

intracellular autofluorescence: a biomarker for epithelial cancer stem cells

Irene Miranda-Lorenzo¹, Jorge Dorado¹, Enza Lonardo¹, Sonia Alcala¹, Alicia G Serrano², Jenifer Clausell-Tormos¹, Michele Cioffi¹, Diego Megias³, Sladjana Zagorac¹, Anamaria Balic¹, Manuel Hidalgo⁴, Mert Erkan⁵, Joerg Kleeff⁵, Aldo Scarpa⁶, Bruno Sainz Jr¹ & Christopher Heeschen^{1,7}

¹Stem Cells and Cancer Group, Molecular Pathology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. ²Melanoma Group, Molecular Pathology Programme, CNIO, Madrid, Spain. ³Confocal Microscopy Unit, Biotechnology Programme, CNIO, Madrid, Spain. ⁴Gastrointestinal Cancer Clinical Research Unit, Clinical Research Programme, CNIO, Madrid, Spain. ⁵Chirurgische Klinik, Technische Universität München, Munich, Germany. ⁶Università degli Studi di Verona, Centro Ricerca Applicata ARC-NET, Verona, Italy. ⁷Centre for Stem Cells in Cancer & Ageing, Barts Cancer Institute, Queen Mary University of London, London, UK. Correspondence should be addressed to B.S. (bruno.sainz@uam.es) or C.H. (c.heeschen@qmul.ac.uk).

cancer stem cells (cscs) are thought to drive tumor growth, metastasis and chemoresistance. Although surface markers such as cd133 and cd44 have been successfully used to isolate cscs, their expression is not exclusively linked to the csc phenotype and is prone to environmental alteration. We identified cells with an autofluorescent subcellular compartment that exclusively showed csc features across different human tumor types. Primary tumor-derived autofluorescent cells did not overlap with side-population (sp) cells, were enriched in sphere culture and during chemotherapy, strongly expressed pluripotency-associated genes, were highly metastatic and showed long-term *in vivo* tumorigenicity, even at the single-cell level. Autofluorescence was due to riboflavin accumulation in membrane-bounded cytoplasmic structures bearing ATP-dependent ABCG2 transporters. In summary, we identified and characterized an intrinsic autofluorescent phenotype in cscs of diverse epithelial cancers and used this marker to isolate and characterize these cells.

The CSC model, which postulates a hierarchical tumor organization such that only a smaller subset of ‘stem-like’ cells is responsible for tumor promotion and cellular heterogeneity¹, provides a plausible explanation for therapeutic resistance and disease relapse following initial tumor regression². CSCs are regularly identified by the use of surface markers, which have emerged as powerful tools for isolating distinct cell populations from freshly harvested primary tumors^{3–6}. Nonetheless, these markers bear the caveats that not only can their expression levels change depending on environmental conditions (such as patient-derived xenografts (PDXs) and primary cell culture) and tissue digestion protocols, but their expression is neither exclusively nor reproducibly linked to a functional CSC phenotype across or even within similar tumor types⁷. Indeed, the use of different surface markers and isolation methodologies has created conflicting data in some settings, emphasizing our lack of sensitive methods to isolate these cells as well as our still immature knowledge of their role in cancer^{8–11}.

Thus, alternate detection and isolation methods based on CSC functional properties would avoid the use of artifact-prone surface markers and should also provide new insights into CSC biology. Toward this end we have identified an intrinsic autofluorescent phenotype in CSCs of diverse human solid tumors and used this marker as a novel and functionally relevant tool to both isolate and characterize these cells. Specifically, autofluorescent cells could be tracked and isolated, by flow cytometry, in tumors of pancreatic ductal adenocarcinoma (PDAC), colorectal carcinoma (CRC), hepatocellular carcinoma (HCC) and non-small-cell lung carcinoma (NSCLC). These cells demonstrated enhanced CSC features and phenotypes, such as exclusive long-term tumorigenicity and invasiveness *in vivo*. We determined that autofluorescence was a result of the accumulation of the fluorescent vitamin riboflavin in ABCG2-coated vesicles exclusively located within the cytoplasm of CSCs. This new autofluorescent marker can be easily used to isolate CSCs across numerous solid tumors and should allow for the eventual development of screening platforms aimed at identifying anticancer compounds that target CSCs.

results

identification of new functional marker for cscs

Although we have previously shown that different cell surface markers (for example, CD133) can be used to isolate CSCs from freshly resected PDAC tissue¹², we found that the expression levels of these markers could vary considerably on the basis of time to analysis and environmental changes, often resulting in variable and modest functional enrichment for CSCs (**Supplementary Fig. 1**). Moreover, although CSCs expressing combinations of surface markers (for example, CD133 and CD44) showed stronger enrichment than cells that lacked surface marker expression, these marker combinations still left behind large populations of cells expressing either one of the two markers, which were themselves only modestly and variably enriched in CSCs (**Supplementary Fig. 1c**). Interestingly, SP cells, which were previously shown to contain CSCs in some established cancer cell lines^{13–15} but not in

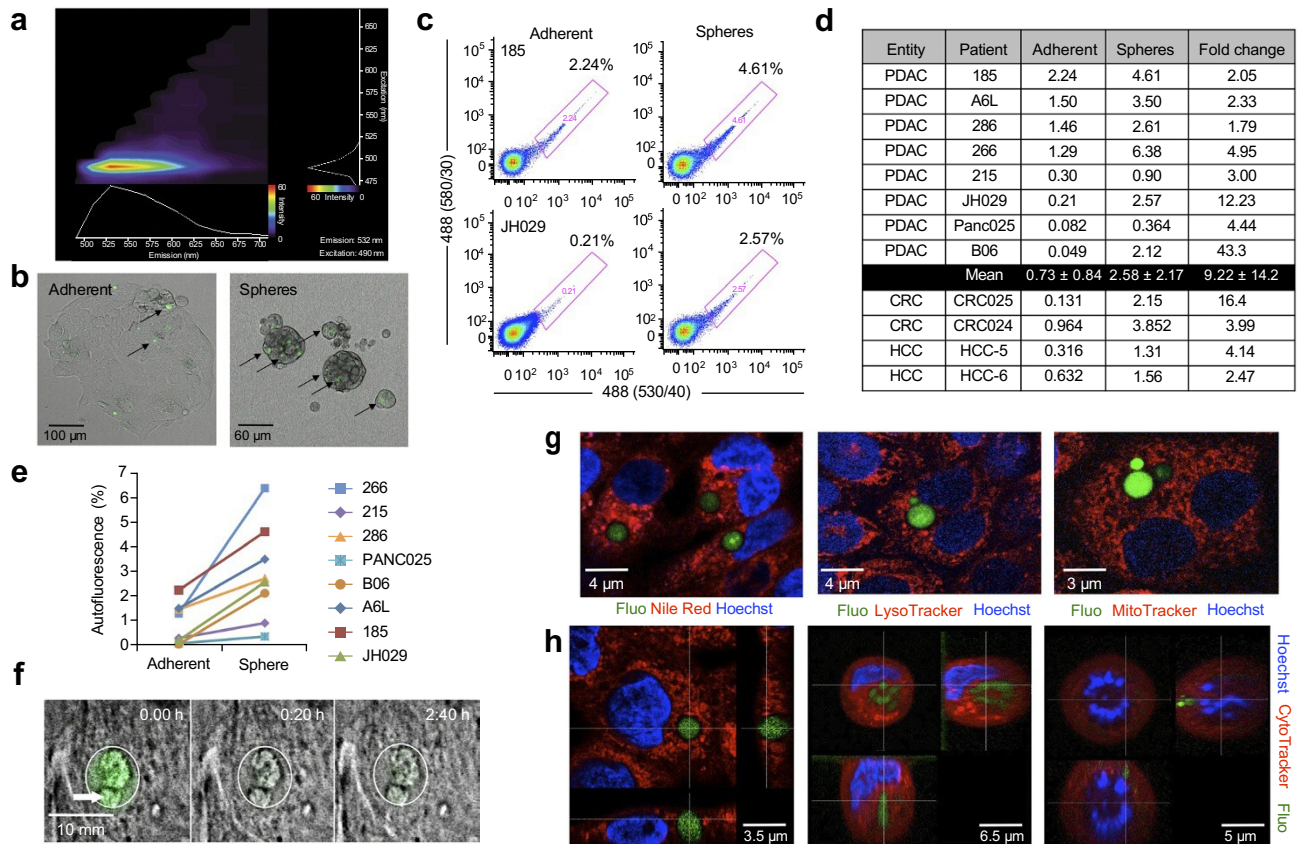


Figure 1 | Identification of autofluorescent cancer stem cells. (a) Spectrum of autofluorescence in primary PDAC PDX-derived cancer cells. (b) Representative images of autofluorescent cells (arrows) in primary PDAC PDX-derived *in vitro* cultures grown as adherent cells or spheres ($n = 6$). (c) Flow cytometry analysis of autofluorescent content in adherent (left) and sphere cultures (right). Autofluorescent cells are excited with a 488-nm blue laser and best selected as the intersection with filters 530/40 and 580/30. (d) Percentage of autofluorescent cells in adherent and sphere cultures for different PDXs ($n = 12$). Fold change is between sphere and adherent percentages. (e) Autofluorescence across different PDAC PDX tumors ($n = 8$) cultured as adherent cells or spheres. (f) Fluorescence recovery after photobleaching in autofluorescent vesicles (time listed as hours:minutes). Arrow indicates bleached area. Circle highlights autofluorescent vesicle. (g) Confocal images of PDAC PDX-derived *in vitro* cultures showing the localization of the indicated markers (Nile red, lipid droplets; LysoTracker, lysosomes; MitoTracker, mitochondria; Hoechst, nuclear staining) relative to autofluorescence (green). (h) z-stack confocal images of CytoTracker-stained cells (red) illustrating the cytosolic localization of the autofluorescent (green) compartment.

others^{16,17}, did not efficiently enrich for CSCs in primary PDAC (Supplementary Fig. 2a). Similarly disappointing *in vivo* results were obtained for other markers such as aldehyde dehydrogenase (Supplementary Fig. 2b). In contrast, spheres were enriched in CSCs (Supplementary Figs. 2b and 3), but sphere formation is not suitable for *ad hoc* identification of CSCs.

We sought to develop new methodologies, devoid of the aforementioned caveats, to efficiently identify CSCs from primary tumors. While searching for such markers, we noticed a small subpopulation of autofluorescent cells in freshly digested tumors from PDAC PDXs and in freshly digested primary patient tumors (PDAC-Tumor) that could be excited (excitation and emission maxima of 490 nm and 532 nm, respectively) with a standard blue laser (for example, 488 nm) (Fig. 1a,b) but not with yellow-green (for example, 561 nm) or red (for example, 640 nm) lasers (Supplementary Fig. 4a). These autofluorescent cells were also visible in PDAC PDX-derived adherent cultures, were strongly enriched in spheres and showed an enhanced *in vivo* tumorigenic phenotype compared to non-autofluorescent cells or cells isolated via more traditional methods (Supplementary Figs. 1c, 2 and 3).

Autofluorescence was previously reported for glioblastoma cells but turned out to be related to contamination with 293T cells

stably expressing GFP¹⁸. To exclude such artifacts, we demonstrated that (i) the autofluorescence spectrum observed was distinct from that of GFP (Supplementary Fig. 4b), (ii) there was no evidence of *gfp* mRNA or GFP protein in our samples (Supplementary Fig. 4c), (iii) autofluorescent and non-autofluorescent cells could not be distinguished by size or granularity (Supplementary Fig. 4d), and (iv) autofluorescent and non-autofluorescent cells were genotypically identical to unsorted bulk cells as determined by TaqMan OpenArray technology (Supplementary Fig. 5a). We also showed that (v) autofluorescent cells were of epithelial origin (i.e., EPCAM⁺) (Supplementary Fig. 5b), (vi) they could not be detected in the tumor stroma (Supplementary Fig. 5c), and (vii) autofluorescence was not a result of dying or apoptotic cells (Supplementary Fig. 6). In sum, autofluorescence was a cell-specific authentic feature present within the bulk tumor cell population.

We next analyzed the autofluorescent phenotype across a large panel of tumors and showed that autofluorescence could be tracked and reproducibly quantified by flow cytometry (Fig. 1c). We observed cells with this phenotype in a large panel of PDAC PDX-derived *in vitro* cultures, and these cells were consistently enriched in culture conditions promoting sphere growth¹⁹

(Fig. 1d,e). We also detected autofluorescent cells in PDX-derived cultures from other tumor types such as CRC, HCC and NSCLC (Fig. 1d and Supplementary Fig. 7). Interestingly, the autofluorescence signal was confined to one or more distinct vesicles but could be eliminated in all vesicles by localized laser-induced photobleaching (Fig. 1f), indicating that the source of the signal was diffusible and neither attached to any structure nor crystalline

in nature, and that the apparently separate autofluorescent compartments present in some cells were interconnected. Moreover, autofluorescence did not colocalize with lipid droplets or lysosomes, nor was it mitochondria-restricted, as demonstrated by costaining with Nile red, LysoTracker or MitoTracker, respectively (Fig. 1g); rather, the autofluorescence was restricted to distinct membrane-bound cytoplasmic vesicles (Fig. 1h).

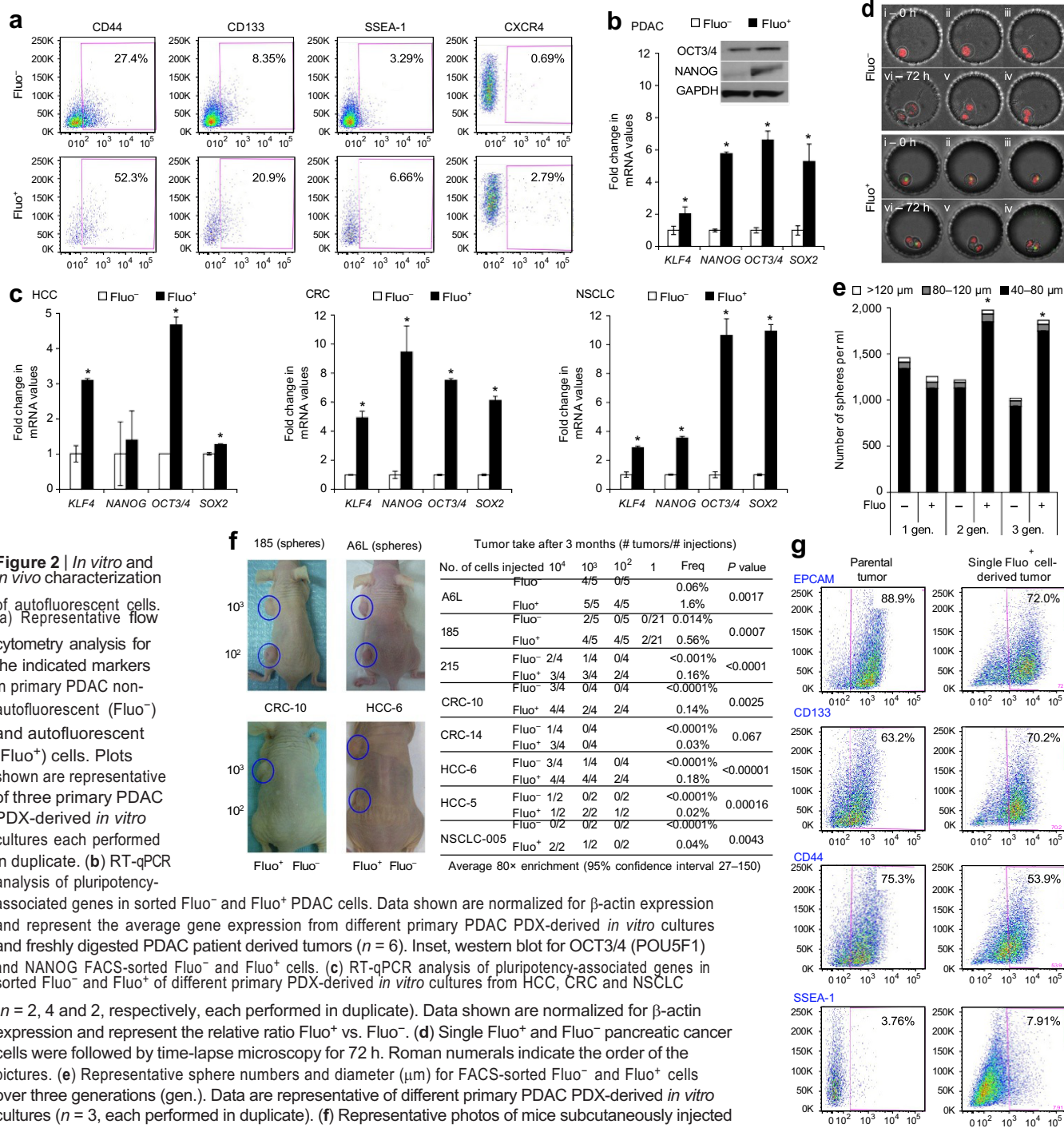


Figure 2 | In vitro and in vivo characterization

of autofluorescent cells. (a) Representative flow cytometry analysis for the indicated markers in primary PDAC non-autofluorescent (Fluo⁻) and autofluorescent (Fluo⁺) cells. Plots shown are representative of three primary PDAC PDX-derived *in vitro* cultures each performed in duplicate. (b) RT-qPCR analysis of pluripotency-associated genes in sorted Fluo⁻ and Fluo⁺ PDAC cells. Data shown are normalized for β-actin expression and represent the average gene expression from different primary PDAC PDX-derived *in vitro* cultures and freshly digested PDAC patient derived tumors (*n* = 6). Inset, western blot for OCT3/4 (POU5F1) and NANOG FACS-sorted Fluo⁻ and Fluo⁺ cells. (c) RT-qPCR analysis of pluripotency-associated genes in sorted Fluo⁻ and Fluo⁺ of different primary PDX-derived *in vitro* cultures from HCC, CRC and NSCLC

(*n* = 2, 4 and 2, respectively, each performed in duplicate). Data shown are normalized for β-actin expression and represent the relative ratio Fluo⁺ vs. Fluo⁻. (d) Single Fluo⁺ and Fluo⁻ pancreatic cancer cells were followed by time-lapse microscopy for 72 h. Roman numerals indicate the order of the pictures. (e) Representative sphere numbers and diameter (μm) for FACS-sorted Fluo⁻ and Fluo⁺ cells over three generations (gen.). Data are representative of different primary PDAC PDX-derived *in vitro* cultures (*n* = 3, each performed in duplicate). (f) Representative photos of mice subcutaneously injected with Fluo⁺ and Fluo⁻ cells (PDAC PDX) (left) and summary of *in vivo* tumorigenicity of subcutaneously injected FACS-sorted Fluo⁻ and Fluo⁺ cells from PDXs of pancreatic (185, A6L, 215), colorectal (CRC-10, CRC-14), liver (HCC-6, HCC-5) and lung (NSCLC-005) cancers (right). (g) Flow cytometry analysis of the indicated markers in parental and single-cell-derived 185 tumors. Error bars (b,c,e), s.d. (*n* = biological replicates). Statistical significance was assessed by Mann-Whitney test, **P* < 0.05.

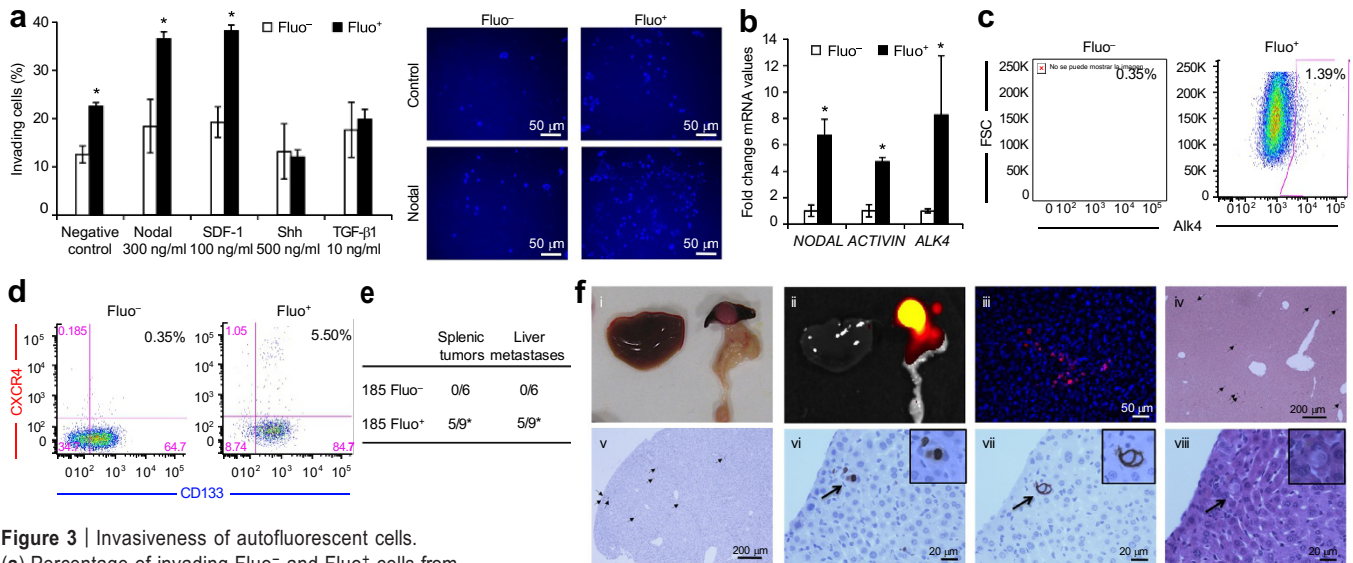


Figure 3 | Invasiveness of autofluorescent cells.

(a) Percentage of invading Fluo⁻ and Fluo⁺ cells from different primary PDAC PDX-derived *in vitro* cultures ($n = 2$, each performed in triplicate) (left) through Matrigel following stimulation with the indicated factors. Representative images of invaded cells (right). (b) RT-qPCR analysis of *NODAL*, *ACTIVIN (INHBA)* and *ALK4 (ACVR1)* in sorted Fluo⁻ and Fluo⁺ cells of primary PDAC PDX-derived *in vitro* cultures ($n = 3$). Data are normalized for β -actin expression and represent the relative ratio Fluo⁺ vs. Fluo⁻. (c) Flow cytometry analysis of ALK4 receptor expression in Fluo⁻ and Fluo⁺ cells. Plots shown are representative of different primary PDAC PDX-derived *in vitro* cultures ($n = 2$, each performed in duplicate). FSC, forward scatter. (d) Flow cytometry analysis of CXCR4 and CD133 expression on Fluo⁻ and Fluo⁺ cells. Plots shown are representative of different primary PDAC PDX-derived *in vitro* cultures ($n = 3$, each performed in duplicate). (e) Quantification of *in vivo* tumorigenicity and subsequent liver metastases for Fluo⁻ and Fluo⁺ cells. (f) Representative images of pancreatic cancer cell human metastases in mouse livers 4 months after intrasplenic injection of PDAC PDX (185) Fluo⁺ sorted cells. Panels show (i) mouse liver, spleen and pancreas; (ii) IVIS (*in vivo* imaging system) for mCherry⁺ cells in the spleen; (iii) *ex vivo* whole-tissue confocal image for mCherry⁺ cells in the liver; (iv) *in situ* hybridization using a human-specific Alu probe; (v) immunohistochemistry for DsRed; and (vi–viii) serial sections stained for DsRed, human cytokeratin19 and hematoxylin-eosin, respectively. Arrows indicate cells stained positive for indicated markers. Error bars (a,b), s.d. ($n =$ biological replicates). Statistical significance was assessed by Mann-Whitney test, $*P < 0.05$.

Autofluorescent cells display functional features of csCs

We next determined whether these cells were phenotypically and functionally distinct from their non-autofluorescent counterparts. First, although CSC surface markers were variably overexpressed in autofluorescent cells, which is in line with the significantly ($P < 0.05$; P values calculated via Mann-Whitney test throughout) stronger enrichment for tumorigenic cells in the autofluorescent population (Supplementary Fig. 3), none of these markers was exclusively restricted to autofluorescent cells (Fig. 2a). Reverse-transcription quantitative PCR (RT-qPCR) analysis revealed that autofluorescent cells also significantly ($P < 0.05$) overexpressed pluripotency-associated genes at both the mRNA and protein level (Fig. 2b,c and Supplementary Fig. 7a–c). Last, using a customized nanowell chip, which allows one to microscopically follow single cells, we observed that over a period of 72 h, autofluorescent cells gave rise to both autofluorescent and non-autofluorescent cells; however, out of more than 400 single non-autofluorescent cells monitored, none gave rise to autofluorescent cells irrespective of whether they underwent cell division (Fig. 2d and Supplementary Fig. 8a). Thus, in contrast to earlier cell population-based studies²⁰, our data argue against the notion that non-CSCs are capable of replenishing CSCs.

Regarding self-renewal capacity, autofluorescent cells formed significantly ($P < 0.05$) more secondary and tertiary spheres *in vitro*, whereas non-autofluorescent cells appeared to lose this capacity during passaging (Fig. 2e). To more rigorously test this, we performed serial *in vivo* transplantation experiments and observed that tumors derived from non-autofluorescent cells disappeared

during serial passaging, whereas autofluorescent cell-derived tumors maintained their intrinsically high self-renewal capacity (Supplementary Fig. 8b), demonstrating that long-term tumorigenic capacity was restricted to autofluorescent cells. Moreover, implantation of decreasing numbers of autofluorescent cells from PDX *in vitro* cultures from various sources (PDAC, HCC, CRC and NSCLC) revealed that autofluorescent cells were more tumorigenic (Fig. 2f). Notably, for tumor PDAC PDX 185, we simultaneously analyzed a large panel of putative CSC markers, which revealed that autofluorescence as a single marker more strongly enriched for CSCs than did any of the other tested single markers or combinations thereof (Supplementary Figs. 1c, 2 and 3). In addition, whereas single non-autofluorescent cells never gave rise to tumors, single autofluorescent cells generated tumors that recapitulated the heterogeneous composition of the original tumors at the level of autofluorescence, expression of cell surface markers (Fig. 2g) and pluripotency-associated genes (Supplementary Fig. 8c). Finally, we sorted autofluorescent cells from several freshly digested PDX tumors as well as from primary tumors from patients and consistently found that these cells exhibited stem-like properties, including the expression of pluripotency-associated genes and significantly ($P < 0.05$) enhanced *in vivo* tumorigenicity without any *in vitro* culturing (Supplementary Fig. 9a,b).

Autofluorescent cells are highly invasive

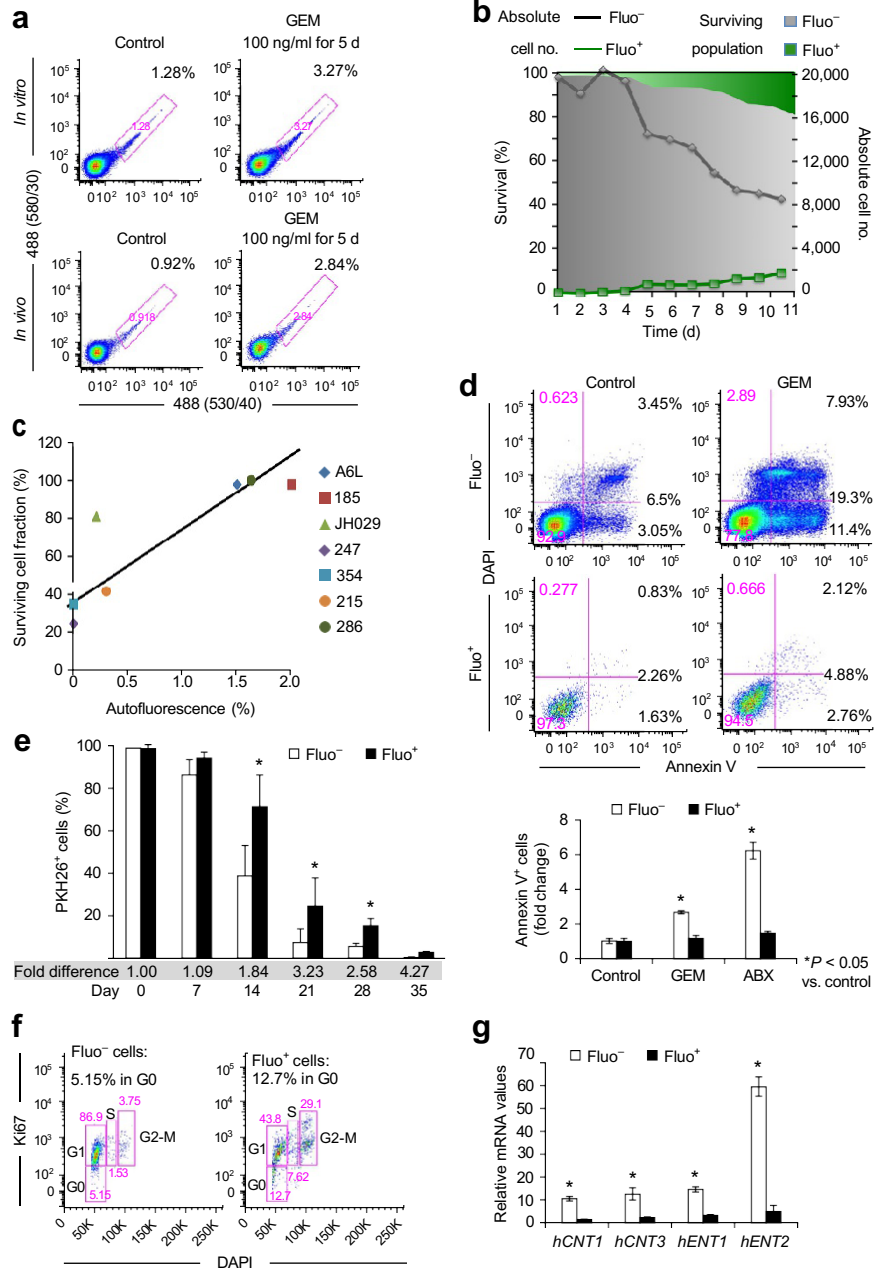
Holding true to the CSC phenotype²¹, *in vitro* invasion assays revealed an overall enhanced invasive capacity of autofluorescent cells (Fig. 3a). Chemoattraction of invading cells by the CXCR4

Figure 4 | Autofluorescent cells are resistant to chemotherapy. **(a)** Representative flow cytometry analysis of autofluorescence in control versus gemcitabine (GEM)-treated cells for different primary PDX-derived (PDAC, CRC) *in vitro* cultures (top, $n = 3, 1$, respectively) and PDAC PDX tumors (bottom, $n = 3$) each performed in triplicate. **(b)** Chemoresistance of autofluorescent PDAC PDX-derived *in vitro* cultures (185) cells during 12 d of treatment with GEM. Areas indicate surviving cell fraction. Lines indicate absolute number of surviving cells. **(c)** Correlation between autofluorescence and chemoresistance for different primary PDAC PDX-derived *in vitro* cultures ($n = 7$)

treated with GEM for 72 h. **(d)** Representative flow cytometry analysis of annexin V staining in GEM-treated cells for different PDAC PDX-derived *in vitro* cultures compared to control-treated cells (top), and quantification of the annexin V staining in GEM-treated and Abraxane (ABX)-treated cells ($n = 2$) each performed in triplicate (bottom). **(e)** Loss of PKH26 labeling in Fluo⁻ and Fluo⁺ cells during 5 weeks of culture. Data are representative of different primary PDAC PDX-derived *in vitro* cultures ($n = 2$, each performed in duplicate). **(f)** Representative flow cytometry analysis for Ki67 and DAPI for sorted Fluo⁻ and Fluo⁺ cells derived from a freshly digested PDAC PDX tumor ($n = 1$, performed in triplicate). **(g)** RT-qPCR analysis of human concentrative nucleoside transporter (hCNT) 1 and 3 and human equilibrative nucleoside transporter (hENT) 1 and 2 (*SLC28A1*, *SLC28A3*, *SLC29A1* and *SLC29A2*) in sorted Fluo⁻ and Fluo⁺ cells of different primary PDAC PDX-derived *in vitro* cultures. Data are normalized for β -actin expression and represent the relative ratio Fluo⁺ vs. Fluo⁻ ($n = 3$, each performed in triplicate). Error bars (**d,e,g**), s.d. ($n =$ biological replicates). Statistical significance was assessed by Mann-Whitney test, * $P < 0.05$.

ligand SDF-1 (ref. 22) and the ALK4 ligand Nodal¹⁹ both resulted in enhanced invasion of these cells, whereas sonic hedgehog and TGF- β 1 had no specific impact.

Consistently, autofluorescent cells overexpressed ALK4 as well as its specific ligands Nodal and Activin (Fig. 3b), and they also harbored a subset of ALK4⁺ cells (Fig. 3c). Moreover, autofluorescent cells (Fluo⁺) also overexpressed CXCR4 (Fig. 2a), and these Fluo⁺CXCR4⁺ cells were additionally enriched in CD133, a subpopulation (CD133⁺CXCR4⁺) that we have previously shown to be metastatic CSCs¹² (Fig. 3d). Intraspinal injections of cells expressing an H2B-mCherry reporter revealed exclusive tumor formation and subsequent cell dissemination to the liver (i.e., micrometastases) for autofluorescent cells only, as determined by RT-qPCR analysis of human *GAPDH* and mCherry-encoding mRNA expression in livers (Fig. 3e and data not shown). These data were further corroborated by *ex vivo* whole liver tissue confocal microscopy showing mCherry⁺ cells and immunohistochemistry demonstrating the presence of cells positive for human-specific ALU, huCK19 and DsRed (Fig. 3f).



Autofluorescent cells are highly chemoresistant

CSCs are believed to be highly chemoresistant. In line with this hypothesis, treatment with the cytotoxic drug gemcitabine resulted in the enrichment of autofluorescent cells *in vitro* and *in vivo* (Fig. 4a). Longer treatment with gemcitabine *in vitro* reduced the non-autofluorescent cell population to ~40% of the total starting population, whereas autofluorescent cells showed no reduction in absolute cell numbers but were rather relatively enriched (Fig. 4b). We expanded these findings to a representative panel of seven primary PDAC PDX tumors, in which overall chemoresistance *in vitro* correlated with autofluorescence content (Fig. 4c). Last, we observed *in vivo* that long-term gemcitabine-treated autofluorescent cells were tumorigenic, whereas their autofluorescent-negative counterparts were not (Supplementary Fig. 9c).

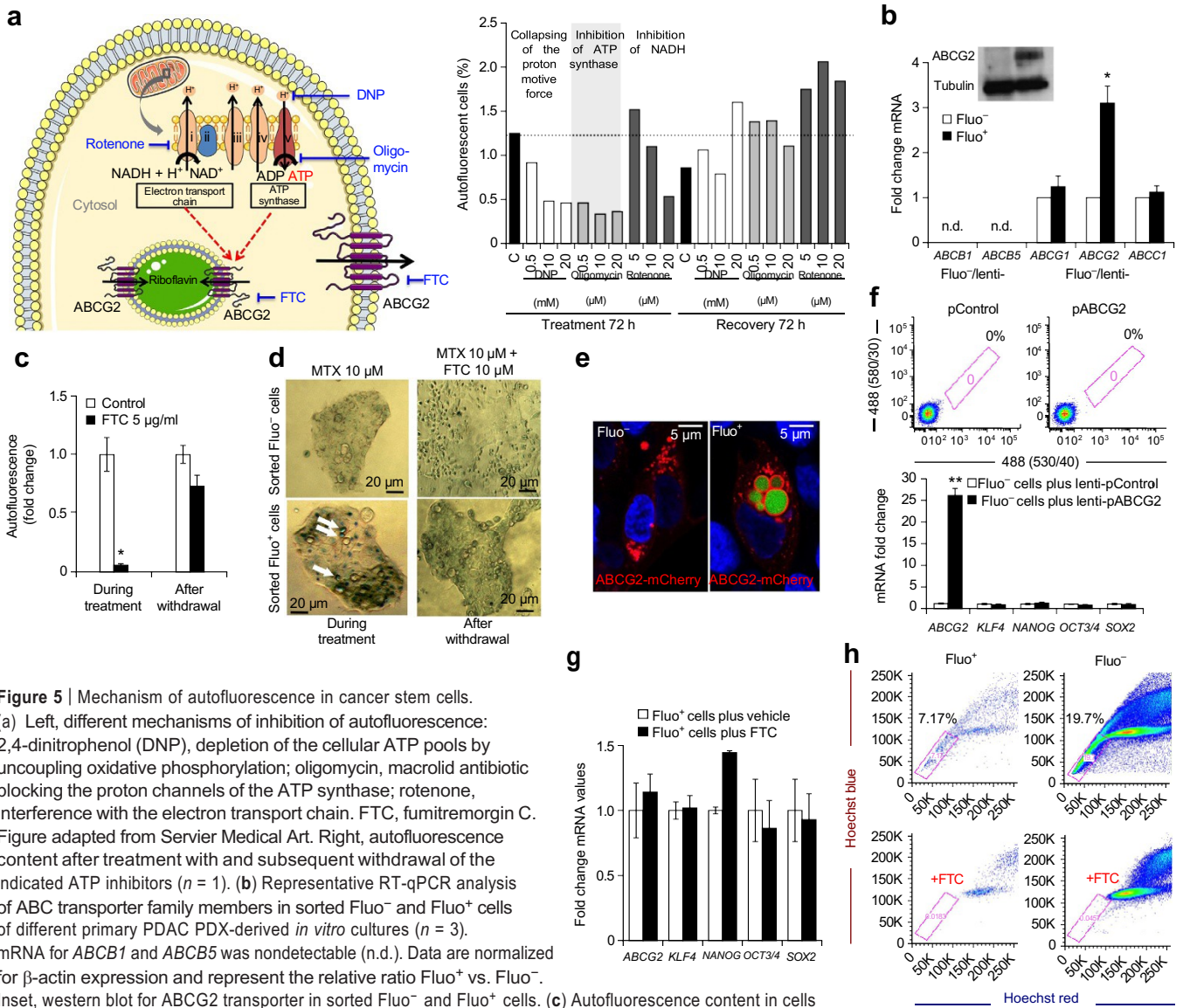


Figure 5 | Mechanism of autofluorescence in cancer stem cells.

(a) Left, different mechanisms of inhibition of autofluorescence: 2,4-dinitrophenol (DNP), depletion of the cellular ATP pools by uncoupling oxidative phosphorylation; oligomycin, macrolid antibiotic blocking the proton channels of the ATP synthase; rotenone, interference with the electron transport chain. FTC, fumitremorgin C. Figure adapted from Servier Medical Art. Right, autofluorescence content after treatment with and subsequent withdrawal of the indicated ATP inhibitors ($n = 1$). (b) Representative RT-qPCR analysis of ABC transporter family members in sorted Fluo⁻ and Fluo⁺ cells of different primary PDAC PDX-derived *in vitro* cultures ($n = 3$). mRNA for *ABCB1* and *ABCB5* was nondetectable (n.d.). Data are normalized for β -actin expression and represent the relative ratio Fluo⁺ vs. Fluo⁻. Inset, western blot for ABCG2 transporter in sorted Fluo⁻ and Fluo⁺ cells. (c) Autofluorescence content in cells during and after FTC treatment. Data are representative of different primary PDAC PDX-derived *in vitro* cultures ($n = 3$, each performed in triplicate). (d) Representative images of mitoxantrone (MTX) accumulation in cytoplasmic vesicles of sorted Fluo⁻ and Fluo⁺ cells in the presence or absence of 10 μ M FTC. (e) Representative confocal images of ABCG2-mCherry expression in Fluo⁻ and Fluo⁺ cells. (f) Top, flow cytometry analysis of autofluorescence content in different PDAC PDX-derived *in vitro* cultures ($n = 2$) sorted for Fluo⁻ cells and subsequently infected with a control or ABCG2-overexpressing lentiviral vector. Bottom, RT-qPCR analysis of pluripotency-associated gene expression in sorted Fluo⁻ control or ABCG2-overexpressing cells of different primary PDAC PDX-derived *in vitro* cultures ($n = 2$) each performed in duplicate. Data are normalized for β -actin expression and represent the relative ratio Fluo⁺ vs. Fluo⁻. (g) RT-qPCR analysis of pluripotency-associated gene expression in sorted Fluo⁺ cells of different primary PDAC PDX-derived *in vitro* cultures ($n = 2$) treated with FTC, each performed in duplicate. Data are normalized for β -actin expression and represent the relative ratio Fluo⁺ vs. Fluo⁻. (h) Representative flow cytometry plots for side-population analysis in sorted Fluo⁻ and Fluo⁺ cells from different PDAC PDX-derived cultures ($n = 3$, each performed in duplicate). ABCG2 was specifically inhibited by FTC. Error bars (b,c,f,g), s.d. ($n =$ biological replicates). Statistical significance was assessed by Mann-Whitney test, * $P < 0.05$, ** $P < 0.01$.

Annexin V staining revealed that PDX-derived *in vitro*-cultured non-autofluorescent cells were more rapidly forced into apoptosis compared to autofluorescent cells (Fig. 4d). We investigated two possible reasons for their enhanced chemoresistance. First, we compared the *in vitro* cycling profile of autofluorescent and non-autofluorescent cells using the label-retaining dye PKH26 (ref. 23) and observed that autofluorescent cells had a significantly ($P < 0.05$) higher label-retaining and slow-cycling phenotype (Fig. 4e). The slow-cycling profile was confirmed *in vivo* as we found a marked enrichment

in the G0 population in autofluorescent cells (12.7%) compared to the non-autofluorescent cells (5.15%) from freshly harvested tumors (Fig. 4f). Second, RT-qPCR analysis for the human concentrative nucleoside (hCNT) and equilibrative nucleoside (hENT) transporters, on which gemcitabine uptake depends²⁴, showed significantly ($P < 0.05$) lower expression in autofluorescent cells (Fig. 4g). Thus, the inherent chemoresistant nature of autofluorescent cells could be explained, in part, by their slow-cycling nature and the lack of expression of gemcitabine transporters.

mechanism of autofluorescence

Because we noted that autofluorescent cells contained higher intracellular ATP levels (Supplementary Fig. 10a), and inhibition of ATP production by different means (oligomycin, rotenone and 2,4-dinitrophenol (DNP)) resulted in reproducible

but reversible loss of autofluorescence (Fig. 5a), we reasoned that the subcellular autofluorescent compartment is driven by ATP-dependent transporters. Indeed, autofluorescent cells over-expressed ABCG2 at both the mRNA and protein level (Fig. 5b), whereas other ATP transporters were not differentially expressed.

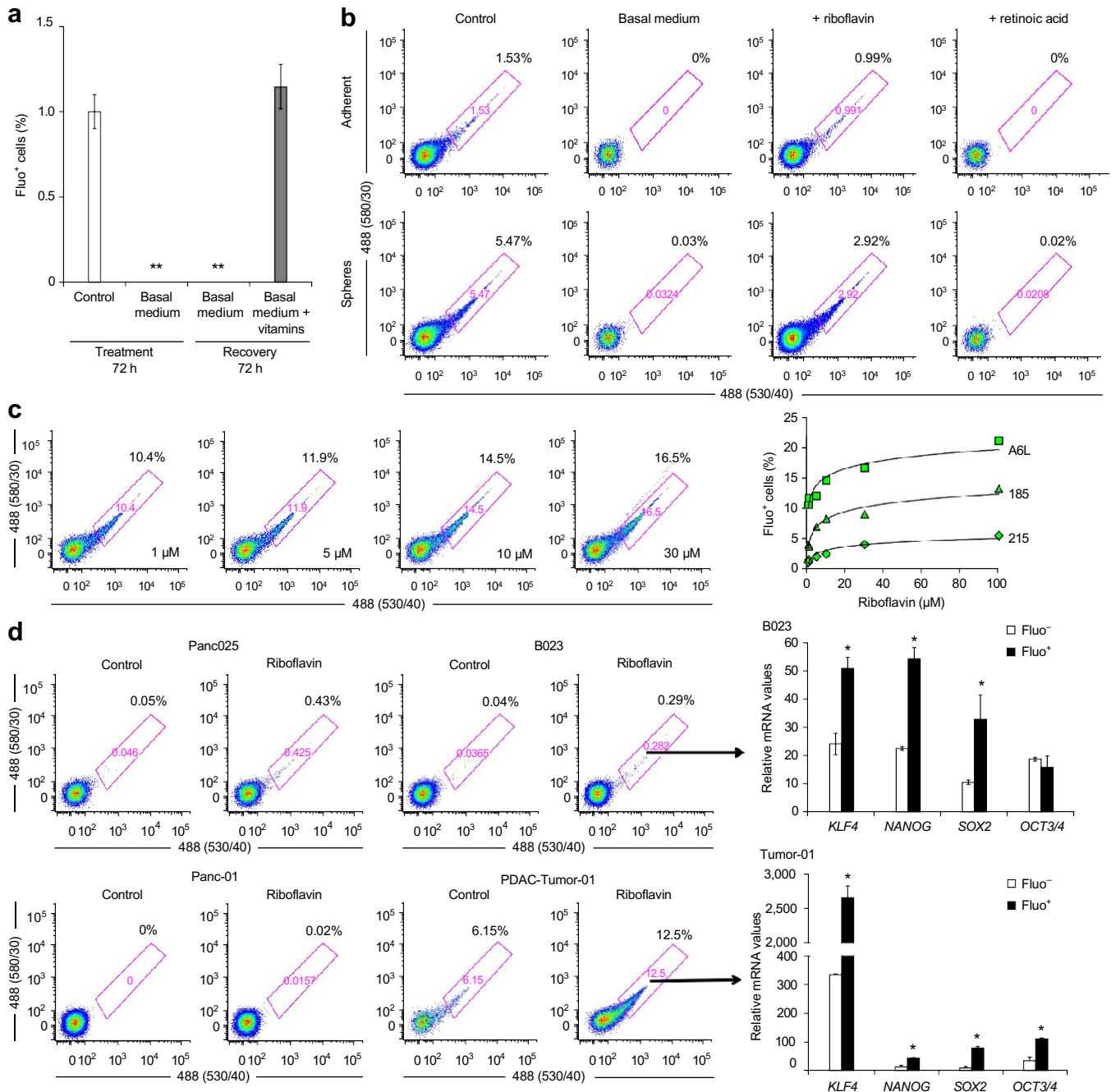


Figure 6 | Source of autofluorescence in cancer stem cells. (a) Percentage of autofluorescent cells in different primary PDAC PDX-derived *in vitro* cultures ($n = 3$, each performed in triplicate) in basal medium (without vitamins) or after recovery in basal medium or basal medium plus vitamin cocktail. **(b)** Representative flow cytometry analysis of autofluorescence recovery in adherent cells or spheres cultured in basal medium or basal medium supplemented with riboflavin or retinoic acid (each at 1 μM). Plots shown are representative of different primary PDAC PDX-derived *in vitro* cultures ($n = 3$, each performed in duplicate). **(c)** Representative flow cytometry plots of cells exposed to increasing concentrations of riboflavin (left), whose effects were quantified for different primary PDAC PDX-derived *in vitro* cultures (right) ($n = 3$). **(d)** Left, representative flow cytometry plots of primary PDAC PDX-derived *in vitro* cultures (Panc025 and B023), the Panc01 cell line and a freshly digested patient tumor (PDAC-Tumor-01), with and without riboflavin treatment (30 μM). Right, RT-qPCR analysis of pluripotency-associated gene expression in Fluo⁻ and Fluo⁺ cells from the indicated samples after exposure to 30 μM riboflavin, each performed in duplicate. Data are normalized for β -actin expression and represent the relative mRNA values of Fluo⁺ and Fluo⁻. Error bars (a,d), s.d. ($n =$ biological replicates). Statistical significance was assessed by Mann-Whitney test, * $P < 0.05$, ** $P < 0.01$.

Moreover, functionally inhibiting ABCG2 with fumitremorgin C (FTC)²⁵ reversibly eliminated autofluorescence (Fig. 5c), and the ABCG2-dependent fluorescent type II topoisomerase inhibitor mitoxantrone²⁶ colocalized with the autofluorescent compartment, suggesting expression of ABCG2 transporters in membranes of autofluorescent vesicles (Fig. 5d). The latter was confirmed by colocalization studies using a lentiviral vector expressing an ABCG2-mCherry fusion protein. ABCG2 transporters were expressed in the membrane of autofluorescent vesicles, whereas in non-autofluorescent cells, ABCG2 expression was detected throughout the cytoplasm and in the cell membrane (Fig. 5e). Notably, in non-autofluorescent cells, overexpression of ABCG2 did not result in *de novo* generation of autofluorescent vesicles, nor did these cells acquire stem-like features such as the overexpression of pluripotency-associated genes (Fig. 5f). As these results suggested that ABCG2 activity was not responsible for the stem-like phenotypes of CSCs, we next functionally inhibited ABCG2 transporter activity with FTC in autofluorescent cells and did not observe any effect on the expression of pluripotency-associated genes or on their sphere-formation capacity (Fig. 5g and data not shown). Last, because overexpression of ABCG2 is a defining feature of SP cells, we investigated the distribution of autofluorescent cells between SP and non-SP cells. Interestingly, autofluorescent cells were more prominent in the non-SP population, indicating that autofluorescent cells represent a distinct ABCG2-expressing subpopulation that is phenotypically different from SP cells (Fig. 5h), which is consistent with our earlier finding that SP-sorted cells were not enriched for CSCs (Supplementary Fig. 2).

source of autofluorescence

Autophagy has been related to autofluorescence²⁷, but neither expression levels of the autophagy-related gene *ATG12* or autophagy related-protein LC3 nor inhibition or induction of autophagy supported this notion in our studies (Supplementary Fig. 10b,c). On the basis of these results and our observations that autofluorescent vesicles colocalized with ABCG2, we reasoned that a specific ABCG2-transported substrate may be responsible for autofluorescence. Culturing primary PDAC PDX-derived cells in basal medium (vitamin free) resulted in complete loss of autofluorescence within 72 h (Fig. 6a; medium composition in Supplementary Table 1), and although basal medium alone could not restore autofluorescence following FTC treatment, basal medium supplemented with a cocktail of essential vitamins could (Fig. 6a), suggesting that one or more vitamins were responsible for the autofluorescence. Using the spectroscopic profile of the autofluorescent CSC feature (excitation, 490 nm; emission, 532 nm (490/532 nm); Fig. 1a), we identified retinoic acid (330/500 nm) and riboflavin (450/520 nm) as the most likely candidates, as both have overlapping spectroscopic profiles. Interestingly, only riboflavin (vitamin B₂), which is a specific substrate for ABCG2 (ref. 28), was capable of restoring autofluorescence (Fig. 6b and Supplementary Fig. 10d). In addition, we found a strong association between the percentage of autofluorescence and riboflavin concentration in the medium, which plateaued at 30 μ M (Fig. 6c). Notably, the enhanced number of autofluorescent cells was not related to nonspecific enrichment of non-CSCs, as riboflavin-enriched autofluorescent cells maintained high expression of pluripotency-associated genes and enhanced

in vivo tumorigenicity (Supplementary Fig. 11a). By adding 30 μ M riboflavin in the isolation medium, we further enhanced our ability to isolate otherwise dim autofluorescent cells from primary PDAC PDX-derived *in vitro* cultures (Panc025 and B023) and from freshly digested tumors, in which autofluorescence was difficult to detect with regular riboflavin concentrations (Fig. 6d). However, the addition of riboflavin to the isolation medium did not enable better identification of the autofluorescent CSC subpopulation in established cell lines such as Panc-1, which did not display autofluorescence under *in vitro* or *in vivo* conditions or following exposure to riboflavin (Fig. 6d, Supplementary Fig. 11b and data not shown). This result further corroborates the notion that higher riboflavin concentrations do not nonspecifically create autofluorescent cells and highlights the importance of working with early-passage primary cells for studying CSCs.

discussion

In this study, we identified a distinct autofluorescent population of self-renewing and highly tumorigenic CSCs with an inherent ability to concentrate the fluorescent vitamin riboflavin in ABCG2-coated intracellular vesicles. This feature allowed for the reliable identification and isolation of CSCs from the large pool of nontumorigenic cancer cells independent of the expression of CD133 or other surface markers and across numerous different tumor types (PDAC, HCC, NSCLC and CRC). These autofluorescent cells could be identified by flow cytometry, purified from primary cultures or freshly digested tumors by FACS without the need for fixation or antibody staining, and sensitivity could be further enhanced by the addition of riboflavin. Autofluorescent cells bore all defining CSC features: they overexpressed pluripotency-associated genes, harbored self-renewal capacity, were highly chemoresistant and metastatic and, most notably, demonstrated *in vivo* tumorigenicity. Autofluorescent cells also demonstrated exclusive long-term self-renewal capacity as shown by serial *in vivo* transplantation and tumorigenicity at the single-cell level.

After we excluded autophagy as a putative source of autofluorescence, we found ABCG2 to be overexpressed in autofluorescent cells. ABCG2 is strongly induced in the mammary gland during lactation^{29,30} and responsible for pumping riboflavin (vitamin B₂) into milk²⁸. ABCG2 also has broad substrate specificity for various drugs (for example, mitoxantrone), carcinogens and dietary toxins^{28–31}. Using autofluorescent CSCs isolated from a variety of carcinomas, we found a marked concentration of riboflavin and mitoxantrone in intracellular vesicles, which was specifically mediated by ABCG2 transporters. The intracellular autofluorescent vesicles identified were coated in ABCG2 and could be depleted using the ABCG2 inhibitor FTC, confirming that the autofluorescent phenotype was indeed ABCG2 dependent. Apart from acting as a sink for ABCG2-dependent substrates, such as mitoxantrone and riboflavin, we did not, however, identify a CSC-specific advantage associated with the presence of autofluorescent phenotype. For example, elimination of the autofluorescent phenotype by functionally inhibiting ABCG2 with FTC did not affect CSC phenotypes such as the expression of pluripotency-associated genes or *in vivo* tumorigenicity. Likewise, overexpression of ABCG2 in non-CSCs did not result in *de novo* formation of autofluorescent vesicles, nor did we observe an increase in CSC features in these cells. Thus, although we cannot definitively state

that autofluorescent vesicles are not biologically important for CSCs, our data do not support a functional role for the accumulation of riboflavin in these vesicles in the biology of CSCs.

Although more in-depth studies are still needed to determine the exact role of autofluorescent vesicles in CSC biology, our studies unequivocally demonstrate that, regardless of their function, autofluorescent vesicles can be used to track, isolate and purify CSCs from several epithelial cancers without the use of antibodies or extensive manipulation. Further exploiting autofluorescence may allow us to more accurately capture the dynamic complexity of CSCs as well as develop new CSC-specific therapies.

methods

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

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

Acknowledgments

We thank A. González-Neira and L. Moreno from the CNIO Human Genotyping-CEGEN Unit for performing the TaqMan OpenArray SNAP analysis, and S.M. Trabulo and M. Tatari for their excellent *in vivo* technical assistance. The research was supported by the ERC Advanced Investigator Grant (Pa-CSC 233460), the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 256974 (EPC-TM-NET) and no. 602783 (CAM-PaC), the Associazione Italiana Ricerca Cancro (AIRC grant no. 12182 to A.S.), the Italian Cancer Genome Project Ministry of University and Research (FIRB RBAP10AHJB to A.S.), the FIMP-Italian Ministry of Health (CUP_J33G13000210001), the Subdirección General de Evaluación y Fomento de la Investigación, Fondo de Investigación Sanitaria (PS09/02129 & PI12/02643) and the Programa Nacional de Internacionalización de la I+D, Subprograma: FCCI 2009 (PLE2009-0105; both Ministerio de Economía y Competitividad (es), Spain). M.C. was supported by the La Caixa Predoctoral Fellowship Program.

Author contributions

I.M.-L. acquired, analyzed and interpreted data as well as developed the study concept and drafted the manuscript; J.D. and A.B. performed the *in vivo* experiments; E.L., S.A., M.C., A.G.S. and S.Z. acquired and analyzed *in vitro* data; J.C.-T. designed the microchip-based single-cell assay; D.M. characterized the autofluorescence signal by confocal microscopy and developed the protocol for automated analysis of autofluorescence; M.H., M.E., J.K. and A.S. provided extensively characterized PDAC samples; B.S. and C.H. developed the study concept, obtained funding, interpreted the data and wrote the manuscript.

Competing financial interests

The authors declare no competing financial interests.

References

- Clarke, M.F. et al. Cancer stem cells—perspectives on current status and future directions: AACR Workshop on cancer stem cells. *Cancer Res.* **66**, 9339–9344 (2006).
- Campbell, L.L. & Polyak, K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle* **6**, 2332–2338 (2007).
- Balic, A., Dorado, J., Alonso-Gómez, M. & Heeschen, C. Stem cells as the root of pancreatic ductal adenocarcinoma. *Exp. Cell Res.* **318**, 691–704 (2012).
- Hermann, P.C., Bhaskar, S., Cioffi, M. & Heeschen, C. Cancer stem cells in solid tumors. *Semin. Cancer Biol.* **20**, 77–84 (2010).
- Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S. & Clarke, M.F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* **100**, 3983–3988 (2003).
- Gallmeier, E. et al. Inhibition of ataxia telangiectasia- and Rad3-related function abrogates the *in vitro* and *in vivo* tumorigenicity of human colon cancer cells through depletion of the CD133⁺ tumor-initiating cell fraction. *Stem Cells* **29**, 418–429 (2011).
- Wicha, M.S., Liu, S. & Dontu, G. Cancer stem cells: an old idea—a paradigm shift. *Cancer Res.* **66**, 1883–1890 (2006).
- Beier, D. et al. CD133⁺ and CD133⁻ glioblastoma-derived cancer stem cells show differential growth characteristics and molecular profiles. *Cancer Res.* **67**, 4010–4015 (2007).
- Joo, K.M. et al. Clinical and biological implications of CD133-positive and CD133-negative cells in glioblastomas. *Lab. Invest.* **88**, 808–815 (2008).
- Ogden, A.T. et al. Identification of A2B5⁺CD133⁻ tumor-initiating cells in adult human gliomas. *Neurosurgery* **62**, 505–514 (2008).
- Wang, J. et al. CD133 negative glioma cells form tumors in nude rats and give rise to CD133 positive cells. *Int. J. Cancer* **122**, 761–768 (2008).
- Hermann, P.C. et al. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell* **1**, 313–323 (2007).
- Zhang, S.N., Huang, F.T., Huang, Y.J., Zhong, W. & Yu, Z. Characterization of a cancer stem cell-like side population derived from human pancreatic adenocarcinoma cells. *Tumori* **96**, 985–992 (2010).
- Zhou, J. et al. Persistence of side population cells with high drug efflux capacity in pancreatic cancer. *World J. Gastroenterol.* **14**, 925–930 (2008).
- Kabashima, A. et al. Side population of pancreatic cancer cells predominates in TGF- β -mediated epithelial to mesenchymal transition and invasion. *Int. J. Cancer* **124**, 2771–2779 (2009).
- Broadley, K.W. et al. Side population is not necessary or sufficient for a cancer stem cell phenotype in glioblastoma multiforme. *Stem Cells* **29**, 452–461 (2011).
- Burkert, J., Otto, W.R. & Wright, N.A. Side populations of gastrointestinal cancers are not enriched in stem cells. *J. Pathol.* **214**, 564–573 (2008).
- Clément, V. et al. Retraction: Marker-independent identification of glioma-initiating cells. *Nat. Methods* **10**, 1035 (2013).
- Lonardo, E. et al. Nodal/Activin signaling drives self-renewal and tumorigenicity of pancreatic cancer stem cells and provides a target for combined drug therapy. *Cell Stem Cell* **9**, 433–446 (2011).
- Iliopoulos, D., Hirsch, H.A., Wang, G. & Struhl, K. Inducible formation of breast cancer stem cells and their dynamic equilibrium with non-stem cancer cells via IL6 secretion. *Proc. Natl. Acad. Sci. USA* **108**, 1397–1402 (2011).
- Hermann, P.C. & Heeschen, C. Metastatic cancer stem cells—*quo vadis?* *Clin. Chem.* **59**, 1268–1269 (2013).
- Hermann, P.C. et al. Multimodal treatment eliminates cancer stem cells and leads to long-term survival in primary human pancreatic cancer tissue xenografts. *PLoS ONE* **8**, e66371 (2013).
- Cicalese, A. et al. The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. *Cell* **138**, 1083–1095 (2009).
- Santini, D. et al. Prognostic role of human equilibrative transporter 1 (hENT1) in patients with resected gastric cancer. *J. Cell. Physiol.* **223**, 384–388 (2010).
- Rabindran, S.K., Ross, D.D., Doyle, L.A., Yang, W. & Greenberger, L.M. Fumitremogin C reverses multidrug resistance in cells transfected with the breast cancer resistance protein. *Cancer Res.* **60**, 47–50 (2000).
- Bell, D.H. Characterization of the fluorescence of the antitumor agent, mitoxantrone. *Biochim. Biophys. Acta* **949**, 132–137 (1988).
- White, E. Deconvoluting the context-dependent role for autophagy in cancer. *Nat. Rev. Cancer* **12**, 401–410 (2012).
- van Herwaarden, A.E. et al. Multidrug transporter ABCG2/breast cancer resistance protein secretes riboflavin (vitamin B2) into milk. *Mol. Cell. Biol.* **27**, 1247–1253 (2007).
- Jonker, J.W. et al. The breast cancer resistance protein protects against a major chlorophyll-derived dietary phototoxin and protoporphyrin. *Proc. Natl. Acad. Sci. USA* **99**, 15649–15654 (2002).
- van Herwaarden, A.E. & Schinkel, A.H. The function of breast cancer resistance protein in epithelial barriers, stem cells and milk secretion of drugs and xenotoxins. *Trends Pharmacol. Sci.* **27**, 10–16 (2006).
- Ifergan, I., Goler-Baron, V. & Assaraf, Y.G. Riboflavin concentration within ABCG2-rich extracellular vesicles is a novel marker for multidrug resistance in malignant cells. *Biochem. Biophys. Res. Commun.* **380**, 5–10 (2009).

32. Mueller, M.T. *et al.* Combined targeted treatment to eliminate tumorigenic cancer stem cells in human pancreatic cancer. *Gastroenterology* **137**, 1102–1113 (2009).
33. Qin, D., Xia, Y. & Whitesides, G.M. Soft lithography for micro- and nanoscale patterning. *Nat. Protoc.* **5**, 491–502 (2010).
34. Sainz, B. Jr. *et al.* Identification of the Niemann-Pick C1-like 1 cholesterol absorption receptor as a new hepatitis C virus entry factor. *Nat. Med.* **18**, 281–285 (2012).
35. Sainz, B. Jr., Barretto, N. & Uprichard, S.L. Hepatitis C virus infection in phenotypically distinct Huh7 cell lines. *PLoS ONE* **4**, e6561 (2009).
36. Greenwood, M. & Yule, G.U. On the statistical interpretation of some bacteriological methods employed in water analysis. *J. Hyg. (Lond.)* **16**, 36–54 (1917).
37. Taswell, C. Limiting dilution assays for the determination of immunocompetent cell frequencies. I. Data analysis. *J. Immunol.* **126**, 1614–1619 (1981).

online methods

Primary human cancer cells. Human PDAC tumors as well as CRC, HCC and NSCLC tumors were obtained with written informed consent (PA 34_2012; Ethics Committee of the Instituto de Salud Carlos III; Madrid, Spain and Comunidad de Madrid, Spain (C.P. CNIO-CTC-11 - C.I. 11/103-E)), minced, mechanically digested (gentleMACS Dissociator; Miltenyi), enzymatically digested with collagenase (Stem Cell Technologies) and subsequently cultured *in vitro* as previously detailed³². PDAC patient-derived xenografts (PDAC PDXs) (185, A6L, 286, 266, 215, JH029, PANC025, B06, B023, 354, 247, CRC025, CRC024, HCC-5, HCC-6) were processed as previously described³². Within the paper, primary PDAC PDX-derived *in vitro* cultures are referred to by their patient number designation (for example: 185, A6L, etc.), freshly digested noncultured PDXs are referred to as 'PDAC PDX' and freshly digested noncultured primary patient tumors are referred to as 'PDAC-Tumor'. Primary cultures were tested for mycoplasma at least every 4 weeks (Gen-Probe-Mycoplasma Tissue Culture- NI Rapid Detection System MTC-NI).

Flow cytometry and cell sorting. Human PDAC tumors, as well as CRC, HCC and NSCLC cells or dissociated cells from primary sphere cultures were adjusted to a concentration of 10^6 cells/ml in Sorting buffer [1X PBS; 3% FBS (v/v); 3 mM EDTA (v/v)] before analysis or sorting with a FACS Canto II or FACS Influx instrument, respectively (BD). Autofluorescent cells were excited with a 488 nm blue laser and selected as the intersection with filters 530/40 and 580/30 as shown in **Figure 1c**. Some non-autofluorescent cells are lost during sorting due to the gate settings that maintain an appropriate distance between autofluorescent and non-autofluorescent cells. However, such settings are required to ensure high purity during sorting. To characterize autofluorescent cells, the following antibodies were used: anti-CD133/1-APC (Miltenyi Biotec); EpCAM-APC, CD44-APC, SSEA-1-APC, CXCR4-APC or appropriate isotype-matched control antibodies (all from BD) and anti-ALK4 (Cell Signaling Technology, Inc.). DAPI was used for exclusion of dead cells. Annexin V-APC (BD) staining was performed according to the manufacturer's instructions. Data were analyzed with FlowJo 9.2 software (Tree Star).

Protocol for acquiring and sorting autofluorescent cells. We used the following protocol to acquire and sort autofluorescent cells from tumors or cultured cells.

1. Mince the tumor into small pieces.
2. Mechanically digest with a gentleMACS Dissociator (Miltenyi).
3. Incubate with collagenase (Stem Cell Technologies) for 60 min at 37 °C and centrifuge for 5 min at 300g.
4. Resuspend cell pellets in 10 ml of RPMI (Invitrogen) supplemented with 10% FBS and 50 units/ml penicillin/streptomycin (referred to hereafter as 'complete medium').
5. Filter the suspension through a 70- μ m strainer and centrifuge for 5 min at 300g.
6. Resuspend the cell pellet in 10 ml of RPMI complete media and add 30 μ M Riboflavin (Sigma) to the suspension. (Note: hereafter is the starting point for cultured primary cells, including an optional addition of 30 μ M of riboflavin to the medium.)

7. Incubate at least 2 h or overnight at 37 °C.
8. Centrifuge at 300g for 5 min and resuspend cell pellets in 1 \times PBS for acquisition or in sorting buffer (1 \times PBS containing 3% (v/v) FBS and 2 mM EDTA) for FACS.
9. Analyze or sort autofluorescent cells using standard flow cytometer equipment (here we used FACS Canto II flow cytometer and BD Influx cell sorter; both BD). Autofluorescent cells are excited with a 488-nm blue laser and best selected as the intersection with filters 530/40 and 580/30 (or similar filters), and a proper distance between gates for autofluorescent and non-autofluorescent cells is required to achieve high purity during sorting. Example cytometry plots including the gating strategy are shown in **Supplementary Figure 6a**.
10. If additional immunophenotyping is required, autofluorescence can be easily covisualized with fluorochromes excitable with a 640-nm red laser without the need for compensations. When combined with antibodies, conjugated PE or alike, autofluorescence can be covisualized with a 488-blue laser (compensations may be required) or using a 561-nm yellow-green laser without the need for compensation as autofluorescence is not excited with this laser.

Side population. Human primary PDAC cells at a concentration of 10^6 cells/ml were stained with Hoechst 33342 (5 μ g/ml) at 37 °C for 2 h in the absence or presence of the ABCG2 transporter inhibitor fumitremorgin C (5 μ g/ml; Sigma). Cells were washed and resuspended in cold 1 \times PBS. Propidium iodide (Sigma) was used to exclude dead cells. SP and non-SP cells were sorted using a FACS Influx sorter (BD).

Immunofluorescence. Primary pancreatic cancer cells and sphere-derived cells were seeded in 96-well culture dishes (Corning) and incubated at 37 °C for 24 h. For tracking the cytosol, mitochondria, lysosomes and lipid droplets, CytoTracker, MitoTracker, LysoTracker, and Nile red, respectively, were used at dilutions of 1:20,000 for 30 min at 37 °C (all from Invitrogen). Following two washes with 1 \times PBS, and 5 min incubation with Hoechst stain (5 μ g/ml; Sigma), cells were analyzed using an SP5 confocal microscope (Leica).

Sphere formation assay. Spheres were generated by culturing $\sim 2 \times 10^4$ human primary PDAC cells in suspension in serum-free DMEM/F12 supplemented with B27 (1:50, Invitrogen), 20 ng/ml bFGF and 50 units/ml pen/strep for a total of 7 d, allowing spheres to reach a size of $>75 \mu$ m. For serial passaging, 7-d-old spheres were harvested using 40- μ m cell strainers, dissociated into single cells and then recultured for 7 additional days as previously described¹⁹.

RNA preparation and qRT-PCR. Total RNAs from human primary pancreatic cancer cells or livers of NSG mice were extracted with Trizol (Life Technologies) according to the manufacturer's instructions. 1 μ g of total RNA was reverse transcribed with SuperScript II reverse transcriptase (Life Technologies) using random hexamers. Quantitative real-time PCR was performed with an Applied Biosystems 7500 real-time thermocycler (Applied Biosystems) using Fast SYBR Green (Qiagen) as per the

manufacturer's instructions. The list of primers used can be found in **Supplementary Table 2**.

Microchip-based single-cell analysis. The microchip pattern was designed with AutoCAD (Autodesk) and was manufactured using standard soft-lithography techniques³³. This post-array is made of poly(dimethylsiloxane) and contains ~4,600 microwells of 80- μ m diameter. The chip was bonded to a 24-well glass-bottom plate using a plasma oven before cell seeding (**Fig. 2d**). For semiautomated analysis of single cells seeded in microwells, we used a custom-developed software (CNIO MSRC). During a first low-resolution fast scan, settings for generating one image per well were established. After image acquisition, the software automatically localizes nanovolume wells containing single cells and records the presence or absence of autofluorescence and the respective well coordinates. Using this spatial information, the application interacts with the SP5 microscope (Leica) and loads high-resolution settings, scanning automatically just the areas of interest at customizable time intervals.

In vivo tumorigenicity and metastasis assays. Human primary PDAC, CRC, HCC and NSCLC cells were sorted for autofluorescence as detailed below. For tumorigenicity assays, serial dilutions of single cells resuspended in Matrigel (BD Biosciences) were subcutaneously injected into 4- to 5-week-old female nude mice (Hsd:ATHymic Nude-Foxn1^{nu}/Foxn1⁺; Harlan Laboratories) and tracked for 3 months for tumor formation because in previous experiments lasting 6 months, we did not observe any tumor formation, specifically with non-autofluorescent cells, after 3 months. For metastasis assays, 10³ FACS-sorted autofluorescent and non-autofluorescent cells, with mCherry labeling directed by histone H2B type 2-E (HIST2H2BE), were resuspended in 1 \times PBS and intrasplenically injected into 4- to 5-week-old female NSG mice (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ; Charles River) as previously described³⁴. All animal experiments were conducted in accordance with FELASA guidelines and approved protocols. Animal treatments were approved (Protocol PA 34_2012) by the Animal Experimental Ethics Committee of the Instituto de Salud Carlos III (Madrid, Spain).

PKH26 assay. Human primary PDAC cells were labeled with PKH26 red fluorescent cell membrane labeling dye (Sigma) according to the manufacturer's instructions. Every 7 d, cells were harvested and PKH26⁺ cells were determined using a FACS Canto II (BD) for a total of 3 weeks.

Invasion and migration assays. Invasion assays were performed using modified Boyden chambers filled with Matrigel (BioCoat, BD Biosciences). Human primary pancreatic cancer cells were added to the Matrigel coated inserts, and 750 μ l of serum-free medium supplemented with recombinant human Nodal (300 ng/ml), SDF-1 (100 ng/ml), Shh (100 ng/ml), or TGF- β (10 ng/ml) was added to the lower chamber. Cells were incubated for 22 h at 37 $^{\circ}$ C, and invaded cells were fixed in 4% PFA and stained with 4',6-diamidino-2-phenylindole (DAPI). The ratio of cells in the lower chamber versus total seeded cells (as a percentage) was calculated. The data are presented as the mean of ten high-power fields (HPF).

Histology and immunohistochemistry. For histopathological analysis, FFPE blocks were serially sectioned (3 μ m thick) and stained with hematoxylin and eosin (H&E). Additional serial sections were probed with antibodies against DsRed (Clontech), human cytokeratin 19 α (Abcam), or *in situ* hybridization was performed using the Alu probe (Qbiogene). Following incubation with primary antibodies, samples were incubated with HRP-conjugated secondary antibodies (Dako), and positive cells were visualized using 3,3-diaminobenzidine tetrahydrochloride plus (DAB⁺) as a chromogen.

Western blot analysis. Cells were harvested in RIPA buffer (Sigma) supplemented with a protease inhibitor cocktail (Roche Applied Science). 50 μ g of protein were resolved by SDS-PAGE and transferred to PVDF membranes (Amersham Pharmacia). Membranes were sequentially blocked with 1 \times TBS containing 5% BSA (w/v), 1% chicken albumin (w/v) and 0.1% Tween 20 (v/v), incubated with a 1:1,000 dilution of antibodies against ABCG2 (ab24115; Abcam), NANOG (D73G4; Cell Signaling), OCT3/4 (H-134. Sc-9081; Santa Cruz Biotechnology), Tubulin (Sigma) or GAPDH (ab8245; Abcam) overnight at 4 $^{\circ}$ C, washed three times with 1 \times PBS containing 0.05% Tween 20 (v/v), incubated with horseradish peroxidase-conjugated goat anti-rat or goat anti-mouse antibody (Sigma), and washed again to remove unbound antibody. Bound antibody complexes were detected with SuperSignal chemiluminescent substrate (Amersham).

Lenti-mCherry-C1-huABCG2. The human ABCG2-mCherry lentiviral expression plasmid (pLVX-mCherry-C1-huABCG2) was cloned by PCR amplifying the human ABCG2 coding sequence (cds) from pOTB7-ABCG2 (CNIO, Genomics Unit collection) with primers 5'-GCC TCG AGC GTC TTC CAG TAA TGT CGA AG-3' (sense) and 5'-GC GGA TCC AGA ATA TTT TTT AAG AAA TAA C-3' (anti-sense). This PCR product was subcloned into pGEM-T Easy (Promega Biotech Ibérica S.L.) for sequence verification and then cloned as a XhoI/BamHI fragment into XhoI/BamHI-digested pLVX-mCherry-C1 (Clontech). The resulting construct contained the human ABCG2 cds ligated in frame to the C-terminal end of the mCherry reporter, so that an ABCG2-mCherry fusion protein was made.

Lentiviruses were produced as previously described³⁵. Briefly, viruses were generated by cotransfection of the pLVX-mCherry-C1-huABCG2 lentivirus expression plasmid, packaging plasmid (Pax2) and envelope plasmid (pCDNA3.1-VSVG) into 293T producer cells. Supernatants were collected 48 h post transfection, filtered through a 0.45- μ m-pore-size filter (BD Biosciences), concentrated by centrifugation for 1 h at 10,062g, divided into aliquots, frozen and subsequently titered by flow cytometry analysis of mCherry expression in 293T cells infected with increasing dilutions of virus. For infection of primary PDAC cells, autofluorescent-positive and autofluorescent-negative FACS-sorted cells were infected in suspension with Lenti-mCherry-C1-huABCG2 at a multiplicity of infection of ten infectious units per cell.

OpenArray genotyping assay. TaqMan OpenArray technology was used for medium-scale 64 SNPs genotyping array format. One array was performed and analyzed uniformly according to the manufacturer's standard protocol. HapMap and sample duplicates

were included in the array to serve as internal controls and to ensure reproducibility of the results. Genotype calling and sample clustering for OpenArray assays was performed in TaqMan Genotyper Software v1.3 (Applied Biosciences).

Statistical analyses. Results for continuous variables are presented as mean \pm s.d. unless stated otherwise, and significance was determined using the Mann-Whitney test. Analyses were performed using SPSS 22.0. Normal distribution was tested before each statistical analysis was performed. Variance was equally distributed between groups. For statistical analysis of *in vivo* tumorigenicity, we used limiting dilution analysis (LDA; <http://bioinf.wehi.edu.au/software/elda/>). LDA is based on the Poisson single-hit model, which assumes that the number of biological active cells in each group varies according to a Poisson distribution, and a single biologically active cell is sufficient for inducing tumor formation^{36,37}. Studies were designed to allow for the detection of differences of 50% as the minimum requirement.

Considering the power for detecting such a difference to be 80%, s.d. 15% and two groups (Fluo⁺ and Fluo⁻) the appropriate sample size is $n = 3$. For the majority of experiments, the experimental group size contained 4 or more samples.