nature genetics

Stromal gene expression defines poor-prognosis subtypes in colorectal cancer

Alexandre Calon[1], Enza Lonardo[1], Antonio Berenguer-Llergo[1], Elisa Espinet[1,6], Xavier Hernando-Momblona[1], Mar Iglesias[2–4], Marta Sevillano[1], Sergio Palomo-Ponce[1], Daniele V F Tauriello[1], Daniel Byrom[1], Carme Cortina[1], Clara Morral[1], Carles Barceló[1], Sebastien Tosi[1], Antoni Riera[1], Camille Stephan-Otto Attolini[1], David Rossell[1,6], Elena Sancho[1] & Eduard Batlle[1,5]

Recent molecular classifications of colorectal cancer (CRC) based on global gene expression profiles have defined subtypes displaying resistance to therapy and poor prognosis. Upon evaluation of these classification systems, we discovered that their predictive power arises from genes expressed by stromal cells rather than epithelial tumor cells. Bioinformatic and immunohistochemical analyses identify stromal markers that associate robustly with disease relapse across the various classifications. Functional studies indicate that cancer-associated fibroblasts (CAFs) increase the frequency of tumor-initiating cells, an effect that is dramatically enhanced by transforming growth factor (TGF)-β signaling. Likewise, we find that all poor-prognosis CRC subtypes share a gene program induced by TGF-β in tumor stromal cells. Using patient-derived tumor organoids and xenografts, we show that the use of TGF-β signaling inhibitors to block the cross-talk between cancer cells and the microenvironment halts disease progression.

About 40–50% of patients who have CRC with locally advanced disease (American Joint Committee on Cancer (AJCC) stage II to III) exhibit resistance to therapy and develop recurrent cancer over the course of treatment. Current CRC staging based on histopathology and imaging has a limited ability to predict prognosis, thus leading to attempts to elaborate molecular classifications. Several independent studies have recently proposed CRC subtypes based on distinct global gene expression profiles[1–4]. Although these studies differed regarding the number of tumor subtypes identified, they all concurred in concluding that poor patient outcome in CRC is associated with the expression of stem cell and mesenchymal genes in CRC cells[1–4]. This stem-like/mesenchymal CRC subtype represents a particular class of highly aggressive CRCs. These findings pave the way to improve the patient staging system and to identify molecules and signaling pathways associated with CRC metastasis and disease recurrence that could be targeted therapeutically.

RESULTS

Expression of poor-prognosis genes in tumor cell types

We studied the expression patterns of genes associated with disease relapse in CRC. To this end, we interrogated the three CRC patient data sets used to establish the above-mentioned molecular classifications and identified all the genes whose expression patterns were significantly associated with decreased disease-free survival intervals after surgery in each data set (hazard ratio for cancer recurrence (HR) > 1, $P < 0.01$) (**Fig. 1a**, gray). A high proportion of the genes found in this analysis were upregulated in the poor-prognosis molecular subtypes (**Fig. 1a**, red). De Sousa E Melo *et al.* distinguished three CRC classes named CCS1, CCS2 and CCS3 (ref. 1). Patients belonging to the CCS3 group have higher risk of recurrence after tumor resection than do patients with the other subtypes[1]. We identified 160 genes upregulated in CCS3 (>2-fold, $P < 0.05$) in comparison to the CCS2 and CCS1 subtypes that were positively associated with disease relapse (HR > 1, $P < 0.01$; **Fig. 1a** and **Supplementary Table 1**). Sadanandam *et al.* defined five distinct molecular CRC subtypes, including a subset of tumors that adopt expression programs similar to those of intestinal stem cells[2]. Patients bearing stem-like CRCs who had not been treated with chemotherapy after surgery had lower disease-free survival[2]. We found 286 genes positively associated with disease relapse that were upregulated (>2-fold, $P < 0.05$) in the stem-like tumor group in comparison to the other four subtypes (**Fig. 1a** and **Supplementary Table 1**). Marisa *et al.* defined six molecular subtypes, including one, named C4, which showed borderline association with increased risk of disease relapse[3]. We found 98 genes upregulated (>2-fold, $P < 0.05$) in the C4 subtype that were positively associated with disease relapse (**Fig. 1a** and **Supplementary Table 1**). All of these transcriptomic data sets were obtained by profiling whole-tumor samples. Thus, the expression of each gene could potentially be contributed by epithelial tumor cells or stromal cells, or by both populations. To investigate these possibilities, we analyzed the gene sets associated with poor prognosis in a transcriptomic data set

[1]Institute for Research in Biomedicine (IRB) Barcelona, Barcelona, Spain. [2]Department of Pathology, Hospital del Mar, Barcelona, Spain. [3]Cancer Research Program, Hospital del Mar Research Institute (IMIM), Barcelona, Spain. [4]Universitat Autónoma de Barcelona, Barcelona, Spain. [5]Institució Catalana de Recerca i Estudis Avançats (iCREA), Barcelona, Spain. [6]Present addresses: Division of Stem Cells and Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany (E.E.) and Department of Statistics, University of Warwick, Warwick, UK (D.R.). Correspondence should be addressed to E.B. (eduard.batlle@irbbarcelona.org).

Received 24 September 2014; accepted 28 January 2015; published online 23 February 2015; doi:10.1038/ng.3225

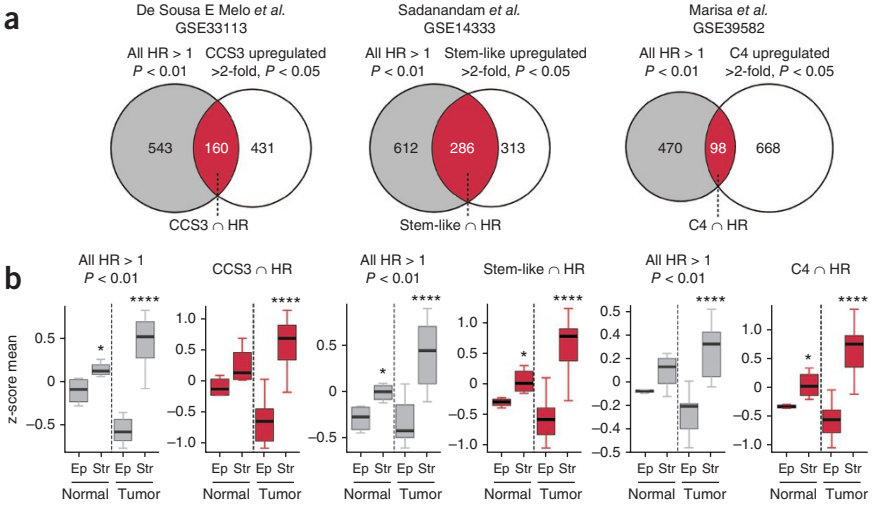

Figure 1 Gene sets defining poor-prognosis CRC subtypes are expressed in the tumor stroma. (**a**) Venn diagrams showing the overlap (red) between genes that predict relapse (all HR > 1, $P < 0.01$; gray) and genes upregulated in the poor-prognosis groups of their respective data sets (white). False discovery rate (FDR) and P values for intersections assuming hypergeometric distributions are FDR = 13.1% and $P < 2.22 \times 10^{-16}$ for CCS3 ∩ HR, FDR = 9.4% and $P < 2.22 \times 10^{-16}$ for stem-like ∩ HR, and FDR = 21.4% and $P < 2.22 \times 10^{-16}$ for C4 ∩ HR. For the GSE14333 cohort, only patients who did not receive chemotherapy were considered. (**b**) Mean expression (z score) of genes that predict relapse (gray) in the three studies in laser capture–microdissected (LCM) tumor epithelial (Ep) and stromal (Str) compartments or stroma and epithelial cells from normal colonic mucosa. Red boxes indicate equivalent analyses for the gene sets defining the CCS3, stem-like and C4 groups that hold predictive power for relapse (****$P < 0.0001$ for tumor stroma versus tumor epithelium, *$P < 0.05$ in normal stroma versus normal epithelium, Mann-Whitney test). Whiskers in box plots extend to minimum and maximum values.

of CRC and normal mucosa samples, in which epithelial cells and stromal cells had been microdissected by laser capture and profiled separately ($n = 13$)[5]. This analysis showed that the global set of genes positively associated with disease recurrence in the different patient data sets (all HR > 1, $P < 0.01$), as well as the gene subset upregulated in the poor-prognosis subtypes, were significantly upregulated in the microdissected tumor stroma in comparison to epithelial tumor areas and, to a lesser extent, in the stroma of the normal mucosa in comparison to the colonic epithelium (**Fig. 1b**). The upregulation of gene sets associated with poor prognosis in the tumor stroma was consistent across multiple thresholds of differential gene expression and significance (**Supplementary Table 2**).

We next analyzed the cell type–specific expression of the gene set associated with poor prognosis. To this end, we isolated various cell types from dissociated human primary CRC samples ($n = 14$) and performed global transcriptomic analysis of each population (**Fig. 2a**)[6]. Differential expression of cell type–specific marker genes in the purified fractions (**Fig. 2b**) confirmed enrichment of epithelial cancer cells (EpCAM$^+$), leukocytes (CD45$^+$), endothelial cells (CD31$^+$) and CAFs (FAP$^+$). Subsequent comparative analysis demonstrated that the global set of genes positively associated with disease recurrence (all HR > 1, $P < 0.01$) in each CRC data set was very significantly upregulated in stromal cell populations in comparison to epithelial tumor cells (**Fig. 2c**). In particular, the expression of these genes was most elevated in CAFs (FAP$^+$), followed by endothelial cells (CD31$^+$) and leukocytes (CD45$^+$) (**Fig. 2c**). This expression pattern was even more evident for the subset of genes upregulated in the poor-prognosis CRC subtypes (**Fig. 2c**). Taken together, these results indicate that tumor-associated stromal cells contribute a major proportion of the transcriptome positively associated with poor prognosis in the three molecular CRC classifications proposed.

Contribution of stromal genes to defining CRC subtypes

The above observations prompted us to study the relative contributions of epithelial and stromal genes to the identification of distinct molecular subtypes. Sadanandam and colleagues applied a 786-gene signature to distinguish between 5 patient subtypes[2]. For each probe set mapping these genes, we annotated association with cancer recurrence (HR), significant expression in the epithelial or stromal fraction from microdissected CRC samples and significant upregulation

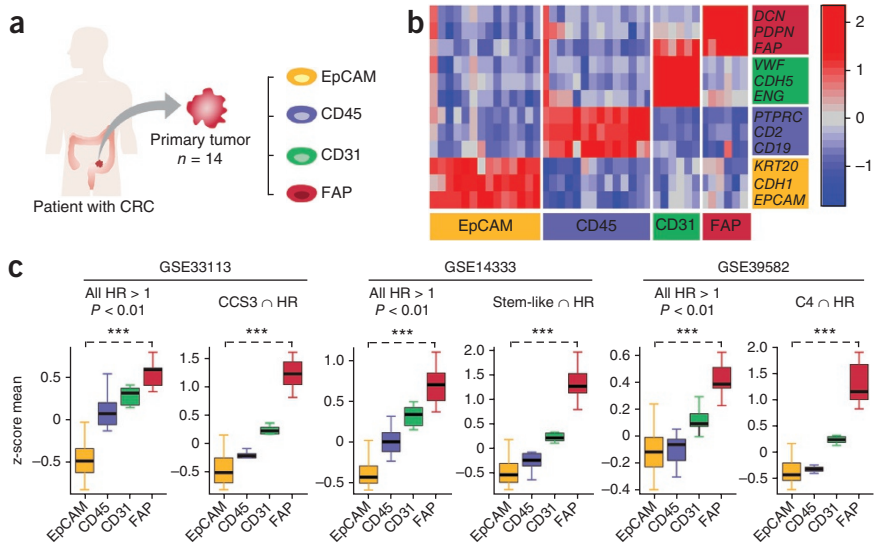

Figure 2 Gene sets defining poor-prognosis CRC subtypes are predominantly expressed in CAFs. (**a**) Scheme depicting the purification of cell populations from disaggregated primary CRC samples, enriching for specific cell types using the indicated markers: EpCAM, epithelial cells; CD45, leukocytes; CD31, endothelial cells; FAP, CAFs. (**b**) Heat map showing the expression levels of epithelial (*CDH1*, *EPCAM*, *KRT20*), leukocyte (*CD2*, *CD19*, *PTPRC*), endothelial (*CDH5*, *ENG*, *VWF*) and CAF (*DCN*, *PDPN*, *FAP*) marker genes in each FACS-purified cell population. Data represent normalized, centered and scaled Affymetrix probe intensities on a $\log_2$ scale. (**c**) z-score means of genes that predict relapse in the three depicted studies within the different human tumor cell types (***$P < 0.001$, Mann-Whitney test). Whiskers in box plots extend to minimum and maximum values.

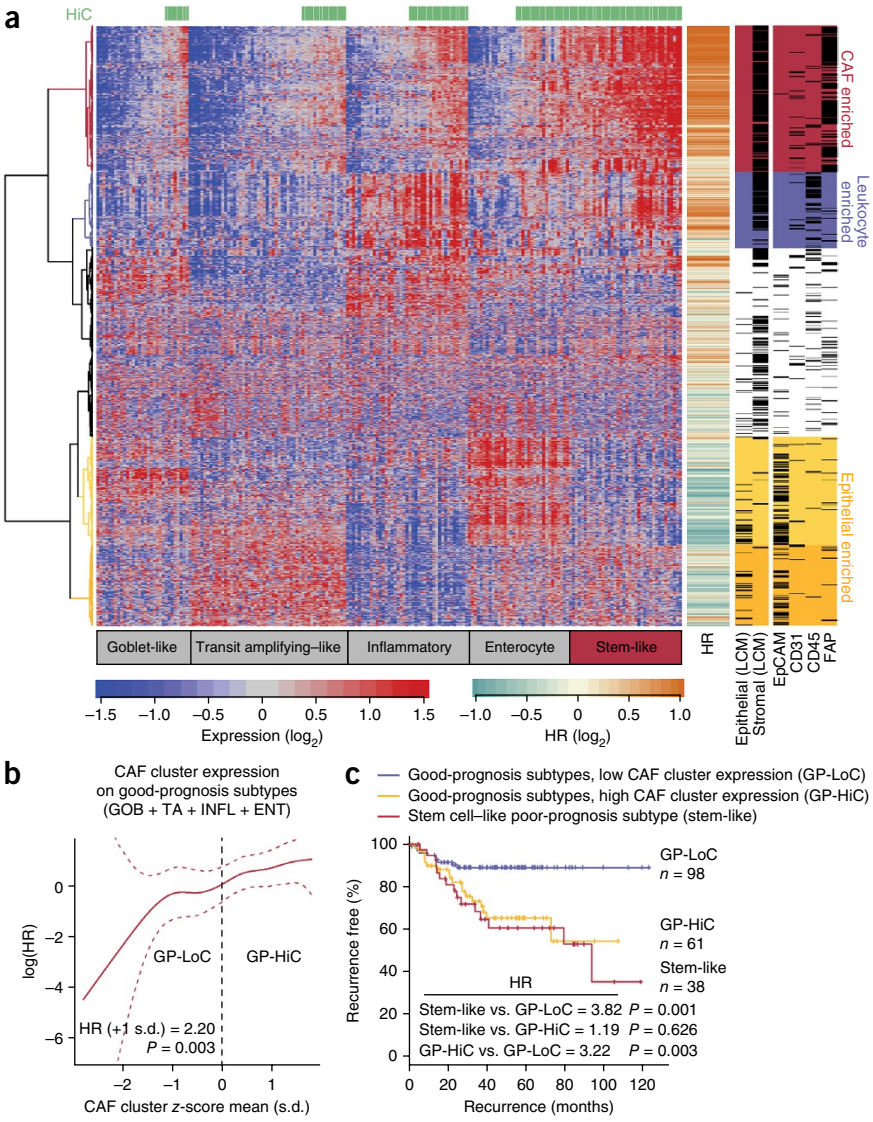

Figure 3 High expression levels of genes characteristic of CAFs identify poor-prognosis patients across CRC subtypes. (**a**) Clustering analysis of the 786-gene signature used to classify patients with CRC into subtypes by Sadanandam *et al.*[2] in the GSE14333 cohort. We allowed unsupervised hierarchical clustering of the 786 genes, while we enforced the classification of patients into subtypes. Data show normalized, centered and scaled Affymetrix probe set intensities on a $\log_2$ scale. The HR lane represents the hazard ratio for the corresponding genes on a $\log_2$ scale. Genes significantly upregulated in microdissected epithelial or stromal compartments[5] are depicted in the epithelial (LCM) and stromal (LCM) lanes, respectively. Genes specifically upregulated in epithelial, endothelial, leukocyte or FAP cell population are represented in the EpCAM, CD31, CD45 and FAP lanes, respectively. Patients across the entire cohort with high average expression (z score > 0) of the CAF-enriched gene cluster are marked in green (GP-HiC). (**b**) Smooth estimates of HR (+1 s.d.) show the higher risk of relapse for patients in good-prognosis subtypes (GOB, goblet-like; TA, transit amplifying–like; INF, inflammatory; ENT, enterocyte) presenting higher average expression of the CAF cluster (GP-HiC group). Dashed lines indicate 95% confidence bands. (**c**) Kaplan-Meier curves showing the recurrence-free survival of patients with good-prognosis cancer subtypes presenting low expression levels of the CAF cluster gene set (GP-LoC) and of patients within good-prognosis cancer subtypes but presenting high expression levels of the CAF cluster gene set (GP-HiC), with both compared to the stem cell–like poor-prognosis subtype (stem like). HR and P values are indicated.

in particular tumor cell populations. We performed hierarchical clustering on these annotated probe sets to explore their contribution to each molecular subtype in the GSE14333 data set (Gene Expression Omnibus (GEO)) (**Fig. 3**). This analysis identified several gene clusters, including two composed largely of genes upregulated in epithelial tumor cells (**Fig. 3a**, epithelial enriched) that were differentially expressed across the distinct CRC subtypes. Another cluster contained genes characteristic of leukocytes ($CD45^+$) that were particularly upregulated in CRCs of the inflammatory subtype (**Fig. 3a**, leukocyte enriched). The gene cluster most highly expressed in the stem-like poor-prognosis subtype contained a large proportion of genes upregulated in CAFs (**Fig. 3a**, CAF enriched). We then reclassified the CRC patient cohort either considering only the expression of genes included in the epithelial cluster or excluding genes belonging to the $CD45^+$ and FAP^+ clusters. These two analyses resulted in the misclassification of 42.1% and 34.5%, respectively, of the patients belonging to the stem-like subtype into good-prognosis subtypes (**Supplementary Table 3**). Therefore, the expression patterns of epithelial genes are not sufficient to accurately define CRC molecular subtypes. We reached an equivalent conclusion from analyses of the classifications of De Sousa E Melo *et al.*[1] and Marisa *et al.*[3] (**Supplementary Table 3**).

As predicted from the data shown in **Figures 1** and **2**, the majority of genes included in the CAF cluster associated positively with cancer recurrence (HR > 1). We noticed that the expression levels of these genes were heterogeneous among patients belonging to good-prognosis subtypes (goblet-like, enterocyte-like, transit amplifying (TA)-like and inflammatory subtypes; **Fig. 3a**). Remarkably, we found a linear association between the average expression of this CAF gene cluster and risk of relapse after therapy in patients with good-prognosis subtypes (**Fig. 3b**). This finding thus enabled further stratification of patients belonging to good-prognosis subtypes into those at high and low risk of relapse depending on high (GP-HiC) or low (GP-LoC) CAF gene cluster expression (**Fig. 3c**). Of note, the disease-free survival progression of the GP-HiC subgroup across all four subtypes with good prognosis overlapped with that of the stem-like subtype with poor prognosis (**Fig. 3c**). Similarly, analysis of the molecular classification proposed by De Sousa E Melo *et al.*[1] showed that elevated expression of CAF gene–enriched clusters identified a subset of patients belonging to good-prognosis subtypes with disease behaving like that in the CCS3 poor-prognosis subtype (**Supplementary Fig. 1**). A similar trend was found in the classification of Marisa *et al.*[3], yet associations with relapse were of borderline significance in this analysis (**Supplementary Fig. 2**). From these observations, we conclude that elevated expression levels of the CAF gene program identify patients with poor prognosis across CRC molecular subtypes.

Analysis of poor-prognosis genes by immunohistochemistry

Eighty-three genes included in the poor-prognosis gene sets were common to at least two of the three molecular classifications (**Fig. 4a** and **Supplementary Table 4**). We interrogated the Human Protein Atlas database[7] and analyzed the tissue expression patterns of the proteins encoded by genes associated with poor prognosis. This database included immunohistochemistry data for antibodies against 76 of the 83 proteins encoded by these genes. Three antibodies (4%) showed no detectable staining in CRC samples, and we therefore focused on the 73 proteins displaying discrete immunohistochemical staining patterns (**Supplementary Fig. 3a** and **Supplementary Table 4**). Of these, only two antibodies (3%) marked exclusively epithelial tumor cells. In contrast, 31% stained solely the tumor stroma in the CRC samples, and the remainder (62%) labeled both stromal and epithelial tumor cells (**Supplementary Fig. 3a** and **Supplementary Table 4**). We performed a more detailed analysis for four genes associated with poor prognosis whose encoded proteins displayed either stromal-specific (CALD1 and POSTN) or mixed epithelial and stromal (FAP and IGFBP7) expression patterns in CRC according to the Human Protein Atlas database (**Fig. 4a**). Bioinformatic analyses indicated that the mRNAs of all four genes were upregulated in CAFs and other stromal cell populations in comparison to epithelial tumor cells (**Supplementary Fig. 3b**). In addition, antibodies selected against the four proteins detected single bands of the expected molecular weights by immunoblot in extracts of normal colon fibroblasts (**Fig. 5e**). We performed immunohistochemistry on 79 tissue samples belonging to patients with stage I, II or III CRC, which confirmed the *in silico* predictions. Overall staining intensity in stromal and epithelial compartments assessed under the microscope is shown in **Figure 4b**. POSTN and CALD1 were exclusively expressed in stromal cells, whereas IGFBP7 and FAP showed higher expression in the stroma than in epithelial tumor cells (**Fig. 4b**; examples in **Fig. 4c**). We assessed association between the intensity of either stromal or epithelial staining and disease-free survival after tumor resection in this patient cohort (**Fig. 4d**,**e** and **Supplementary Tables 5** and **6**). This analysis showed that elevated expression of CALD1, FAP or IGFPB7 in stromal cells predicted robustly shorter disease-free intervals, either as a linear (**Fig. 4d**) or categorized (**Fig. 4e**) variable. Of note, expression of these stromal proteins was a prognostic factor independent of the main clinical variables, including AJCC stage, treatment with adjuvant chemotherapy, age, tumor location and sex (**Supplementary Table 5**). In contrast, epithelial expression of IGFBP7 or FAP was not positively associated with disease recurrence (**Supplementary Table 6**). Overall, these results suggest that CRC subtypes with good and poor prognosis can be identified on the basis of the expression levels of a small subset of stromal proteins.

TGF-β signaling in CRC subtypes

TGF-β signaling was identified as one of the biological processes enriched in the poor-prognosis molecular CRC subtypes[1–3]. We confirmed

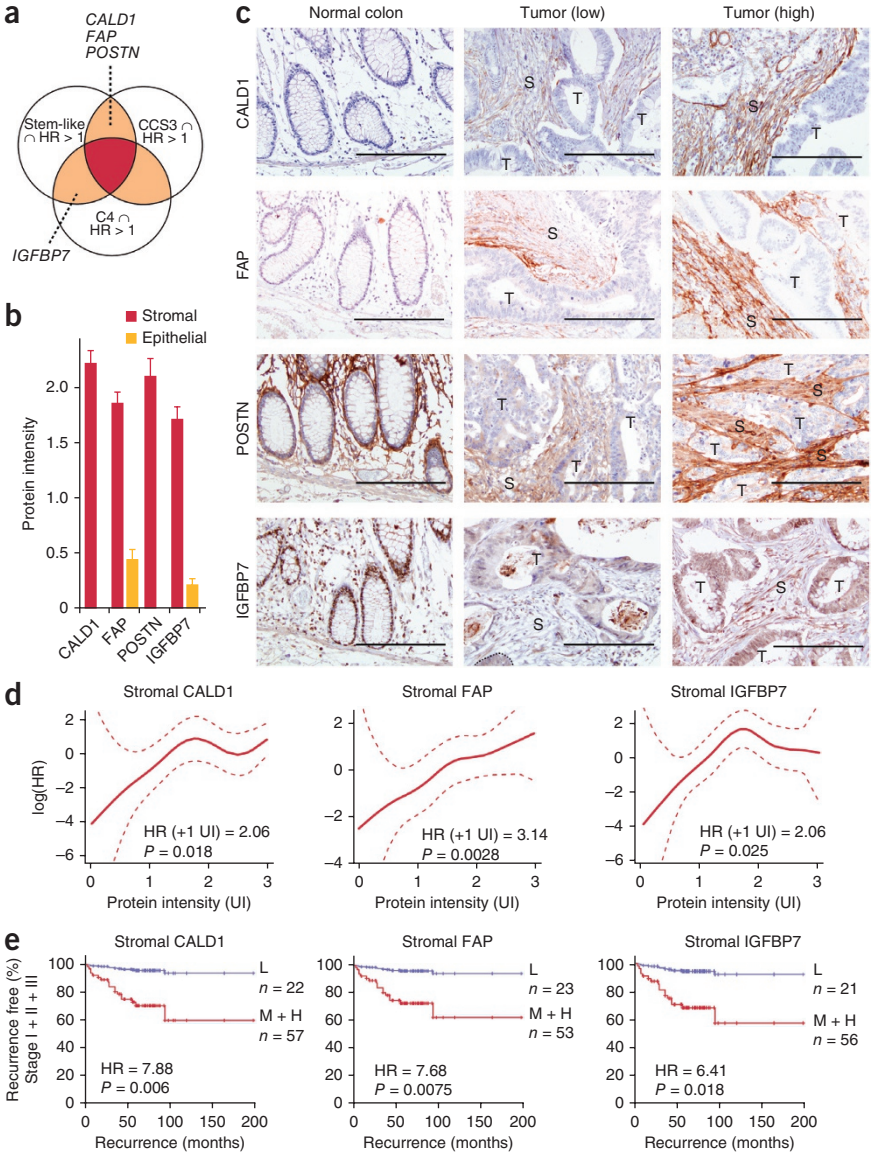

Figure 4 Identification of poor-prognosis patients by immunohistochemistry. (**a**) Venn diagram of the set of genes shared by the CCS3 and stem-like (*CALD1*, *FAP*, *POSTN*) and the stem-like and C4 (*IGFBP7*) poor-prognosis signatures. (**b**) Quantification of tumor microarray (TMA) analysis (79 patients) of common stromal genes in both tumor epithelium and stroma, with staining intensity for the protein scored from 0 to 3 as units of intensity (UI). Values are means ± s.e.m. (**c**) Immunostaining for CALD1, FAP, POSTN and IGFBP7 in representative human normal colon and tumor samples with lower (low) or higher (high) staining intensities (S, stroma; T, tumor). Scale bars, 200 µm. (**d**) HR (+1 UI; increase in recurrence risk for every unit of staining intensity) as a smoothened function of protein intensity based on TMA scores for CALD1, FAP and IGFBP7. Indicated are 95% confidence bands (red dashed lines). HRs for the continuous model and corresponding *P* values are indicated. (**e**) Kaplan-Meier curves showing the recurrence-free survival of patients with different levels of stromal protein intensity (L, low; M, medium; H, high) for CALD1, FAP and IGFBP7. HRs and corresponding *P* values are indicated (with the low-intensity category as reference; blue lines).

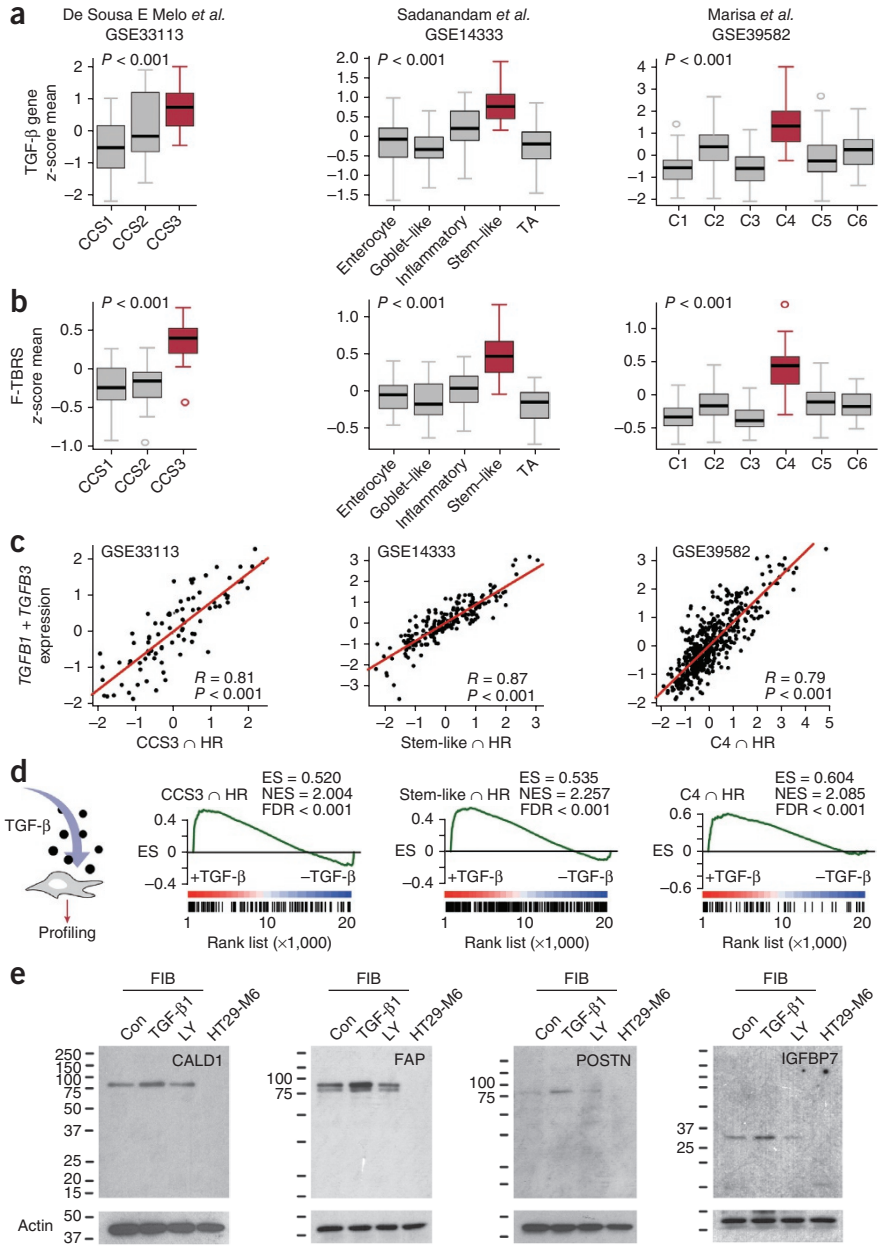

Figure 5 High levels of TGF-β and F-TBRS expression characterize poor-prognosis CRC subtypes. (**a**) Overall TGF-β gene mRNA expression levels in the three CRC patient data sets shown as *z*-score means computed for selected *TGFB1*, *TGFB2* and *TGFB3* probe sets ($P < 0.001$ for all comparisons involving the CCS3, stem-like and C4 subtypes; see **Supplementary Table 7** for pairwise comparisons). (**b**) *z*-score means of F-TBRS gene expression in the three data sets for each molecular subtype ($P < 0.001$ for all comparisons involving the CCS3, stem-like and C4 subtypes; see **Supplementary Table 7** for pairwise comparisons). Poor-prognosis groups for each data set are depicted in red in **a** and **b**. Whiskers represent upper and lower quartiles. (**c**) Correlation of *TGFB1* and *TGFB3* mRNA levels with poor-prognosis signatures in the three data sets. Spearman's correlations (R) and P values are indicated. (**d**) GSEA of gene sets associated with poor prognosis in control fibroblasts or ones treated with TGF-β (ES, enrichment score; NES, normalized enrichment score). (**e**) Top, levels of CALD1, FAP, POSTN and IGFBP7 proteins in fibroblasts (FIB) untreated (Con) or treated with recombinant TGF-β1 or LY2157299 (LY) and in HT29-M6 cells. Protein blots were performed with the antibodies used for the immunohistochemistry analyses in **Figure 4**. Bottom, actin protein levels as normalization controls.

elevated expression of overall TGF-β levels (**Fig. 5a**) and particularly of *TGFB1* and *TGFB3* in the CCS3, stem-like and C4 subtypes (**Supplementary Fig. 4a** and **Supplementary Table 7**). We have recently reported that genes upregulated by TGF-β in stromal cells are robust predictors of cancer recurrence and metastasis in CRC[6]. Several of the stromal TGF-β response signatures[6] (TBRSs; see **Supplementary Table 8** for gene lists) predicted disease relapse in the three CRC patient cohorts analyzed, independently of the main clinical parameters (**Supplementary Table 9**). We thus explored the association of these TBRSs with the proposed molecular classifications of CRC. Our analyses also showed that the genes induced by TGF-β in normal colonic fibroblasts (F-TBRSs) were upregulated in all poor-prognosis CRC subtypes (**Fig. 5b** and **Supplementary Table 7**). We obtained equivalent results using TBRSs derived from macrophages and T cells (Ma-TBRSs and T-TBRSs, respectively) (**Supplementary Fig. 4b,c**).

The average expression of the CCS3, stem-like and C4 gene sets associated with poor prognosis was tightly correlated with combined expression of *TGFB1* and *TGFB3* in the three CRC patient cohorts (**Fig. 5c**). To investigate whether these gene sets are directly regulated by TGF-β, we profiled the global expression of normal colon fibroblasts before and after induction with TGF-β. GSEA indicated that TGF-β treatment elevated the expression of the gene sets associated with poor prognosis in fibroblasts (**Fig. 5d**). We also analyzed expression of the markers used in **Figure 4** to identify poor-prognosis patients by immunohistochemistry. The levels of these proteins (CALD1, POSTN, FAP and IGFBP7) were upregulated by TGF-β in colon fibroblasts (**Fig. 5e**). All together, these data show that TGF-β signaling in stromal cells is a defining feature of poor-prognosis CRC subtypes.

Analysis of TGF-β signaling in the tumor stroma

The above correlations prompted us to analyze the interactions of CRC cells with their microenvironment and specifically with fibroblasts through TGF-β (**Fig. 6**). To functionally dissect this effect without the interference of TGF-β signaling in epithelial cancer cells, we used the CRC cell line HT29-M6, which carries mutations in *SMAD4* that inactivate the response to TGF-β (ref. 6). We enforced TGF-β signaling in the tumor microenvironment by engineering HT29-M6 cells to secrete active TGF-β1. These HT29-M6$^{TGF\text{-}\beta}$ cells displayed no autocrine responses to TGF-β1 secretion (**Supplementary Fig. 5a,b**)[6]. Control HT29-M6 cells or HT29-M6$^{TGF\text{-}\beta}$ cells were not tumorigenic upon subcutaneous injection into immunodeficient mice at low numbers (**Fig. 6a**). In contrast, coinoculation of HT29-M6$^{TGF\text{-}\beta}$ cells

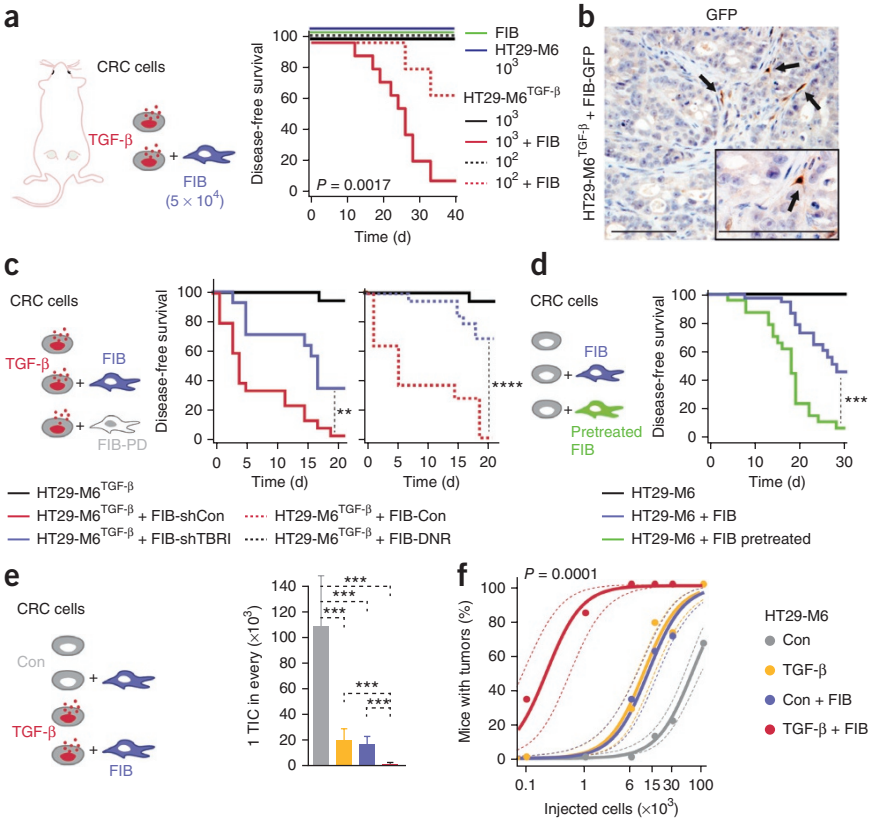

Figure 6 TGF-β–activated fibroblasts promote tumor initiation. (**a**) Disease-free survival plots for mice coinoculated with HT29-M6$^{TGF\text{-}\beta}$ cells (1,000 cells, $n = 12$ inoculations; 100 cells, $n = 6$ inoculations) and fibroblasts (FIB; 5×10^4 cells) in comparison to mice inoculated with fibroblasts alone (5×10^4 cells, $n = 6$ tumor cell inoculations), control HT29-M6 cells alone (1,000 cells, $n = 6$ inoculations) or HT29-M6$^{TGF\text{-}\beta}$ cells alone (1,000 cells, $n = 20$ inoculations; 100 cells, $n = 6$ inoculations). (**b**) GFP immunostaining of a representative tumor derived from HT29-M6$^{TGF\text{-}\beta}$ cells coinoculated with eGFP-expressing fibroblasts. Arrows point to GFP$^+$ fibroblasts. Inset, magnification of a GFP$^+$ fibroblast. Scale bars, 100 µm. (**c**) Disease-free survival plots for mice after coinoculation of 6×10^3 HT29-M6$^{TGF\text{-}\beta}$ cells together with control fibroblasts (FIB-shCon (expressing control shRNA), $n = 20$ inoculations; FIB-Con, $n = 14$) or TGF-β pathway–defective fibroblasts (FIB-PD) in comparison with HT29-M6$^{TGF\text{-}\beta}$ cells alone ($n = 14$). $^{**}P < 0.01$, $^{****}P < 0.0001$. FIB-PD could be either fibroblasts expressing shRNA to *TGFBR1* (FIB-shTBRI; $n = 14$) or fibroblasts expressing a dominant-negative *TGFBR2* (FIB-DNR; $n = 20$). (**d**) Disease-free survival plots of mice coinoculated with HT29-M6 cells (15×10^3) and 5×10^4 fibroblasts prestimulated over 4 d with recombinant TGF-β1 ($n = 24$ inoculations) or with untreated fibroblasts ($n = 40$) and of mice inoculated with HT29-M6 cells alone ($n = 14$). $^{***}P < 0.0001$. (**e**) Estimated frequency of TICs calculated using extreme limiting dilution analysis (ELDA) (Online Methods). $^{***}P < 0.001$ for comparisons against the TGF-β + FIB group; 95% confidence intervals are indicated. (**f**) Solid lines indicate the percentage of mice developing tumors according to estimations provided by a logistic regression model (Online Methods), including the interaction term (likelihood-ratio test, $P = 0.0001$). Dashed lines are 95% confidence bands. Detailed information on conditions and numbers of mice for **e** and **f** are provided in **Supplementary Figure 7c**.

with normal colonic fibroblasts very dramatically enhanced the frequency of xenograft formation (**Fig. 6a**). We analyzed the fate of coinoculated fibroblasts by engineering them to constitutively express eGFP. At experimental end points, we detected a small proportion of eGFP-labeled cells in HT29-M6$^{TGF\text{-}\beta}$–derived xenografts, which constituted 1–5% of all stroma cells ($n = 18$; example in **Fig. 6b**). Therefore, coinoculated fibroblasts enhanced the capacity of HT29-M6$^{TGF\text{-}\beta}$ cells to engraft in recipient mice, yet they did not contribute extensively to the formation of the tumor stroma.

Knockdown of *TGFBR1* or expression of dominant-negative *TGFBR2* in fibroblasts blocked the compounded effect of TGF-β on xenograft formation (**Fig. 6c**). We obtained equivalent results using KM12L4a$^{TGF\text{-}\beta}$ cells (**Supplementary Fig. 5c**), a CRC cell line that carries biallelic *TGFBR2* loss-of-function mutations and also does not respond to TGF-β (**Supplementary Fig. 5a**,**b**)[6,8]. Notably, addition of wild-type or TGF-β pathway–defective fibroblasts did not modify xenograft growth rates (**Supplementary Fig. 6a**,**b**). Moreover, tumors displayed equivalent histological traits and were formed by similar proportions of epithelial and stromal cells (**Supplementary Fig. 6c**–**e**). We also coinoculated *TGFBR1*-knockdown fibroblasts with control HT29-M6 cells. In these experiments, both fibroblasts with control short hairpin RNA (shRNA) and those lacking a TGF-β response showed a minor capacity to support tumor formation (**Supplementary Fig. 7a**). In a second approach, we coinoculated HT29-M6 control cells with normal colon fibroblasts that had been pretreated *in vitro* with TGF-β for 96 h before inoculation (**Fig. 6d**). TGF-β upregulated the expression of genes associated with poor prognosis in these fibroblasts (**Fig. 5d**,**e**) while it slowed down their proliferation (**Supplementary Fig. 7b**). Fibroblasts prestimulated with TGF-β reduced tumor latency and increased engraftment of control HT29-M6 cells (**Fig. 6d**). In this setting, transient activation of fibroblasts before inoculation suggests that TGF-β–activated stroma operates at the initial phase of tumor formation.

It has been shown that the frequency of tumor-initiating cells (TICs) is a surrogate for cancer stem cell activity[9,10]. This tumor cell population is believed to mediate disease relapse and metastasis[11], which are the major drivers of poor prognosis in CRC. We measured the impact of TGF-β–activated fibroblasts on TIC frequency by performing limiting dilution assays (**Supplementary Fig. 7c**). Groups of mice were injected subcutaneously with control HT29-M6 or HT29-M6$^{TGF\text{-}\beta}$ cells at different cell concentrations with or without normal colon fibroblasts, and we assessed tumor initiation. We calculated that control HT29-M6 cells contained 1 TIC for every 1.08×10^5 cells (**Fig. 6e** and **Supplementary Fig. 7d**). Secretion of TGF-β increased TIC frequency by fivefold over the level of control cells upon inoculation, which was roughly equivalent to the frequency obtained upon coinoculation of control HT29-M6 cells with normal colon fibroblasts (**Fig. 6e** and **Supplementary Fig. 7d**). Remarkably, coinoculation of HT29-M6$^{TGF\text{-}\beta}$ cells with fibroblasts increased the TIC frequency to 1 for every 493 cells, which represents a 200-fold increase in comparison to control HT29-M6 cells alone (**Fig. 6e** and **Supplementary Fig. 7d**). These findings show a synergistic as opposed to an additive effect of TGF-β and fibroblasts on TIC frequency ($P = 0.0001$; **Fig. 6f**). In addition, we calculated tumor size

over time starting from the day that tumors were palpable, yet we did not find significant differences in the growth rates of xenografts arising under the four experimental conditions (**Supplementary Fig. 5d**). We thus conclude that TGF-β signaling in fibroblasts acts by specifically enhancing the tumor-initiating potential of CRC cells.

Analysis of TGF-β signaling in epithelial tumor cells

In addition to regulating gene programs in stromal cells, TGF-β might possibly also control the behavior of epithelial CRC cells that carry a wild-type pathway. TGF-β is a well-established inducer of epithelial-to-mesenchymal transition (EMT)[12], and it therefore could elevate the expression of mesenchymal genes in CRCs, contributing to the acquisition of a more aggressive phenotype. An important limitation in testing this hypothesis is the fact that virtually all CRC cell lines available have lost the response to TGF-β as a consequence of inactivating mutations in TGF-β pathway components[13–15]. To overcome this obstacle, we took advantage of a recently developed methodology that enables *in vitro* propagation of primary CRCs as three-dimensional primary cultures (**Fig. 7a**)[16,17]. Under these conditions, CRC cells grew indefinitely as epithelial structures lacking stromal cells called patient-derived tumor organoids (PDOs) (**Fig. 7b** and data not shown). We expanded organoids from eight primary CRCs. The main clinicopathological features of the CRC of origin for each tumor organoid culture are detailed in **Supplementary Table 10**. Exome sequencing analysis confirmed that these organoids had distinct combinations of mutations in the main driver pathways (**Supplementary Table 10**). Organoid technology allowed us to detect primary CRC samples that displayed a TGF-β response. In four tumor organoids (PDO2, PDO3, PDO4 and PDO6), addition of TGF-β induced robust cytostasis that included decreased organoid-forming capacity (**Fig. 7c**) and reduced growth rates (**Supplementary Fig. 8**). This response was blocked by LY2157299, a TGF-βR1–specific inhibitor[18]. In these four tumor organoids, treatment with TGF-β upregulated the CDK4 and CDK6 inhibitors *CDKN1A*, *CDKN2A* and *CDKN2B*, whereas it downregulated *MKI67* and *MYC* (**Fig. 7e**), indicating a tumor-suppressor response similar to that previously shown for keratinocytes[19,20]. Exome sequencing of PDO5 showed no apparent alteration in the expression of genes encoding TGF-β pathway components, yet this organoid neither halted proliferation nor upregulated expression of cell cycle inhibitors in the presence of TGF-β (**Fig. 7c–e** and **Supplementary Fig. 8**). PDO7 and PDO8 displayed *SMAD4* loss-of-function alterations in a homozygous state (**Supplementary Table 10**), and TGF-β did not modify their growth rates (**Fig. 7c–e** and **Supplementary Fig. 8**). PDO1 carried *TGFBR2* loss-of-function mutations in both alleles (**Supplementary Table 10**) and was completely insensitive to the action of TGF-β (**Fig. 7c–e** and **Supplementary Fig. 8**). Of note, even after 1 week in the presence of TGF-β, none of the tumor organoids acquired mesenchymal traits and they all remained epithelial, as shown by the maintenance of a cuboidal morphology and basolateral E-cadherin staining (**Fig. 7d**). There were minimal changes in the expression of epithelial markers (*CDH1*, *CLDN1* and *EPCAM*) or EMT master genes (*SNAIL*, *TWIST1* and *ZEB1*)[12] (**Fig. 7e**). Therefore, CRCs appear to be largely resilient to the induction of EMT by TGF-β, which nonetheless triggered a robust cytostatic response in tumor cells with a wild-type TGF-β pathway.

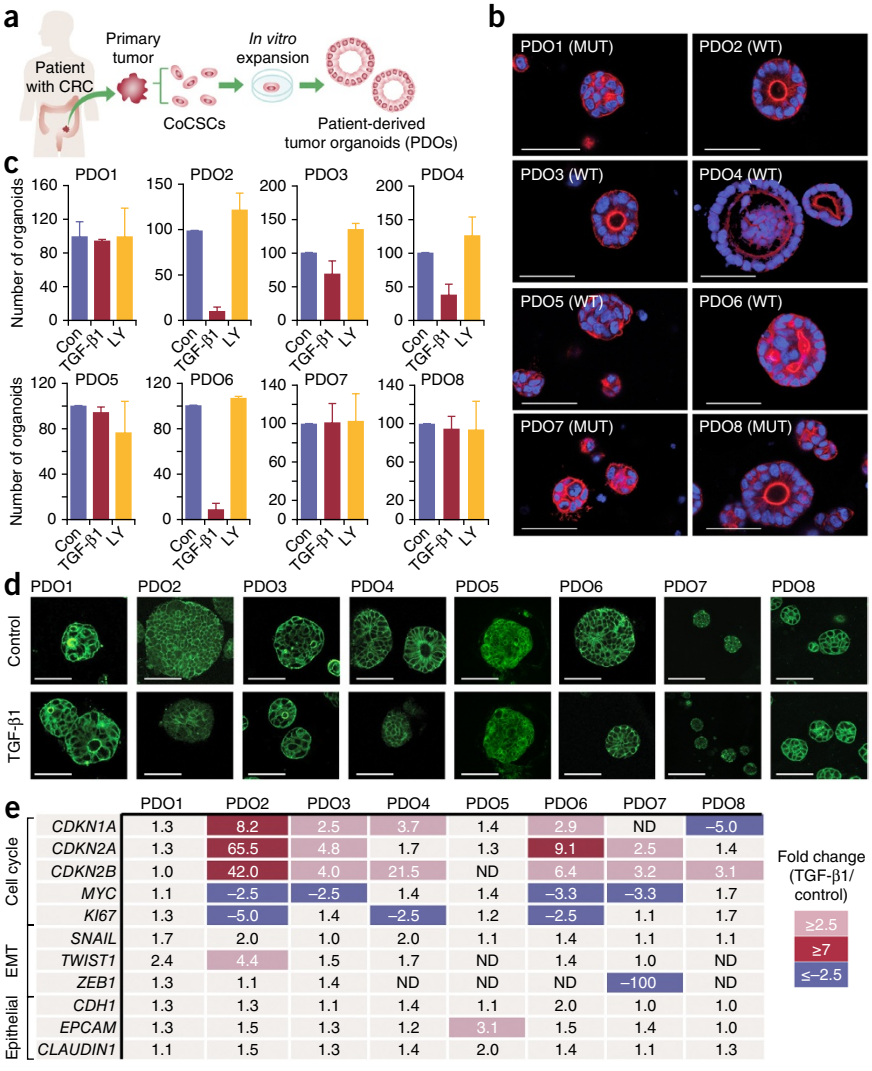

Figure 7 TGF-β induces a cytostatic response but not EMT in tumor organoids with a wild-type TGF-β pathway. (**a**) CRC stem cells (CoCSCs) were isolated from fresh patient biopsies and cultured in Matrigel and niche factors. Under *in vitro* optimal culture conditions, CoCSCs consistently form dysplastic organoid structures. (**b**) Confocal images for phalloidin (red) and nuclei (blue) of organoids derived from eight patients. TGF-β pathway status is indicated: WT, wild type; MUT, mutant. Scale bars, 50 µm. (**c**) Tumor organoid formation capacity in the presence or absence of recombinant TGF-β1 protein (5 ng/ml) or LY2157299 TGF-βRI inhibitor (LY; 1 µM). Values are means ± s.d. ($n \geq 6$ independent organoid cultures). (**d**) Confocal images for E-cadherin (green) of organoids treated or untreated with recombinant TGF-β1 for 7 d. Scale bars, 50 µm. (**e**) Quantitative RT-PCR (qRT-PCR) analysis of cell cycle, EMT and epithelial genes in organoids treated or untreated for 7 d with recombinant TGF-β1. Data are normalized to *PPP1CA* and are presented as fold changes versus untreated cells ($n \geq 6$ independent tumor organoid cultures). ND, not detectable.

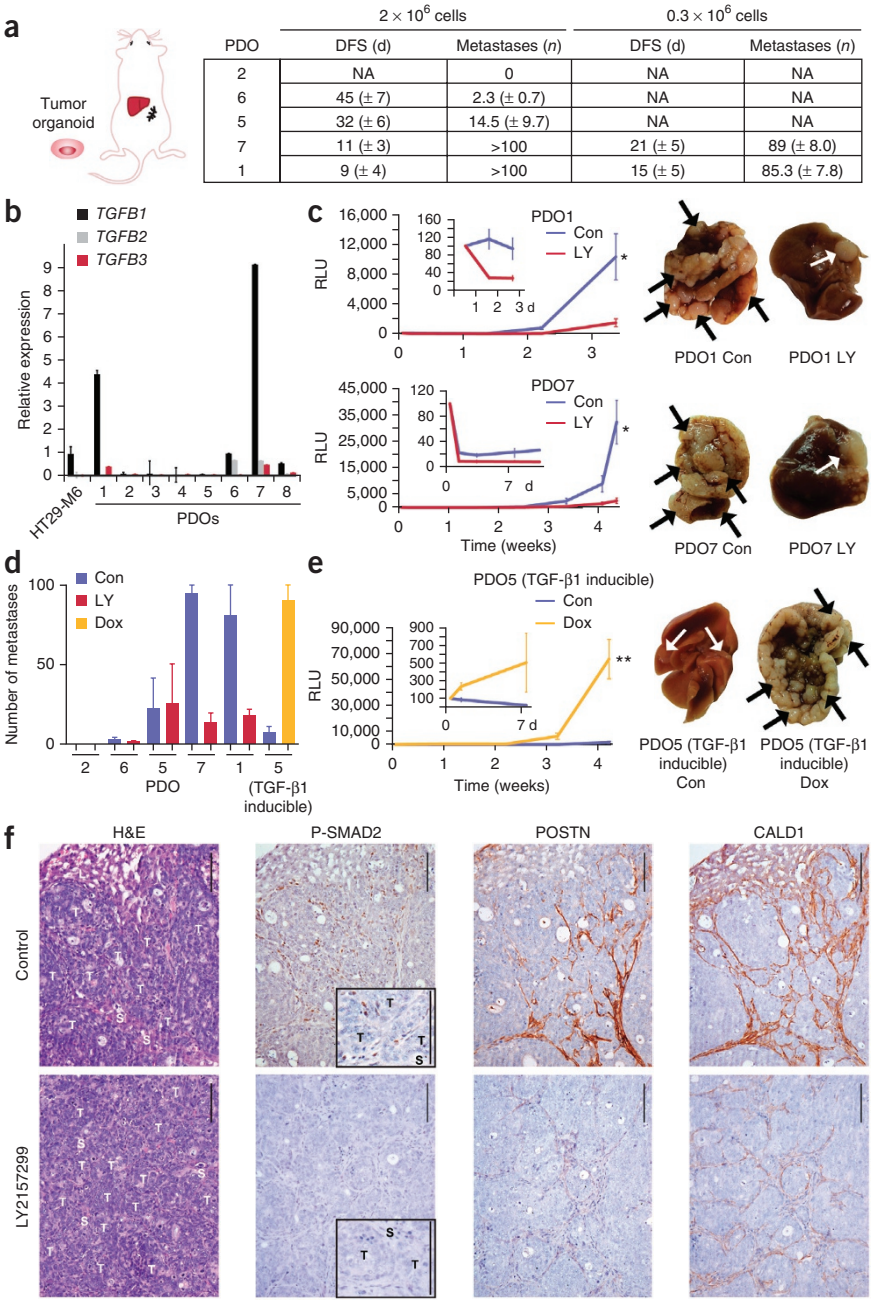

Figure 8 TGF-β–activated stroma promotes tumor initiation, an effect reversed by its chemical inhibition with LY2157299. (**a**) Intrasplenic injection of PDO2, PDO6, PDO5, PDO7 or PDO1 (2×10^6 cells, $n = 5$ mice; 0.3×10^6 cells, $n = 5$ mice). The table shows disease-free survival (DFS) and the incidence of distant metastases after inoculation (NA, not available). (**b**) Relative expression levels of the *TGFB1*, *TGFB2* and *TGFB3* genes measured by qRT-PCR in the indicated PDOs cultured *in vitro*. Values are means ± s.d.; $n = 3$ cultures per condition. (**c**,**e**) Normalized bioluminescence (RLU) over time from mice with intrasplenic inoculation with cells from PDO1 or PDO7 and treated with LY2157299 (LY; PDO1, $n = 5$ mice; PDO7, $n = 6$ mice) or with vehicle (Con; PDO1, $n = 5$ mice; PDO7, $n = 6$ mice) (**c**) or with PDO5 TGF-β1–inducible cells (2×10^6 cells) and treated with doxycycline (Dox; $n = 6$ mice) or sucrose (Con; $n = 6$ mice) (**e**). Signal intensities were normalized to day 0, which was set arbitrarily to 100. Values are means ± s.e.m. (*$P < 0.05$, **$P < 0.01$, Student's t test). Macroscopic pictures of representative livers are shown. Arrows point to metastatic nodules. (**d**) Quantification of liver metastases at the time of sacrifice in **c** and **e**. Values are means ± s.e.m. (**f**) Hematoxylin and eosin (H&E) staining and immunostaining for phosphorylated SMAD2 (P-SMAD2), POSTN and CALD1 in representative liver metastases obtained from mice treated with carrier (top) or LY2157299 (bottom) (S, stroma; T, tumor). Insets, magnifications of stromal positive or negative staining for phosphorylated SMAD2. Scale bars, 100 μm.

Use of TGF-β signaling inhibitors to block metastasis

The ultimate stage of CRC progression is metastatic disease, which is caused by disseminated cancer cells that hold the capacity to initiate a new tumor within a foreign tissue[11], mainly the liver and lungs. To assess the ability of the different tumor organoids to generate metastases, we inoculated them as dissociated cells through the spleen of immunodeficient mice (**Fig. 8**). Despite thousands of tumor cells entering the portal vein and reaching the liver within minutes of inoculation in this experimental model, most tumor organoids produced few or no metastases (**Fig. 8a**). The exceptions were PDO1 and PDO7, both of which expressed elevated TGF-β levels (**Fig. 8b**) and were highly metastatic (**Fig. 8a**; pictured in **Fig. 8c**). LY2157299 (ref. 18) is a TGF-βR1–specific inhibitor that is currently in phase 2 clinical trials for the treatment of hepatocellular carcinoma[21]. Treatment of mice with this drug reduced the number of metastases formed by these two tumor organoids (**Fig. 8c**,**d**). Reciprocally, enforced secretion of active TGF-β1 by PDO5 massively increased the metastatic burden (**Fig. 8d**,**e**). Analysis of the kinetics of metastasis using luciferase-labeled organoids indicated that LY2157299 acted by blocking the capacity of tumor cells to thrive in the liver over the colonization phase (during the first few days after inoculation) (insets in **Fig. 8c**). Of note, PDO1 and PDO7 exhibited biallelic inactivation of the *TGFBR2* and *SMAD4* genes, respectively (**Supplementary Table 10**), and both were insensitive to TGF-β (**Fig. 7** and **Supplementary Fig. 8**). Therefore, reduced metastatic capacity caused by LY2157299 treatment in these tumors can only be attributed to inhibition of TGF-β signaling in the tumor microenvironment. Indeed, we observed decreased staining for phosphorylation of SMAD2 and reduced expression of the TGF-β target genes *CALD1* and *POSTN* in the stroma of the liver metastases upon LY2157299 treatment (**Fig. 8f**). Remarkably, despite the *in vitro* cytostatic effect imposed by TGF-β in tumor organoids with an intact TGF-β pathway, we did not observe enhancement of metastasis *in vivo* upon treatment of mice with LY2157299 (**Fig. 8d**; PDO2 and PDO6). LY2157299 promoted neither initiation nor growth of TGF-β–responsive tumor organoids upon subcutaneous inoculation in mice (**Supplementary Fig. 9**).

DISCUSSION

Recently developed molecular classification systems offer great opportunities to improve the stratification and treatment of patients with CRC[1–3]. The finding that poor-prognosis subtypes are characterized by elevated expression of mesenchymal genes has led to the speculation that EMT might be responsible for their aggressiveness[1–3]. Our analyses show that elevated expression of mesenchymal genes associated with poor prognosis in CRC samples is mainly contributed by tumor-associated stromal cells rather than by epithelial tumor cells. This discovery does not invalidate the possibility that individual tumor cells undergo EMT, particularly at invasion fronts. Yet, it argues against generalized expression of a mesenchymal poor-prognosis gene program in epithelial CRC cells. Sadanandam *et al.* proposed that the poor-prognosis subtype is characterized by upregulation of a gene program similar to that of intestinal stem cells. This conclusion was based on the expression of genes upregulated in microdissected crypt bottoms compared to crypt tops in stem cell–like tumors[22]. However, the signature of crypt bottoms used in this study also contained mesenchymal genes expressed by pericryptal fibroblasts, the presumptive stem cell niche cells, which were microdissected together with the epithelial cells[22]. Our bioinformatic and immunohistochemistry analyses indicate that several genes associated with poor prognosis are expressed by stromal cells in the normal mucosa (**Figs. 1b** and **4c**), which may suggest that the CRC microenvironment reproduces some features of the crypt niche. We propose that, to a large extent, the various molecular classifications published so far distinguish CRCs of good and poor prognosis on the basis of distinct features of the tumor stroma. Notably, even those patients classified as having good-prognosis subtypes that expressed elevated levels of the CAF program displayed increased risk of relapse and disease-free survival intervals, similar to those of patients with poor-prognosis subtypes. These observations could help unify the multiple CRC subtypes identified in each molecular classification. Our conclusions are further supported by various studies that link elevated expression of particular stromal-specific genes with poor outcome in CRC[23–27].

We showed that a common feature of all poor-prognosis subtypes is elevated TGF-β expression. Our functional data underscore a dual role for the TGF-β pathway in epithelial versus stromal tumor cells. TGF-β signaling slows down proliferation of epithelial CRC cells without triggering EMT. A large fraction of CRCs avert this block by losing sensitivity to TGF-β via mutations and possibly through additional mechanisms. On the other hand, TGF-β in CAFs boosts the tumor-initiating capacity of CRC cells, a property that is connected with increased metastatic potential and the ability to regenerate disease after therapy[11]. Indeed, our data suggest that TGF-β target genes are an important fraction of the genes sets associated with poor prognosis. We show that pharmacological inhibition of TGF-β signaling in the tumor microenvironment prevents metastasis formation by patient-derived tumor organoids. These findings corroborate our previous results demonstrating a dependency of TGF-β signaling in stromal cells during metastasis[6]. Our work further warrants the development of anti–TGF-β therapies for the treatment of poor-prognosis CRCs. The observation that LY2157299 did not boost xenograft growth even in TGF-β–responsive tumor organoids indicates that the use of this inhibitor may be safe for a wide range of patients with CRC.

URLs. French Ligue Nationale Contre le Cancer, http://cit.ligue-cancer.net/; Fiji Trainable Weka Segmentation, http://fiji.sc/Trainable_Weka_Segmentation; custom macro for image processing, http://adm.irbbarcelona.org/image-j-fiji; Rsamtools: binary alignment (BAM), FASTA, variant call (BCF) and tabix file import (R package version 1.18.2), http://bioconductor.org/packages/release/bioc/html/Rsamtools.html; BSgenome.Hsapiens.UCSC.hg19: full genome sequences for *Homo sapiens* (UCSC version hg19; R package version 1.4.0), http://www.bioconductor.org/packages/release/data/annotation/html/BSgenome.Hsapiens.UCSC.hg19.html.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Exome sequencing data for tumor organoids have been deposited in the European Nucleotide Archive (ENA) under accession PRJEB7932. Expression data for colonic fibroblasts treated with TGF-β1 for 8 h are accessible under Gene Expression Omnibus (GEO) accession GSE64192.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank G. Stassi (University of Palermo) for providing PDO7 and PDO8, L. Wakefield (US National Cancer Institute) for providing the plasmid encoding DNR, I. Joval for assistance in mounting the figures, M. Virtudes Cespedes and R. Mangues (IIB Sant Pau) for logistic support with CRC samples, and all members of the Batlle laboratory for support and discussions We are grateful for the excellent assistance of the IRB Barcelona core facilities for Histology, Functional Genomics and Advanced Digital Micropscopy. D.V.F.T. holds a Juan de la Cierva postdoctoral fellowship, from the Spanish Ministry of Economy and Competitiveness, and E.L. holds a fellowship from Fundación Olga Torres and Asociación Española contra el Cáncer (AECC). This work has been supported by grants from the Doctor Josef Steiner Foundation, AECC, Red Temática de Investigación Cooperativa en Cáncer, Instituto de Salud Carlos III (RTICC:RD12/0036/0024) and grant SAF2011-27068, the latter two from the Spanish Ministry of Economy and Competitiveness, and by 'Xarxa de Bancs de Tumors' sponsored by Pla Director d'Oncologia de Catalunya (XBTC).

AUTHOR CONTRIBUTIONS

A.C., E.L. and E.E. designed, planned and performed experiments and analyzed the results. X.H.-M. and S.P.-P. provided crucial assistance with *in vivo* experiments. M.S. performed immunohistochemistry. M.I. scored the tumor microarray. S.T. developed an algorithm to quantify xenograft images. D.V.F.T., C.C., C.B. and C.M. performed experiments and/or analyzed results. D.V.F.T., D.B. and A.R. synthesized the LY2157299 inhibitor. A.B.-L., C.S.-O.A. and D.R. designed and performed biostatistical analyses. E.B. conceptualized and supervised the project, analyzed results and wrote the manuscript, with the assistance of E.S.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. De Sousa E Melo, F. *et al.* Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. *Nat. Med.* **19**, 614–618 (2013).
2. Sadanandam, A. *et al.* A colorectal cancer classification system that associates cellular phenotype and responses to therapy. *Nat. Med.* **19**, 619–625 (2013).
3. Marisa, L. *et al.* Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med.* **10**, e1001453 (2013).
4. Sadanandam, A. *et al.* Reconciliation of classification systems defining molecular subtypes of colorectal cancer: interrelationships and clinical implications. *Cell Cycle* **13**, 353–357 (2014).
5. Nishida, N. *et al.* Microarray analysis of colorectal cancer stromal tissue reveals upregulation of two oncogenic miRNA clusters. *Clin. Cancer Res.* **18**, 3054–3070 (2012).
6. Calon, A. *et al.* Dependency of colorectal cancer on a TGF-β–driven program in stromal cells for metastasis initiation. *Cancer Cell* **22**, 571–584 (2012).
7. Uhlen, M. *et al.* Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **28**, 1248–1250 (2010).
8. Mouradov, D. *et al.* Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. *Cancer Res.* **74**, 3238–3247 (2014).

9. O'Brien, C.A., Pollett, A., Gallinger, S. & Dick, J.E. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* **445**, 106–110 (2007).
10. Ricci-Vitiani, L. *et al.* Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111–115 (2007).
11. Oskarsson, T., Batlle, E. & Massague, J. Metastatic stem cells: sources, niches, and vital pathways. *Cell Stem Cell* **14**, 306–321 (2014).
12. Thiery, J.P., Acloque, H., Huang, R.Y. & Nieto, M.A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
13. Markowitz, S.D. & Bertagnolli, M.M. Molecular origins of cancer: molecular basis of colorectal cancer. *N. Engl. J. Med.* **361**, 2449–2460 (2009).
14. Grady, W.M. & Markowitz, S.D. Genetic and epigenetic alterations in colon cancer. *Annu. Rev. Genomics Hum. Genet.* **3**, 101–128 (2002).
15. Markowitz, S. *et al.* Inactivation of the type II TGF-β receptor in colon cancer cells with microsatellite instability. *Science* **268**, 1336–1338 (1995).
16. Jung, P. *et al.* Isolation and *in vitro* expansion of human colonic stem cells. *Nat. Med.* **17**, 1225–1227 (2011).
17. Sato, T. *et al.* Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762–1772 (2011).
18. Bueno, L. *et al.* Semi-mechanistic modelling of the tumour growth inhibitory effects of LY2157299, a new type I receptor TGF-β kinase antagonist, in mice. *Eur. J. Cancer* **44**, 142–150 (2008).
19. Seoane, J., Le, H.V. & Massague, J. Myc suppression of the $p21^{Cip1}$ Cdk inhibitor influences the outcome of the p53 response to DNA damage. *Nature* **419**, 729–734 (2002).
20. Seoane, J. *et al.* TGFβ influences Myc, Miz-1 and Smad to control the CDK inhibitor $p15^{INK4b}$. *Nat. Cell Biol.* **3**, 400–408 (2001).
21. Giannelli, G., Villa, E. & Lahn, M. Transforming growth factor-β as a therapeutic target in hepatocellular carcinoma. *Cancer Res.* **74**, 1890–1894 (2014).
22. Kosinski, C. *et al.* Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. *Proc. Natl. Acad. Sci. USA* **104**, 15418–15423 (2007).
23. Berdiel-Acer, M. *et al.* A 5-gene classifier from the carcinoma-associated fibroblast transcriptomic profile and clinical outcome in colorectal cancer. *Oncotarget.* **5**, 6437–6452 (2014).
24. O'Shannessy, D.J. *et al.* Influence of tumor microenvironment on prognosis in colorectal cancer: tissue architecture–dependent signature of endosialin (TEM-1) and associated proteins. *Oncotarget.* **5**, 3983–3995 (2014).
25. Francí, C. *et al.* Snail1 protein in the stroma as a new putative prognosis marker for colon tumours. *PLoS ONE* **4**, e5595 (2009).
26. Ngan, C.Y. *et al.* Quantitative evaluation of vimentin expression in tumour stroma of colorectal cancer. *Br. J. Cancer* **96**, 986–992 (2007).
27. Calon, A., Tauriello, D.V. & Batlle, E. TGF-β in CAF-mediated tumor growth and metastasis. *Semin. Cancer Biol.* **25**, 15–22 (2014).

ONLINE METHODS

Description of the CRC transcriptomic data sets. We used three Affymetrix data sets publicly available in NCBI GEO[28] with available clinical data and follow-up information: GSE14333 (ref. 29), GSE33113 (ref. 30) and GSE39582 (ref. 3). GSE14333 contains a pool of 290 patients with CRC treated at 2 different hospitals: the Peter MacCallum Cancer Center (Australia) and the H. Lee Moffitt Cancer Center (United States). This data set was used to define the molecular classification by Sadanandam *et al.*[2]. The GSE33113 data set includes disease-free survival information for 90 patients with AJCC stage II disease collected at the Academic Medical Center in Amsterdam (the Netherlands), and it was used to define the molecular classification by De Sousa E Melo *et al.*[1]. Finally, GSE39582 includes expression and clinical data for 566 patients with CRC collected for the Cartes d'Identité des Tumeurs (CIT) program, from the French Ligue Nationale Contre le Cancer. This data set was used to build the CRC classification by Marisa *et al.*[3]. All patients with stage 1, 2 and 3 (but not stage 4) disease in each cohort were taken into consideration for analyses. See the **Supplementary Note** for details about data processing.

Expression profile of poor-prognosis molecular subtypes. To characterize each poor-prognosis subtype (CCS3, stem-like and C4), we defined a transcriptomic profile on the same data originally used to define their corresponding classification (GSE33113, GSE14333 and GSE39582, respectively). Genes included in these profiles were required to meet stringent statistical criteria: at least 2-fold overexpression in the poor-prognosis subtype and FDR <0.01.

Association of poor-prognosis subtypes with clinical outcome. Association of gene expression with relapse was assessed in the CRC transcriptomic data sets using a univariate Cox proportional hazards model. For the Sadanandam *et al.* signature, an association with recurrence was reported only in patients not treated with adjuvant chemotherapy. Therefore, this analysis was restricted to this group of patients in data set GSE14333. When assessing overlap with subtype profiles, risk-associated genes (HR > 1) showing a *P* value less than 0.01 were considered. A gene was located in the intersection when some of its probe sets were found in both profiles. The significance of the overlapping genes was assessed assuming a hypergeometric distribution, and the median and 2.5% and 97.5% percentiles of FDR were computed in each case, assuming that occurrences in the intersection were due to randomness.

Expression of genes associated with the poor-prognosis subtypes in subpopulations. Three GEO data sets were used to characterize the subtype gene profiles according to specific gene expression in tumoral cell subpopulations: GSE39395 (ref. 6), GSE39396 (ref. 6) and GSE35602 (ref. 5). In GSE39395 and GSE39396 and as we described previously[6], FACS was used to separate the following populations from 14 fresh CRC samples: $CD45^+EpCAM^-CD31^-FAP^-$, $CD45^-EpCAM^+CD31^-FAP^-$, $CD45^-EpCAM^-CD31^+FAP^-$ and $CD45^-EpCAM^-CD31^-FAP^+$. GSE35602 contains transcriptomic data for epithelial and stromal cells microdissected from 13 CRC tissue samples and 4 adjacent morphologically normal colorectal mucosae (>5 cm from the tumor). The signatures derived from the poor-prognosis subtypes were summarized and evaluated in these samples. Population groups were then compared using Kruskal and Mann-Whitney tests. In addition, we explored the sensitivity of these results to the thresholds used to define the subtypes and recurrence signatures (see the **Supplementary Note** for details).

Contribution of genes expressed by epithelial or stromal cells to the molecular classification of CRC. Gene signatures defining the molecular classifications derived by Sadanandam *et al.*[2] (786 genes), De Sousa E Melo *et al.*[1] (146 Affymetrix HG U133 Plus 2.0 probe sets) and Marisa *et al.*[3] (1,459 Affymetrix HG U133 Plus 2.0 probe sets) were retrieved from their original publications. All these probe sets were annotated according to corresponding estimations of recurrence risk (HR for continuous variables), possible upregulation in microdissected epithelial or stromal compartments in data set GSE35602 (fold change >1.5, *P* value <0.05 after Benjamini and Yekutieli correction[31]) and eventual specific upregulation in epithelial, endothelial, leukocyte or FAP cell populations according to data sets GSE39395 and GSE39396 (minimum fold change >1.5 compared to any other cell population; all raw *P* values <0.05).

For each classification, hierarchical clustering was performed on probe sets using correlation distance and Ward's agglomeration method. Clustering results were shown in a heat map in which sample grouping was enforced according to molecular subtype. CAF-, leukocyte- and epithelial-enriched clusters were identified according to the annotation of the probe sets that constituted them.

To evaluate the role of epithelial and stromal genes in the definition of the molecular subtypes, we trained a Prediction Analysis for Microarrays (PAM) classifier[32] to predict these subtypes using the probe sets included in the subtype signatures as predictors. For each classification, this analysis was performed using (i) all probe sets in the profile; (ii) all probe sets but excluding those part of the CAF- or leukocyte-enriched cluster; or (iii) epithelial-enriched clusters only, when existing. Global and subtype-specific error rates were computed in each case.

For each molecular classification, CAF-enriched clusters were summarized and their prognosis values were evaluated after excluding the poor-prognosis molecular subtype from the data set. In addition, patients not harboring the poor-prognosis subtype were classified according to the expression shown for the CAF-enriched cluster gene set (low, less than zero; high, greater than zero). Differences in relapse risk between these groups and the poor-prognosis subtype were then assessed. In all cases, association with recurrence was analyzed using Kaplan-Meier estimates and Cox proportional hazards models (see the **Supplementary Note** for further details).

Association of poor-prognosis subtypes with TGF-β. To evaluate the association of TGF-β genes with poor-prognosis subtypes (CCS3, stem-like and C4), we assessed the association of *TGFB1*, *TGFB2* and *TGFB3* expression with each molecular classification in data sets GSE33113, GSE14333 and GSE39582, respectively. We described previously which probe sets on the Affymetrix Human Genome U133 Plus 2.0 array correlate better with measurements of *TGFB1*, *TGFB2* and *TGFB3* expression obtained from qRT-PCR data[6]. The probe sets selected were as follows: *TGFB1*, 203085_s_at; *TGFB2*, 220406_at; *TGFB3*, 209747_at. Because of the generalized low expression intensity of *TGFB2*, overall expression of TGF-β genes was summarized using only *TGFB1* and *TGFB3* expression. Subtype comparisons were performed for *TGFB1*, *TGFB3* and overall TGF-β gene expression levels using standard linear models and two-sided *t* tests.

Association of poor-prognosis subtypes with stromal TGF-β response signatures. Generation of TBRSs has been described previously[6]. To assess association with poor-prognosis subtypes, we summarized each TBRS separately in data sets GSE33113, GSE14333 and GSE39582. Subtype comparisons were conducted using standard linear models and two-sided *t* tests.

Correlation of poor-prognosis subtypes with TGF-β. Correlation between *TGFB1* or *TGFB3* expression and the poor-prognosis subtypes (CCS3, stem-like and C4) was independently assessed in the data where they were originally described (GSE33113, GSE14333 and GSE39582, respectively). Summaries of similarities between subtypes and TGF-β gene expression profiles were evaluated via Spearman correlation. Summaries of signatures were centered and scaled before computing Spearman's correlation.

Tissue microarray construction. Formalin-fixed, paraffin-embedded tissue blocks of colorectal adenocarcinomas were retrieved from the archives of the Servei de Patologia from Hospital del Mar. Samples were obtained under informed consent and approval of the Tumor Bank Committees according to Spanish ethical regulations. The study followed the guidelines of the Declaration of Helsinki, and patient identity for pathological specimens remained anonymous in the context of this study. Different areas of invasive carcinoma and, when possible, different histological tumor patterns (cribriform, mucinous, poorly differentiated); adenomatous lesions from the same surgical sample, when available; and normal mucosa, located far from the carcinoma, were identified on corresponding slides stained with hematoxylin and eosin.

doi:10.1038/ng.3225

Immunohistochemistry. Immunostainings were carried out using 4-µm tissue sections according to standard procedures. Briefly, after antigen retrieval, samples were blocked with Peroxidase-Blocking Solution (Dako, S202386) for 10 min at room temperature, and primary antibodies were then incubated with samples overnight. Slides were washed with EnVision FLEX Wash Buffer (Dako, K800721), and the corresponding secondary antibody was incubated with the sample for 45 min at room temperature. Samples were developed using 3,3′-diaminobenzidine, counterstained with hematoxylin and mounted. Antibodies against CALD1 (HPA008066, Sigma), POSTN (HPA012306, Sigma) and IGFBP7 (HPA002196, Sigma) were used at 1:200 dilutions. Antibodies to EpCAM (AF960, R&D Systems) and FAP (MABS1004, Vitatex) were used at 1:50 dilutions.

Association of markers with recurrence in immunohistochemistry data. Immunohistochemistry was performed on TMAs of stage I, II and III CRC samples obtained from patients (n = 79) from Hospital del Mar distributed in 2 independent TMAs. TMAs contained two to five representative samples per tumor. Qualitative categorization of the samples according to staining in stromal versus epithelial cancer cells was performed by an expert pathologist (M.I.) in a blinded fashion with respect to clinical information. An intensity score from 0 to 3 (units of intensity, or UI) was given for each sample, and an average intensity score per tumor was calculated for each patient. For tumor-associated stroma, all cell types stained were taken into consideration.

We assessed the association between tumor relapse and the intensity of protein staining (scored from 0 to 3) using Kaplan-Meier estimates and Cox proportional hazard models. Staining values were evaluated as continuous (whenever possible) and categorical covariates (**Supplementary Note**). As samples were distributed across two independent TMAs, this information was included as an adjusting covariate in the analyses. In addition, Cox models were fitted adjusting by AJCC stage, treatment with adjuvant chemotherapy, age, tumor location and sex (see the **Supplementary Note** for further details).

Protein blots. Immunoblotting was carried out according to standard procedures. Antibodies against POSTN (HPA012306, Sigma), FAP (MABS1004, Vitatex) and IGFBP7 (HPA002196, Sigma) were used at 1:250 dilutions. Antibodies to CALD1 (HPA008066, Sigma) and β-actin (ab20272, Abcam) were used, respectively, at 1:1,000 and 1:30,000 dilutions.

Detection of markers for poor-prognosis subtypes in the Human Protein Atlas database. Qualitative categorization of the samples according to staining in stromal versus epithelial cancer cells was performed by an expert pathologist (M.I.) in a blinded fashion. For tumor-associated stroma, all cell types stained were taken into consideration.

Culture of patient-derived tumor organoids. Human biological samples used to expand tumor organoids were obtained from individuals treated at Hospital del Mar and Hospital de la Santa Creu i Sant Pau, under informed consent and approval of the Tumor Bank Committees according to Spanish ethical regulations. The study followed the guidelines of the Declaration of Helsinki, and patient identity for pathological specimens remained anonymous in the context of this study. CRC stem cells (CRCSCs) were identified and purified using the same protocol that we developed for normal colon stem cells (CoSCs)[16] and later used to isolate stem cell–like cells from CRC samples[33]. Briefly, CRC cells with high EphB2 levels were FACS sorted from dissociated tumors and cultured embedded in Matrigel (Basement Membrane Matrix Low Concentration, BD) with a simplified version of the normal CoSC-specific medium described by Jung *et al.*[16] (Advanced DMEM/F12, 10 mM HEPES, 1× Glutamax; 1× B-27 without retinoic acid, 1× N-2, 20 ng/ml bFGF (basic fibroblast growth factor); 50 ng/ml EGF (epidermal growth factor), 1 µM LY2157299 and 10 µM Y-27632). Under these conditions, cells with high EphB2 levels expanded as tumor organoids that we could propagate indefinitely, whereas cells with medium or low EphB2 levels did not. PDO7 and PDO8, a kind gift of G. Stassi (University of Palermo), were obtained from the dissociation of whole CRCs in suspension as described elsewhere[34]. They were cultured with the medium described above upon arrival to our laboratory. All cells were tested weekly for mycoplasma contamination with negative results.

Cell lines. Cell lines were maintained in DMEM supplemented with 10% FBS. KM12L4a and HT29-M6, CCD-18co fibroblast[35,36] cells, control or TGF-β secreting, have been described elsewhere. For bioluminescent tracking, cancer cells were infected with a virus encoding a fusion protein reporter construct consisting of red fluorescent protein (mCherry) and firefly luciferase. For cell tracking, fibroblasts were infected with a construct encoding eGFP. Knockdown experiments were conducted using shRNA targeting *TGFBR1* (Sigma Aldrich). Non-targeting shRNA sequence was used as a control (Sigma Aldrich). The construct for dominant-negative mutant form of the type II TGF-β receptor (DNR)[37] cloned in the pcDNA 3 vector was provided by L. Wakefield (US National Cancer Institute). All genetically manipulated cells were produced via lentiviral infection. All cell lines were tested weekly for mycoplasma contamination with negative results.

Fibroblast treatments. For association with each poor-prognosis subtype, a differential expression analysis was performed by comparing colonic fibroblasts treated or not with TGF-β1 for 8 h (fold change >2, Benjamini-Yekutieli P value <0.05). This gene set was obtained from an Affymetrix Human Genome U133 Plus 2.0 expression data set including control fibroblast samples and fibroblasts stimulated with recombinant TGF-β1 (n = 2). Genes were ranked according to the t statistics derived from this comparison. This ranking was then tested for enrichment of genes found to be highly upregulated in each poor-prognosis subtype and associated with prognosis using GSEA[38]. These data were submitted in GEO and are accessible through GEO accession GSE64192.

For *in vivo* coinjections, prestimulation of fibroblasts before subcutaneous injection into nude mice was performed as follows. Cells were cultured in DMEM supplemented with 5% FBS in the presence or absence of recombinant human TGF-β1 (5 ng/ml) for a total of 96 h. Media were changed every 24 h.

Quantification of stromal and epithelial tumor compositions with Masson's trichrome staining. Masson's trichrome staining was carried out using 4-µm tissue sections according to standard procedures (Sigma, ref. HT15). Fiji Trainable Weka segmentation, an ImageJ plugin based on the Weka[39] machine learning Java library, was used to classify images on the basis of local colorimetric, textural and structural features in the neighborhood of each pixel. Images were processed with a custom macro (see URLs). Pairwise comparisons were explored using a two-sided t test.

Orthotopic mouse studies. All experiments with mouse models were approved by the Animal Care and Use Committee of Barcelona Science Park (CEEA-PCB) and the Catalan government. Cells were injected subcutaneously or intrasplenically[40] into 5- to 6-week-old females NSG or Swiss nude mice (Jackson Laboratories), which were followed until sacrifice. Sample size was predetermined empirically according to previous experience using the same strains and treatments. Generally, we used $n > 6$ mice per genotype and condition. We ensured that experimental groups were balanced in terms of mouse age, sex and weight. The experiments were not randomized. Mice were caged together and treated in the same way. Neither the technician nor the investigator could distinguish them during the experiment or when assessing outcomes. The general condition of the mice was monitored using animal fitness and weight controls throughout the experiment. When deteriorating clinical alterations were observed, the mice were excluded from the study and sacrificed. Subcutaneous tumor appearance was assessed by palpation. Tumor volume was measured twice a week until sacrifice. Tumorigenesis after intrasplenic injection was assessed by bioluminescent imaging (see "Bioluminescent imaging and analysis").

Estimation of tumor-initiating cell frequency and assessment of synergistic interactions. Serial dilutions of control cancer cells (HT29-M6) or cancer cells with forced TGF-β1 expression (HT29-M6$^{TGF\text{-}\beta}$) were coinjected subcutaneously with or without fibroblasts (FIB; 5×10^4). The proportion of mice showing development of tumors was modeled using a logistic regression model in which TGF-β status, fibroblast inoculation and dose were considered as covariates. Dose was included in the model on a logarithmic scale. An interaction term was also included in the model to assess whether a synergistic effect between TGF-β and fibroblasts existed on tumor development.

A likelihood-ratio test (LRT) was used to assess the significance of the corresponding interaction term as opposed to an additive effect of TGF-β and fibroblasts. For each condition, the interaction model was used to estimate the logit functional relationships between the proportion of mice developing tumors and the number of injected cells, as well as their 95% confidence bands. For each sample condition, the proportion of TICs and its 95% confidence interval was estimated using ELDA[41], a methodology that assumes the Poisson single-hit model. An LRT test was used to assess pairwise differences in TIC numbers between conditions.

Pharmacological TGF-β inhibition by LY2157299 treatment *in vivo*. Mice were treated twice a day with a dose of 4 mg *per os*, starting 3 d before cancer cell inoculation and continued until the end of the experiment. Control mice were treated with vehicle. The drug was synthesized in house and prepared as previously described[6].

Exome sequencing analysis. Genomic DNA from each tumor organoid (3 μg; quantitated by Qubit fluorometer DNA Hs Assay) was fragmented, and exome capture was performed using Nimblegen Sequence Capture EZ Human Exome Library v2.0. Libraries were sequenced on a HiSeq 2000 platform (Illumina), using 2 × 100-bp paired-end sequencing. Fastqc files were generated by FastQC software v0.10.1.

SNP and copy number calling in exome sequencing data. Samples were aligned to human reference genome version GRCh37 (ref. 42) using BWA aligner software[43] with default parameters. Preprocessing included removing duplicate reads[44] and base quality recalibration using BaseRecalibrator from the Genome Analysis Toolkit[45]. Local realignment was performed around indels (using IndelRealigner from ref. 45). The UnifiedGenotyper algorithm from ref. 45 was used to call SNPs in the merged file (see refs. 46,47 for details). SNPs were filtered with the parameters recommended for best practice by the GATK website[48], removing candidates that fulfilled at least one of the following criteria: QD <2.0, FS >60.00, MQ <40.0 and DP <10. Variants were annotated using dbSNP database version 13, UCSC human RefGene69 and the software snpEFF[49].

Copy number alteration discovery. Total counts per probe in the enrichment array were computed for each sample after removing duplicates using the Rsamtools package for R[50]. Gene annotations for the probes were downloaded using the UCSC browser website[51]. GC content was computed for each probe using the package BSgenome.Hsapiens.UCSC.hg19 and the rtracklayer package[52]. Probe counts were normalized by exon length. Normalized counts were used to correct for GC content–related sequencing biases. A generalized additive model was fitted to the data through the function gam in the mgcv package[53]. Normalized and GC content–corrected counts were used to compute gene-level copy number estimates by calculating the mean across exon estimates.

Genomes were segmented on the basis of corrected counts using the segment function from the DNAcopy R package[54]. Segments with a mean lower than −3.5 (higher than 0.5) were considered deletions (amplifications). The remaining segments were classified as diploid.

28. Barrett, T. & Edgar, R. Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. *Methods Enzymol.* **411**, 352–369 (2006).
29. Jorissen, R.N. *et al.* Metastasis-associated gene expression changes predict poor outcomes in patients with Dukes stage B and C colorectal cancer. *Clin. Cancer Res.* **15**, 7642–7651 (2009).
30. De Sousa E Melo, F. *et al.* Methylation of cancer-stem-cell-associated Wnt target genes predicts poor prognosis in colorectal cancer patients. *Cell Stem Cell* **9**, 476–485 (2011).
31. Benjamini, Y., Krieger, A.M. & Yekutieli, D. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**, 491–507 (2006).
32. Tibshirani, R., Hastie, T., Narasimhan, B. & Chu, G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc. Natl. Acad. Sci. USA* **99**, 6567–6572 (2002).
33. Merlos-Suárez, A. *et al.* The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* **8**, 511–524 (2011).
34. Lombardo, Y. *et al.* Bone morphogenetic protein 4 induces differentiation of colorectal cancer stem cells and increases their response to chemotherapy in mice. *Gastroenterology* **140**, 297–309 (2011).
35. Morikawa, K., Walker, S.M., Jessup, J.M. & Fidler, I.J. *In vivo* selection of highly metastatic cells from surgical specimens of different primary human colon carcinomas implanted into nude mice. *Cancer Res.* **48**, 1943–1948 (1988).
36. Morikawa, K. *et al.* Influence of organ environment on the growth, selection, and metastasis of human colon carcinoma cells in nude mice. *Cancer Res.* **48**, 6863–6871 (1988).
37. Böttinger, E.P., Jakubczak, J.L., Haines, D.C., Bagnall, K. & Wakefield, L.M. Transgenic mice overexpressing a dominant-negative mutant type II transforming growth factor β receptor show enhanced tumorigenesis in the mammary gland and lung in response to the carcinogen 7,12-dimethylbenz-[*a*]-anthracene. *Cancer Res.* **57**, 5564–5570 (1997).
38. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
39. Witten, I.H. *et al.* Weka: practical machine learning tools and techniques with Java implementations. *Proceedings of the ICONIP/ANZIIS/ANNES'99 Workshop on Emerging Knowledge Engineering and Connectionist-Based Information Systems* 192–196 (1999).
40. Warren, R.S., Yuan, H., Matli, M.R., Gillett, N.A. & Ferrara, N. Regulation by vascular endothelial growth factor of human colon cancer tumorigenesis in a mouse model of experimental liver metastasis. *J. Clin. Invest.* **95**, 1789–1797 (1995).
41. Hu, Y. & Smyth, G.K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347**, 70–78 (2009).
42. Lander, E.S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
43. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
44. O'Brien, C.A. *et al.* ID1 and ID3 regulate the self-renewal capacity of human colon cancer–initiating cells through p21. *Cancer Cell* **21**, 777–792 (2012).
45. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
46. DePristo, M.A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
47. Pathmanathan, S., Hamilton, E., Atcheson, E. & Timson, D.J. The interaction of IQGAPs with calmodulin-like proteins. *Biochem. Soc. Trans.* **39**, 694–699 (2011).
48. Clevers, H. & Batlle, E. EphB/EphrinB receptors and Wnt signaling in colorectal cancer. *Cancer Res.* **66**, 2–5 (2006).
49. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w^{1118}; iso-2; iso-3. *Fly* (*Austin*) **6**, 80–92 (2012).
50. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2014).
51. Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
52. Lawrence, M., Gentleman, R. & Carey, V. rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics* **25**, 1841–1842 (2009).
53. Wood, S.N. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. Royal Statist. Soc. (B)* **73**, 3–36 (2011).
54. Venkatraman, E.S. & Olshen, A.B. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics* **23**, 657–663 (2007).