). The availability of different retinoids (i.e: Tamibarotene) could be interesting approaches for the treatment of cancer, especially to hamper the crosstalk between stroma and tumour cells. Another interesting gene differentially expressed in CAFs *versus* counterpart NCF to be targeted is *TGFB2*. At present there are therapies targeting specifically this growth factor (belagenpumatucel-L), currently in phase III in non-small cell lung cancer (Nemunaitis et al., 2009; Decoster et al., 2012).

Interestingly, there was a remarkable overlap considering our 108 deregulated genes from data previously obtained from breast (Bauer et al., 2010), lung (Navab et al., 2011) and oesophageal carcinoma (Zhang et al., 2009), whereby *TNFSF4*, *ST6GALNAC5*, *TGFB2*, *TFAP2C* and *LEF1* are among the most overexpressed genes, and *NOVA1*, *PDE3B*, *SLIT3*, *AKR1C1*, *AKR1C2* are among the most underexpressed. The degree of overlap is highly remarkable when relaxing the stringent criterion of SAM false-discovery rate q-value of <0.1. A fundamental consideration of the biological pathways involved in such transcriptional programs compared with previous data from breast, lung and oesophagus, (focal adhesion, Wnt signaling pathway, axon guidance, among others), leads us to consider the stimulated common processes in these site-specific CAFs. In other words, the crosstalk between tumour cells and fibroblasts in the microenvironment might depend on similar events and signalling pathways, with the consequences that this may have for finding good druggable targets to disrupt the communication between the two cell types. These observations suggest that, roughly, similar transcriptomic programs might be active in the transition from normal adjacent fibroblast to CAFs independently of the anatomic demarcation, and strongly support the existing correlation between cancer and the "wound-healing phenotype".

Another aspect is the degree of activation of such processes, since fibroblasts are highly heterogeneous cells as reported by Herrera et al. (2013). Other authors have suggested that such heterogeneity might be related to their possibly miscellaneous origin (Worthley et al., 2010). Regarding this heterogeneity, we have also demonstrated that the score of the 108-gene signature is correlated with functional parameters of CAFs, such as their migratory or proliferative capacity, and that this may link the signature with a desmoplastic reaction. We propose that fibroblasts with higher scores are those that are recruited in large quantities in highly desmoplastic colorectal tumours. Some authors have reported that the more fibroblasts there are in a tumour, the worse is the prognosis

(de Kruijf et al., 2011; Huijbers et al., 2013). In our study, we found an association with disease-free survival in two independent datasets of whole-tumour samples, whereby the higher the level of expression of those DEG, the worse the prognosis. Interestingly those genes have a prognostic value when considering Stage II patients (GSE33113). Since controversy exists when treating stage II patients with chemotherapy, CAFs might provide relevant information to take clinical decisions. Additionally, the performance of the signature in independent, whole tumour-derived data sets indicates that although the prognostic ability of those genes is basically specific of tumour stroma, the signal can be detected in and the signature can predict outcome in datasets from whole tissue. Additionally, we confirmed that the prognostic ability of our signature score was not simply a reflection of tumours with high fibroblast content by computing the quantity of fibroblasts from the expression values of *COL1A1* and *COL3A1* (as surrogate markers) in a cohort of 142 cases of whole-tumour colorectal specimens, and showing that high collagen tumours have the same risk of recurrence as low collagen tumours (data not shown). It should be noted that DEG have predictive value for patient outcome in non-advanced and advanced-stage tumours, indicating that the heterotypic interactions between stroma and epithelium are already relevant at the initial steps of the tumorogenic process. CAFs prognostic information could be then useful to select patients even at earlier stages that may benefit from treatment (i.e. Stage II risk tumours, T4N0).

Another important aspect might involve the fibroblast specificity of a given gene, especially considering possible biomarkers for target discovery. For some deregulated genes, expression was checked by means of quantitative RT-PCR in 13 CRC cell lines and also in four isolated populations (CAFs, epithelial cells, endothelial cells and CD45+ cells) obtained from six colorectal cancer patients. Approximately the 50% of deregulated probes belong to genes specifically expressed in fibroblasts and only 15% were genes mainly expressed by epithelial cells. Another group of genes could be considered to be stromal genes since they are also expressed in endothelial and inflammatory cells and even a group of ubiquitously expressed genes.

In order to find a reliable CAF gene classifier for prognostic purposes, the present work, using matched cases sharing the same genetic background, constitutes a filter to discard genes with less variation with counterpart NCF and genes ubiquitously expressed.

6. Conclusions

This study contributes evidence of the role of CAFs in the biology and progression of colorectal carcinoma and identifies an association between these phenotypic features and changes in gene expression compared with paired NCFs. We also identify new biomarkers that could differentiate normal adjacent fibroblasts from tumour fibroblasts and that could be useful for finding strategies to break the shell that protects tumour cells in desmoplastic tumours. Additionally, the subset of deregulated genes provides an opportunity to develop a gene classifier for categorising patients according to their level of risk.

Funding

ISCIII (PI07/0657) Spanish Government.

RSP has a grant funded by the AECC (Spanish Association Against Cancer) Scientific Foundation.

Conflict of interest statement

None declared.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.molonc.2014.04.006.

REFERENCES

Bauer, M., Su, G., Casper, C., He, R., Rehrauer, W., Friedl, A., 2010. Heterogeneity of gene expression in stromal fibroblasts of human breast carcinomas and normal breast. Oncogene 29, 1732–1740.

Berdiel-Acer, M., Bohem, M.E., Lopez-Doriga, A., Vidal, A., Salazar, R., Martinez-Iniesta, M., Santos, C., Sanjuan, X., Villanueva, A., Mollevi, D.G., 2011. Hepatic carcinoma-associated fibroblasts promote an adaptative response in colorectal cancer cells that inhibit proliferation and apoptosis: nonresistant cells die by nonapoptotic cell death. Neoplasia 13, 931–946.

Bissell, M.J., Hall, H.G., Parry, G., 1982. How does the extracellular matrix direct gene expression? J. Theor. Biol. 99, 31–68.

Bissell, M.J., Hines, W.C., 2011. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. Nat. Med. 17, 320–329.

Calon, A., Espinet, E., Palomo-Ponce, S., Tauriello, D.V., Iglesias, M., Cespedes, M.V., Sevillano, M., Nadal, C., Jung, P., Zhang, X.H., Byrom, D., Riera, A., Rossell, D., Mangues, R., Massague, J., Sancho, E., Batlle, E., 2012. Dependency of colorectal cancer on a TGF-beta-driven program in stromal cells for metastasis initiation. Cancer Cell 22, 571–584.

Chang, H.Y., Chi, J.T., Dudoit, S., Bondre, C., van de Rijn, M., Botstein, D., Brown, P.O., 2002. Diversity, topographic differentiation, and positional memory in human fibroblasts. Proc. Natl. Acad. Sci. U.S.A. 99, 12877–12882.

Chang, H.Y., Sneddon, J.B., Alizadeh, A.A., Sood, R., West, R.B., Montgomery, K., Chi, J.T., van de Rijn, M., Botstein, D., Brown, P.O., 2004. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS Biol. 2, E7.

Cirri, P., Chiarugi, P., 2012. Cancer associated fibroblasts: the dark side of the coin. Am. J. Cancer Res. 1, 482–497.

Costea, D.E., Hills, A., Osman, A.H., Thurlow, J., Kalna, G., Huang, X., Pena Murillo, C., Parajuli, H., Suliman, S., Kulasekara, K.K., Johannessen, A.C., Partridge, M., 2013. Identification of two distinct carcinoma-associated fibroblast

subtypes with differential tumor-promoting abilities in oral squamous cell carcinoma. Cancer Res. 73, 3888–3901.
de Kruijf, E.M., van Nes, J.G., van de Velde, C.J., Putter, H., Smit, V.T., Liefers, G.J., Kuppen, P.J., Tollenaar, R.A., Mesker, W.E., 2011. Tumor-stroma ratio in the primary tumor is a prognostic factor in early breast cancer patients, especially in triple-negative carcinoma patients. Breast Cancer Res. Treat. 125, 687–696.
De Robertis, M., Massi, E., Poeta, M.L., Carotti, S., Morini, S., Cecchetelli, L., Signori, E., Fazio, V.M., 2011. The AOM/DSS murine model for the study of colon carcinogenesis: from pathways to diagnosis and therapy studies. J. Carcinogenesis 10, 9.
de Sousa, E.M.F., Colak, S., Buikhuisen, J., Koster, J., Cameron, K., de Jong, J.H., Tuynman, J.B., Prasetyanti, P.R., Fessler, E., van den Bergh, S.P., Rodermond, H., Dekker, E., van der Loos, C.M., Pals, S.T., van de Vijver, M.J., Versteeg, R., Richel, D.J., Vermeulen, L., Medema, J.P., 2011. Methylation of cancer-stem-cell-associated Wnt target genes predicts poor prognosis in colorectal cancer patients. Cell Stem Cell 9, 476–485.
Decoster, L., Wauters, I., Vansteenkiste, J.F., 2012. Vaccination therapy for non-small-cell lung cancer: review of agents in phase III development. Ann. Oncol. 23, 1387–1393.
Elenbaas, B., Weinberg, R.A., 2001. Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation. Exp. Cell Res. 264, 169–184.
Guo, X., Oshima, H., Kitmura, T., Taketo, M.M., Oshima, M., 2008. Stromal fibroblasts activated by tumor cells promote angiogenesis in mouse gastric cancer. J. Biol. Chem. 283, 19864–19871.
Hawinkels, L.J., Paauwe, M., Verspaget, H.W., Wiercinska, E., van der Zon, J.M., van der Ploeg, K., Koelink, P.J., Lindeman, J.H., Mesker, W., ten Dijke, P., Sier, C.F., 2014. Interaction with colon cancer cells hyperactivates TGF-beta signaling in cancer-associated fibroblasts. Oncogene 33, 97–107.
Hawsawi, N.M., Ghebeh, H., Hendrayani, S.F., Tulbah, A., Al-Eid, M., Al-Tweigeri, T., Ajarim, D., Alaiya, A., Dermime, S., Aboussekhra, A., 2008. Breast carcinoma-associated fibroblasts and their counterparts display neoplastic-specific changes. Cancer Res. 68, 2717–2725.
Herrera, M., Islam, A.B., Herrera, A., Martin, P., Garcia, V., Silva, J., Garcia, J.M., Salas, C., Casal, I., Garcia de Herreros, A., Bonilla, F., Pena, C., 2013. Functional heterogeneity of Cancer-associated fibroblasts from human colon tumors shows specific prognostic gene expression signature. Clin. Cancer Res. 19, 5914–5926.
Huijbers, A., Tollenaar, R.A., v Pelt, G.W., Zeestraten, E.C., Dutton, S., McConkey, C.C., Domingo, E., Smit, V.T., Midgley, R., Warren, B.F., Johnstone, E.C., Kerr, D.J., Mesker, W.E., 2013. The proportion of tumor-stroma as a strong prognosticator for stage II and III colon cancer patients: validation in the VICTOR trial. Ann. Oncol. 24, 179–185.
Hwang, R.F., Moore, T., Arumugam, T., Ramachandran, V., Amos, K.D., Rivera, A., Ji, B., Evans, D.B., Logsdon, C.D., 2008. Cancer-associated stromal fibroblasts promote pancreatic tumor progression. Cancer Res. 68, 918–926.
Jorissen, R.N., Gibbs, P., Christie, M., Prakash, S., Lipton, L., Desai, J., Kerr, D., Aaltonen, L.A., Arango, D., Kruhoffer, M., Orntoft, T.F., Andersen, C.L., Gruidl, M., Kamath, V.P., Eschrich, S., Yeatman, T.J., Sieber, O.M., 2009. Metastasis-associated gene expression changes predict poor outcomes in patients with Dukes stage B and C Colorectal Cancer. Clin. Cancer Res. 15, 7642–7651.
Liu, X., Nugoli, M., Laferriere, J., Saleh, S.M., Rodrigue-Gervais, I.G., Saleh, M., Park, M., Hallett, M.T., Muller, W.J., Giguere, V., 2011. Stromal retinoic acid receptor beta promotes mammary gland tumorigenesis. Proc. Natl. Acad. Sci. U.S.A. 108, 774–779.
Louie, E., Chen, X.F., Coomes, A., Ji, K., Tsirka, S., Chen, E.I., 2013. Neurotrophin-3 modulates breast cancer cells and the microenvironment to promote the growth of breast cancer brain metastasis. Oncogene 32, 4064–4077.
Navab, R., Strumpf, D., Bandarchi, B., Zhu, C.Q., Pintilie, M., Ramnarine, V.R., Ibrahimov, E., Radulovich, N., Leung, L., Barczyk, M., Panchal, D., To, C., Yun, J.J., Der, S., Shepherd, F.A., Jurisica, I., Tsao, M.S., 2011. Prognostic gene-expression signature of carcinoma-associated fibroblasts in non-small cell lung cancer. Proc. Natl. Acad. Sci. U.S.A. 108, 7160–7165.
Nemunaitis, J., Nemunaitis, M., Senzer, N., Snitz, P., Bedell, C., Kumar, P., Pappen, B., Maples, P.B., Shawler, D., Fakhrai, H., 2009. Phase II trial of Belagenpumatucel-L, a TGF-beta2 antisense gene modified allogeneic tumor vaccine in advanced non small cell lung cancer (NSCLC) patients. Cancer Gene Ther. 16, 620–624.
Olumi, A.F., Grossfeld, G.D., Hayward, S.W., Carroll, P.R., Tlsty, T.D., Cunha, G.R., 1999. Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. Cancer Res. 59, 5002–5011.
Orimo, A., Gupta, P.B., Sgroi, D.C., Arenzana-Seisdedos, F., Delaunay, T., Naeem, R., Carey, V.J., Richardson, A.L., Weinberg, R.A., 2005. Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. Cell 121, 335–348.
Peng, Q., Zhao, L., Hou, Y., Sun, Y., Wang, L., Luo, H., Peng, H., Liu, M., 2013. Biological characteristics and genetic heterogeneity between carcinoma-associated fibroblasts and their paired normal fibroblasts in human breast cancer. PLoS One 8, e60321.
Quante, M., Tu, S.P., Tomita, H., Gonda, T., Wang, S.S., Takashi, S., Baik, G.H., Shibata, W., Diprete, B., Betz, K.S., Friedman, R., Varro, A., Tycko, B., Wang, T.C., 2011. Bone marrow-derived myofibroblasts contribute to the mesenchymal stem cell niche and promote tumor growth. Cancer Cell 19, 257–272.
Rickelt, S., Winter-Simanowski, S., Noffz, E., Kuhn, C., Franke, W.W., 2009. Upregulation of plakophilin-2 and its acquisition to adherens junctions identifies a novel molecular ensemble of cell-cell-attachment characteristic for transformed mesenchymal cells. Int. J. Cancer 125, 2036–2048.
Smith, J.J., Deane, N.G., Wu, F., Merchant, N.B., Zhang, B., Jiang, A., Lu, P., Johnson, J.C., Schmidt, C., Bailey, C.E., Eschrich, S., Kis, C., Levy, S., Washington, M.K., Heslin, M.J., Coffey, R.J., Yeatman, T.J., Shyr, Y., Beauchamp, R.D., 2010. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. Gastroenterology 138, 958–968.
Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U.S.A. 102, 15545–15550.
Torres, S., Bartolome, R.A., Mendes, M., Barderas, R., Fernandez-Acenero, M.J., Pelaez-Garcia, A., Pena, C., Lopez-Lucendo, M., Villar-Vazquez, R., de Herreros, A.G., Bonilla, F., Casal, J.I., 2013. Proteome profiling of cancer-associated fibroblasts identifies novel proinflammatory signatures and prognostic markers for colorectal cancer. Clin. Cancer Res. 19, 6006–6019.
Tusher, V.G., Tibshirani, R., Chu, G., 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. U.S.A. 98, 5116–5121.

Wilson, C.L., Miller, C.J., 2005. Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. Bioinformatics 21, 3683–3685.

Worthley, D.L., Giraud, A.S., Wang, T.C., 2010. Stromal fibroblasts in digestive cancer. Cancer Microenviron. 3, 117–125.

Zhang, C., Fu, L., Fu, J., Hu, L., Yang, H., Rong, T.H., Li, Y., Liu, H., Fu, S.B., Zeng, Y.X., Guan, X.Y., 2009. Fibroblast growth factor receptor 2-positive fibroblasts provide a suitable microenvironment for tumor development and progression in esophageal carcinoma. Clin. Cancer Res. 15, 4017–4027.