

## Extensive characterization of sphere models established from colorectal cancer cell lines

Ada Collura · Laetitia Marisa · Diletta Trojan · Olivier Buhard ·  
Anaïs Lagrange · Arnaud Saget · Marianne Bombled · Patricia Méchighel ·  
Mira Ayadi · Martine Muleris · Aurélien de Reynies · Magali Svrcek ·  
Jean-François Fléjou · Jean-Claude Florent · Florence Mahuteau-Betzer ·  
Anne-Marie Faussat · Alex Duval

Received: 27 May 2012/Revised: 3 September 2012/Accepted: 4 September 2012/Published online: 25 September 2012  
© Springer Basel AG 2012

**Abstract** Links between cancer and stem cells have been proposed for many years. As the cancer stem cell (CSC) theory became widely studied, new methods were developed to culture and expand cancer cells with conserved determinants of “stemness”. These cells show increased ability to grow in suspension as spheres in serum-free medium supplemented with growth factors and chemicals. The physiological relevance of this phenomenon in established cancer cell lines remains unclear. Cell lines have traditionally been used to explore tumor biology and serve as preclinical models for the screening of potential therapeutic agents. Here, we grew cell-forming spheres (CFS) from 25 established colorectal cancer cell lines. The molecular and cellular characteristics of CFS were compared to the bulk of tumor cells. CFS could be isolated from 72 % of the cell lines. Both CFS and their parental

CRC cell lines were highly tumorigenic. Compared to their parental cells, they showed similar expression of putative CSC markers. The ability of CRC cells to grow as CFS was greatly enhanced by prior treatment with 5-fluorouracil. At the molecular level, CFS and parental CRC cells showed identical gene mutations and very similar genomic profiles, although microarray analysis revealed changes in CFS gene expression that were independent of DNA copy-number. We identified a CFS gene expression signature common to CFS from all CRC cell lines, which was predictive of disease relapse in CRC patients. In conclusion, CFS models derived from CRC cell lines possess interesting phenotypic features that may have clinical relevance for drug resistance and disease relapse.

**Keywords** Colorectal cancer · Colon cancer cell lines · Spheres · Microarray

**Electronic supplementary material** The online version of this article (doi:10.1007/s00018-012-1160-9) contains supplementary material, which is available to authorized users.

A. Collura · D. Trojan · O. Buhard · A. Lagrange · A. Saget ·  
P. Méchighel · M. Muleris · M. Svrcek · J.-F. Fléjou · A. Duval  
Inserm, UMRS\_938, Centre de Recherche Saint-Antoine,  
Equipe “Instabilité des Microsatellites et Cancers”,  
75012 Paris, France

A. Collura · D. Trojan · O. Buhard · A. Lagrange · A. Saget ·  
P. Méchighel · M. Muleris · M. Svrcek · J.-F. Fléjou ·  
A.-M. Faussat · A. Duval  
Université Pierre et Marie Curie-Paris6, Paris, France

L. Marisa · M. Ayadi · A. de Reynies  
Programme “Cartes d’Identité des Tumeurs”,  
Ligue Nationale Contre le Cancer, Paris, France

M. Bombled · J.-C. Florent · F. Mahuteau-Betzer  
UMR 176 CNRS/Institut Curie, Université Paris-Sud,  
Orsay, France

M. Svrcek · J.-F. Fléjou  
AP-HP, Hôpital Saint-Antoine, Service d’Anatomie et Cytologie  
Pathologiques, Paris, France

M. Svrcek · J.-F. Fléjou  
AP-HP, Hôpital Saint-Antoine, Tumorotheque CancerEst,  
75012 Paris, France

A. Collura (✉) · A. Duval (✉)  
INSERM UMRS 938 Team “Microsatellite Instability and  
Cancer”, Hôpital Saint-Antoine, 184, rue du Faubourg  
Saint-Antoine, 75571 Paris cedex 12, France  
e-mail: ada.collura@inserm.fr

A. Duval  
e-mail: alex.duval@inserm.fr

## Abbreviations

5-FU	5-Fluorouracil
CFS	Cell-forming-spheres
CIN	Chromosomal instability
CRC	Colorectal cancer
CSC	Cancer stem cells
DFS	Disease-free survival
MMR	Mismatch-repair
MSI	Microsatellite instability
MSS	Microsatellite stability

## Introduction

The cancer stem cell (CSC) theory has generated much interest in both the research and clinical communities, notably for colorectal cancer (CRC) [1–3]. According to this hierarchical model, CSCs are defined by their ability to self-renew indefinitely, while also being able to differentiate and generate both tumorigenic and non-tumorigenic daughter cells that constitute the bulk of the tumor [4]. CSCs are thought to have a low rate of division and proliferation that helps them resist various chemotherapies and radiation. Both these forms of treatment preferentially affect highly proliferative cells, thus potentially making CSCs a major reason for the failure of anticancer treatment [5]. In tumors including CRC, the presence and survival of CSCs has been suggested as a key mechanism underlying chemoresistance and disease relapse [6, 7]. This original interpretation of the CSC theory has recently been challenged, however. Some authors have highlighted the plasticity of the CSC phenotype and suggested that it could be induced through dedifferentiation processes influenced by the tumor cell environment [8].

Cancer cell lines have been widely used to explore tumor biology and as preclinical models for the screening of potential therapeutic agents. They are a valuable resource that can be used repeatedly and have also been well characterized with respect to mutational and gene expression profiles [9]. Similar frequencies of gene mutation have been reported in primary tumors and in cancer cell lines derived from the same primary site. Cancer cell lines are not contaminated with stromal tissue, which can sometimes affect the interpretation of data obtained from primary tumors [10]. Furthermore, cancer cell lines often faithfully represent the tumor from which they were isolated [11, 12]. It remains to be determined whether cancer cell lines are relevant biological tools to study the role of CSCs in tumorigenesis. A number of authors have hypothesized the existence of cancer stem-like cells in these cellular models [13, 14]; however, their phenotype is still poorly characterized. Moreover, it is not known

whether cancer stem-like cells from cell lines have any clinical relevance [15, 16].

In order to study cells from cancer cell lines that could display a stem-like phenotype, the first requirement is to have a system in which they can be propagated. The ability to grow in suspension as spheres in serum-free medium supplemented with specific growth factors and chemicals has been described for the expansion of neuronal stem cells [17]. Sphere culture has also been proposed as a valuable method for isolating cancer cells with conserved stemness determinants that are able to propagate in defined media [18–21]. In the present study, we have used this method to grow cell-forming spheres (CFS) from a panel of 25 CRC cell lines. These cell lines were selected to reflect the heterogeneity of CRC in terms of showing microsatellite stability (MSS) or instability (MSI). CRC is a complex tumor entity that includes distinctive molecular phenotypes associated with different clinical features, including response to chemotherapy [22–24]. We investigated the cellular and molecular phenotypes of CFS derived from this panel of CRC cell lines, with particular reference to treatment resistance and CSC features.

## Materials and methods

### Tissue collection and preparation of xenografts

Human colon tissue fragments were obtained in accordance with the ethical standards of the institutional committee on human experimentation from 15 patients undergoing a colon resection for CRC at the Saint-Antoine hospital in Paris. A biobank collection of 30 tumors stored at  $-80^{\circ}\text{C}$  was used to obtain fresh tumor tissue after engraftment in 5-week-old female nude mice (nu/nu). Tumor implantation procedures were performed as previously described [38]. Twelve tumor xenografts were grown for between 1 and 4 months after engraftment. Cancer tissues were intensively washed four times in PBS solution containing antibiotics and then incubated overnight in DMEM/F12 (PAA) containing penicillin (500 U/ml), streptomycin (500  $\mu\text{g/ml}$ ), amphotericin B (0.25  $\mu\text{g/ml}$ ), and ceftazidime (50  $\mu\text{g/ml}$ ). Enzymatic digestion was performed using collagenase (1.5 mg/ml; Sigma) and hyaluronidase (20  $\mu\text{g/ml}$ ; Sigma) in PBS for 1 h. These digests were used for FACS analysis.

### Cell culture

Colon cancer cell lines were cultured in DMEM media supplemented with 10 % FCS (20 % for Caco-2 cell line), 100 U/ml penicillin G, and 100  $\mu\text{g/ml}$  streptomycin. For the culture of CFS, cell lines were grown in serum-free

DMEM/F12 media supplemented with 100 U/ml penicillin G, 100 µg/ml streptomycin, 6 g/l glucose, 1 mg/ml NaHCO<sub>3</sub>, 5 mM HEPES, 2 mM L-glutamine, 4 µg/ml Heparin, 4 mg/ml BSA, 60 mmol/l putrescine, 20 nmol/l progesterone, 30 nmol/l sodium selenite, 25 µg/ml, insulin, 100 µg/ml apo-transferrin, and human recombinant EGF and FGF-2 (Sigma), both at a final concentration of 20 µg/ml (sphere-medium).

#### CFS formation assay

The CFS capacity of colon cancer cell lines tested in this study was derived from monolayer culture or floating culture (for Colo320 colon cancer cell line). It was assessed by plating  $2 \times 10^5$  cells in a T25 flask (8,000 cells/cm<sup>2</sup>). In CFS<sup>Positive</sup> cell lines, CFS were observed 3–7 days after plating. To obtain pure CSC-like cells, the culture period in sphere-medium was extended to 10 passages. To evaluate the CFS capacity of sorted CD166<sup>+</sup>CD44<sup>+</sup>EpCAM<sup>high</sup> or CD166<sup>-</sup>CD44<sup>-</sup>EpCAM<sup>low</sup> HCT116 cell subpopulations (see Fig. 4c, below), 1,000 cells/well were plated in 96-well culture dishes in 200 µl of sphere-medium. The number of CFS in each well was evaluated after 5 days.

#### Proliferation and chemosensitivity assay

Rates of proliferation and sensitivity to 5-FU were assessed using the cell proliferation reagent WST-1 (Roche). Briefly, 10<sup>4</sup> cells of each cell line or from CFS cultures were plated per well in 24-well plates in 2 ml of media with or without 5-FU. After 5 days, WST-1 reagent was added at a 1:10 final dilution and incubated for 4 h at 37 °C. The relative survival fraction of cells was compared between treated and untreated cells.

#### CFS assay following 5-FU treatment

Inoculation of  $2 \times 10^6$  cells into a T75 flask was made with different concentrations of 5-FU to obtain 10 % cell survival after 5 days of incubation. Cells were then washed detached and 10<sup>5</sup> cells were inoculated into a T25 flask with sphere-medium (4,000 cells/cm<sup>2</sup>). After 3 days observation, photographs were taken to determine the proportion of CFS amongst the 5-FU resistant cells. For experimental controls, untreated cells were plated at the same concentration in sphere-medium.

#### Chemical screening

The multi-step strategy used to screen the Institut Curie/CNRS chemical library is shown in Fig. 3a (see below). This bank contains 8,560 compounds in a 96-well format at 10 mg/ml in DMSO (i.e. mean concentration 10 mM). Screening was

performed at 10 and 1 µM final concentrations in 96-well plates and in a final volume of 200 µl of medium. CRC cell lines were incubated with the chemical bank at 700 cells/well for 5 days in standard culture conditions. The Wst-1 assay was used to indirectly estimate cell survival according to the indicated procedure (Roche). Confirmation of the 15 compounds (validation step) was obtained by starting with drug powders in order to reach the correct initial concentration and then performing a cell survival test in 24-well plates with 2 ml of medium. CRC cell lines and CFS were incubated with the chemical bank at 700 and 1,500 cells/well, respectively, for 5 days in standard culture conditions.

#### Subcutaneous transplantation of colon cancer cell lines

Colon cancer cell lines and CFS were suspended in 200 µl PBS-Matrigel (1:1) mixture. They were injected subcutaneously in the flank of 5-week-old nude mice (nu/nu; 1 injection/flank). Experiments were performed in triplicates (3 mice/each sample). Tumor formation was evaluated using a caliper starting on the third week after injection and then weekly for 4 weeks. Animals were sacrificed when the tumor size was between 15 and 20 mm in diameter, or 7 weeks after the injection.

#### Microsatellite analysis

Non-coding microsatellite repeat markers were used to detect instability in five microsatellites (NR27, NR21, NR24, Bat25, and Bat26 comprising the pentaplex PCR system) in CRC cell lines and their CFS counterparts, as previously described [39]. The *TGFBR2*, *BAX*, *MSH3*, and *MSH6* genes containing coding repeats were amplified as previously described [40]. Four other genes containing coding mononucleotide repeat were also amplified with specific primers (sequences available on request). Amplified PCR products were run on an Applied Biosystems PRISM 3100 Genetic Analyzer automated capillary electrophoresis DNA sequencer. Allelic sizes were estimated using gene mapper software (Applied Biosystems).

#### RNA and DNA extraction

Total RNA from CFS and CRC cell lines was extracted using Trizol (Invitrogen) and DNA was extracted using a standard phenol–chloroform procedure. Both RNA and DNA were assessed for integrity and quantity following stringent quality control criteria (CIT program protocols <http://cit.ligue-cancer.net>).

#### Flow cytometry

Flow cytometry was performed on adherent cell lines or CFS cultures after dissociation with accutase (PAA). Cells

were washed once in PBS supplemented with 1 % BSA (Sigma) and resuspended in PBS/1 % BSA at a concentration of  $10^6$  cells/100  $\mu$ l. Cells were stained with IgG-PE/PECy5/FITC/APC (BD Biosciences) to detect non-specific binding of antibodies and autofluorescence. Primary antibodies used were: CD44-FITC 1:75 (clone G44-26; BD Biosciences), CD133-PE 1:100 (clone AC133; Miltenyi Biotec), EpCAM-APC 1:200 (clone HEA-125; Miltenyi Biotec), CD24-FITC 1:100 (clone ML5; BD Biosciences) and CD29-PECy5 1:100 (clone MAR4; BD Biosciences). Cells were incubated for 30 min at 4 °C in the dark and then washed in buffer (PBS, 1 % BSA, 1 mM EDTA). Expression of cell surface markers was detected with a FACScan flow cytometer (BD Biosciences). Cell line suspensions were sorted according to their CD166 and/or CD44 and/or EpCAM expression with a FACS Coulter (BD Biosciences). Separated subpopulations were reanalyzed for purity.

#### Genomic and gene expression arrays and analysis

##### *Data preparation*

Gene expression analysis using arrays was carried out on the IGBMC microarray platform (Strasbourg, France). Total RNA was amplified, labeled, and hybridized to Affymetrix Human Genome U133 plus2 GeneChips following the manufacturer's protocol (Affymetrix, Santa Clara, CA, USA). The chips were scanned with the Affymetrix GeneChip Scanner 3000 and raw intensities were quantified from subsequent images using GCOS 1.4 software (Affymetrix). Data were normalized using the Robust Multi-array Average method and implemented in the R package *affy* [41].

Genomic arrays were performed on the Integragen Platform (Evry, France). DNAs from 13 CFS/adherent cell lines (i.e. LIM2405, HCT8, HCT116, HCT15, TC-7, CO115, RKO, LS411, V9P, HT29, SW620, Colo320, FET) were hybridized on IlluminaSNP HumanCNV610 chips according to instructions provided by the array manufacturer (Illumina, San Diego, CA, USA). Data were normalized and processed as described in supplemental methods [42]. Data are available in the ArrayExpress database ([www.ebi.ac.uk/arrayexpress](http://www.ebi.ac.uk/arrayexpress)).

All analyses were performed using R software (<http://www.R-project.org>)

##### *Unsupervised analysis of gene expression data*

To evaluate the distance between CRC cell lines and CFS, PCA and consensus hierarchical clustering analysis of the 13 pairs (26 samples) were performed on probe-sets present [ $\log_2$  intensity >  $\log_2(3.5)$ ] in at least 5 % of the samples

and having a robust coefficient of variation significantly different from the median variance of all probe sets ( $rCV > 0.05$  and  $P$  value variance test <0.01).

##### *CFS signature*

Genes differentially expressed between CRC cell lines and CFS were assessed using Limma paired moderated  $t$  test [43]. Genes having a  $P$  value <0.001 define the CFS signature ( $n = 359$  probe sets).

##### *CFS functional analysis*

All KEGG pathways and gene sets functionally related to stemness from KEGG, Biocarta, GeneOntology, Molecular Signatures database and Stanford Microarray database were tested for enrichment of up- and down-regulated genes in CFS. Enrichment of the top 100–500 up/down deregulated probe sets was evaluated by computing a hypergeometric test. The median  $P$  value across up/down top probe set lists was used to select pathways and gene sets of interest ( $P$  value <0.05).

##### *Analysis of deregulated regions*

To define up/down regulation regions separately for each CFS/CRC cell line pair, the genome was segmented into overlapping windows of 5 Mb. In each window, the enrichment of up- or down-regulated genes ( $\log_2(FC) > 0.5$ ) for the given pair was assessed by a Fisher test between up/down genes in the windows and up/down regulated genes in the rest of the genome. To compute the frequency across pairs, each region was assigned  $-1$ (down-regulated)/0 (not modified)/1 (up-regulated) depending on the significance ( $P$  value <0.05) of the enrichment.

##### *Survival analysis*

Two publicly available Affymetrix U133P2 datasets with Recurrence Free Survival annotations were used: [25] dataset GSE17536 and GSE17537 comprising 148 samples of Stage II/III CRC), and [26] GSE14333 comprising 99 samples of Stage B and C CRC not contained in the previous dataset. These were normalized by RMA and by clinical center. To evaluate the survival impact of the CFS signature in those datasets, a subset of genes from the original signature (see Supplementary Table S1) was selected based upon their high fold-change [ $|\log_2(CFS/cell\ line)| > 0.8$ ;  $n = 55$ ]. To define a prognostic CFS signature, the 55-gene signature was reduced to probe sets significantly associated with prognosis in the first dataset using univariate Cox models ( $\log$  rank  $P$  value <0.05;  $n = 8$ ). For both CFS signatures, an average expression score per sample was then defined. Normalized intensity values of

selected genes were each centered to zero by subtracting the median expression of each gene. Genes that were down-regulated in CFS were multiplied by  $(-1)$ , thus both up- and down-regulated genes can be used in the score. All genes from the given signature were then averaged. A higher score corresponded to a higher deregulation of CSF genes. The score was then divided into high and low score groups by taking tails of the score distribution in the considered dataset, i.e. 30 % of the highest and 30 % of the lowest scores.

Survival curves were calculated according to the Kaplan–Meier method with an end-point at 5 years. Differences between curves were assessed using the log-rank test. Univariate and multivariate associations for outcome were performed using the Cox regression model.

## Results

The ability of CRC cell lines to grow as CFS in serum-free medium enriched with growth factors and chemical supplements is highly variable

We tested the ability of established MSI and MSS CRC cell lines ( $n = 25$ ) to grow as CFS in sphere-medium. CFS were obtained from 10 MSI and 8 MSS cell lines (total 18/25, 72 %) (Fig. 1a). In the remaining 7 CRC cell lines (4 MSI, 3 MSS), a small number of cells remained afloat but died after 1–4 passages (1–6 weeks) in sphere-medium (data not shown). Morphologically, the CFS were quite heterogeneous and ranged from densely packed spheres with almost indiscernible individual cells to more loosely packed spheres or individual floating cells (Fig. 1b).

CFS display identical clonal genomic alterations to their parental CRC cells

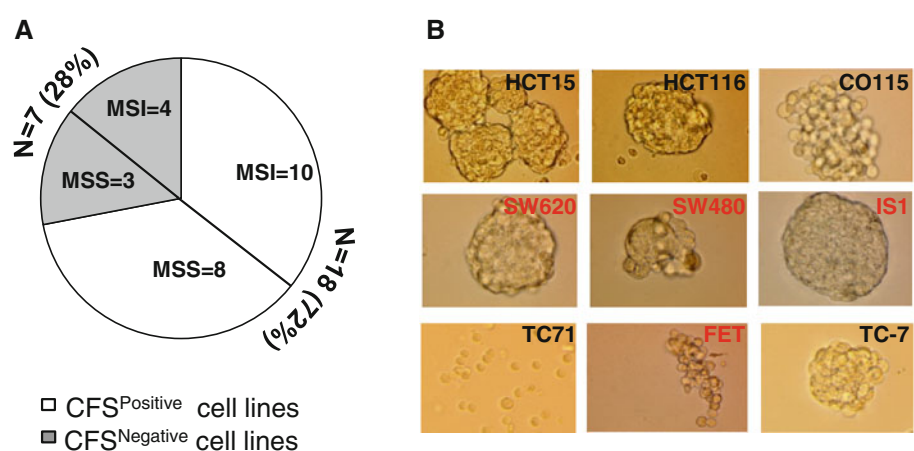
Extensive analysis using SNP microarrays revealed that genomic alterations due to chromosomal instability (CIN)

were similar between parental CRC cell lines and their related CFS. This was observed both with MSI and MSS CRC cell lines displaying low and high levels of CIN, respectively. SNP analysis was performed on 13 CFS/adherent cell lines (LIM2405, HCT8, HCT116, HCT15, TC-7 CO115, RKO, LS411, V9P, HT29, SW620, Colo320) (Fig. 2a; and data not shown). The status of DNA microsatellites that constitute accurate markers of the ‘history’ of each tumor cell in CRC cell lines displaying MSI was also analyzed. Mutation analysis of DNA microsatellite sequences in 10 CFS<sup>Positive</sup> MSI cell lines (HCT116, CO115, HCT15, HCT8, ISHI, LIM2405, LS411, RKO, TC-7, TC71) revealed identical patterns of alteration in non-coding (BAT26, BAT25, NR21, NR25, NR27) and coding (*TGFBR2*, *RAD50*, *MSH6*, *MSH3*, *MBD4*, *BAX*, *ATR*, *BLM*) repeats between parental cells and their corresponding CFS (Fig. 2b). Thus, CRC cell lines and their corresponding CFS progeny are clonally identical, indicating the CFS phenotype arises from the selection of pre-existing clones in the parental cell lines that are able to grow under specific conditions of serum deprivation.

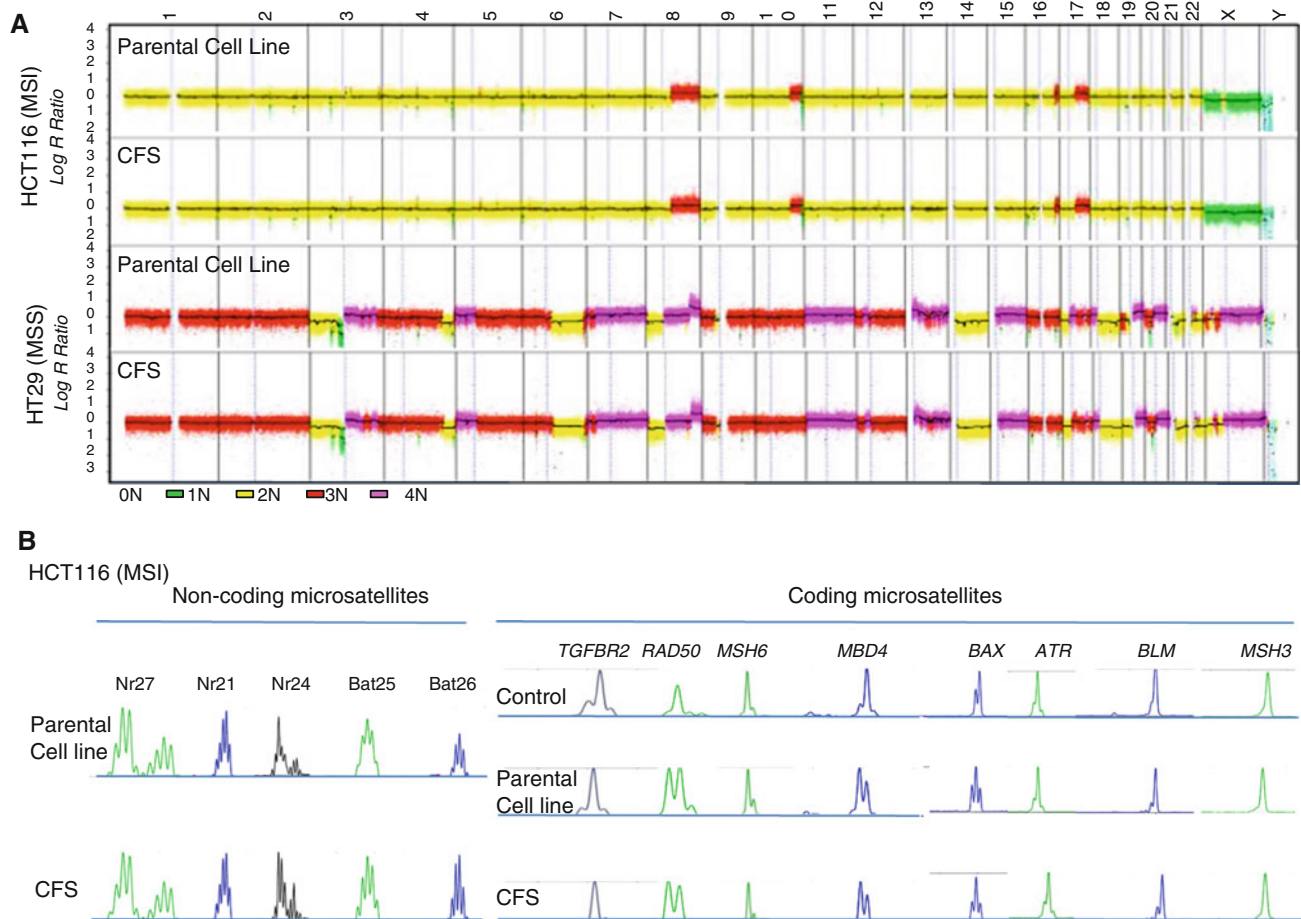
Acquisition of the CFS phenotype by CRC cells is associated with specific changes in gene expression

The gene expression profiles of 13 CFS<sup>Positive</sup> CRC cell lines (8 MSI: LIM2405, HCT8, HCT116, HCT15, TC-7 CO115, RKO, LS411; 5 MSS: V9P, HT29, SW620, Colo320, FET) and their corresponding CFS populations were compared using microarrays. Principal component analysis of the profiles revealed the CFS grouped together with their parental CRC cell lines (Supplementary Fig. S1A). A total of 264 genes displayed significant down- or up-regulation in CFS ( $P$  value  $<0.001$ , paired moderated  $t$  test) and are listed in Supplementary Table S1. Specific signaling and metabolic pathways were associated with the CFS<sup>Positive</sup> gene expression signature (Table 1). As

**Fig. 1 a** Results of the CFS assay in 25 CRC cell lines (MSS = 11; MSI = 14). **b** Morphological features of CFS from 9 CRC cell lines (MSI *black*, MSS *red*) grown in sphere-medium. The morphology of CFS derived from T71 is peculiar and these grow in suspension as single cells







**Fig. 2 a** Extensive microarray analysis revealed that chromosomal aberrations due to CIN were similar between 13 parental CRC cell lines (LIM2405, HCT8, HCT116, HCT15, TC-7, CO115, RKO, LS411, V9P, HT29, SW620, Colo320, FET) and their corresponding CFS. Examples of chromosomal instability in HCT116 (MSI cell line), HT29 (MSS cell line) and their corresponding CFS. **b** Microsatellite instability in HCT116 parental cells and their corresponding

CFS. Mutation analysis of DNA microsatellites showed identical patterns of alteration in both non-coding (BAT26, BAT25, NR21, NR25, NR27) and coding (*TGFBR2*, *RAD50*, *MSH6*, *MSH3*, *MBD4*, *BAX*, *ATR*, *BLM*) repeats in the parental CRC cell lines and their corresponding CFS. This was observed in HCT116 and in 10 CFS<sup>Positive</sup> MSI CRC cell lines (HCT116, CO115, HCT15, HCT8, ISHI, LIM2405, LS411, RKO, TC-7, TC71, not shown)

expected, a number of genes from this signature reflected the different culture conditions used for adherent and CFS cells and were linked mainly to cell metabolism and growth factor pathways (Table 1). The CFS signature also included genes related to stemness and to cellular mechanisms associated with treatment resistance, such as transmembrane transporters, apoptosis and DNA damage (Table 1). Specific chromosomal regions were enriched with genes from the CFS signature (Supplementary Fig. S1B). These regions displayed a similar genomic status (DNA copy-number) in CFS and the parental CRC cell line, suggesting that gene deregulation occurred via epigenetic processes that were independent of DNA copy number. Regions that were frequently down or up-regulated in CFS (>30 %) are shown in the lower panel of Figure S1B.

Both CFS and their parental CRC cell lines are highly tumorigenic and they show similar expression of CSC markers

Serial engraftments of 8 CRC cell lines (5 CFS<sup>Positive</sup>: HCT15, HCT116, LoVo, V9P, RKO; 3 CFS<sup>Negative</sup>: LS174T, LIM1215, KM12) were performed in nude mice. All 8 cell lines tested were found to be highly tumorigenic, even when only 200 CRC cells were injected (Fig. 3a). CFS cells derived from HCT116 and LoVo cell lines were also highly tumorigenic. Low numbers of injected CFS cells (200 or 500 cells) derived from both these MSI CRC cell lines were less tumorigenic compared to the parental cell lines (Fig. 3b; and data not shown). However, no difference was apparent when 1000 cells were injected. The expression of

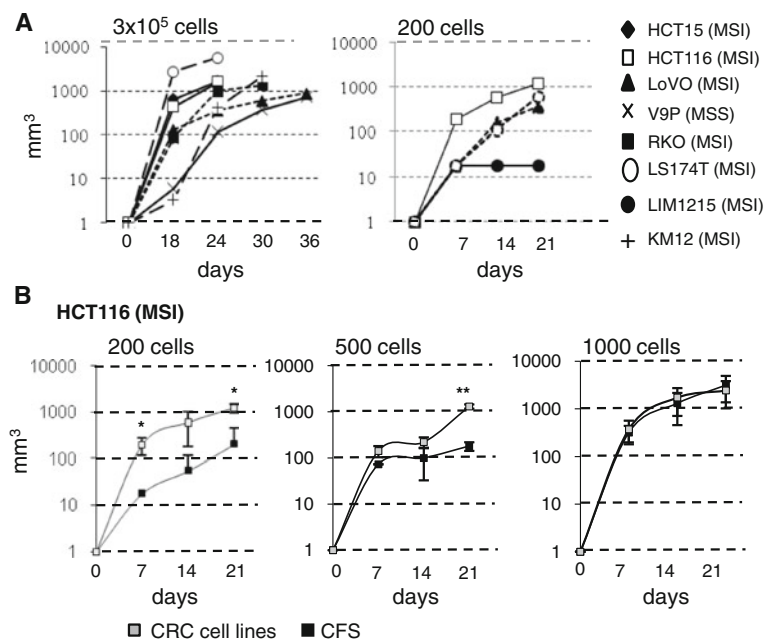
**Table 1** Specific signaling and metabolic pathways significantly associated with the CFS gene signature

	Category	Subcategory	Src	Gene Sets	P value up	P value down
Significant KEGG pathways	Metabolism	Overview	KEGG	Biosynthesis of steroids	2.6E-27	-
		Metabolism of Terpenoids and Polyketides	KEGG	Terpenoid biosynthesis	1.3E-07	-
		Energy Metabolism	KEGG	Reductive carboxylate cycle (CO2 fixation)	2.4E-04	-
		Lipid Metabolism	KEGG	Fatty acid biosynthesis	1.4E-03	-
			KEGG	Synthesis and degradation of ketone bodies	3.3E-03	-
			KEGG	Glycerolipid metabolism	3.4E-03	-
			KEGG	Biosynthesis of unsaturated fatty acids	4.3E-03	-
			KEGG	Fatty acid metabolism	9.6E-03	-
			KEGG	Glycerophospholipid metabolism	2.7E-02	-
		Carbohydrate Metabolism	KEGG	Propanoate metabolism	5.6E-03	-
			KEGG	Pyruvate metabolism	8.6E-03	-
			KEGG	Citrate cycle (TCA cycle)	3.4E-02	-
			KEGG	Valine, leucine and isoleucine degradation	1.0E-02	-
		Amino Acid Metabolism	KEGG	Glycosaminoglycan degradation	1.2E-02	-
			KEGG	Other glycan degradation	1.0E-02	-
	Metabolism of Other Amino Acids	KEGG	Glutathione metabolism	3.9E-02	-	
		KEGG	Adipocytokine signaling pathway	4.0E-03	-	
	Organismal Systems	Endocrine System	KEGG	PPAR signaling pathway	2.9E-02	-
	Environmental Information Processing	Signal Transduction	KEGG	MAPK signaling pathway	1.6E-02	-
	Genetic Information Processing	Folding, Sorting and Degradation	KEGG	Proteasome	-	9.3E-04
			KEGG	Ubiquitin mediated proteolysis	-	1.3E-04
		Replication and Repair	KEGG	DNA replication	-	1.0E-04
			KEGG	Mismatch repair	-	1.4E-04
	Metabolism	Carbohydrate Metabolism	KEGG	Glycolysis / Gluconeogenesis	-	1.3E-04
		Nucleotide Metabolism	KEGG	Pyrimidine metabolism	-	3.6E-02
	Cellular Processes	Cell Growth and Death	KEGG	Cell cycle	-	1.1E-02
	Environmental Information Processing	Signal Transduction	KEGG	Hedgehog signaling pathway	-	2.7E-02
			KEGG	TGF-beta signaling pathway	-	4.0E-02
	Stem Cell Targeted Gene Sets	Stem Cell	MSigDB	BHATTACHARYA_EMBRYONIC_STEM_CELL (C2)	3.5E-02	-
			MSigDB	JAATINEN_HEMATOPOIETIC_STEM_CELL_UP (C2)	4.4E-02	-
MSigDB			WONG_EMBRYONIC_STEM_CELL_CORE (C2)	-	9.3E-04	
MSigDB			OSWALD_HEMATOPOIETIC_STEM_CELL_IN_COLLAGEN_GEL_DN (C2)	-	1.2E-04	
MSigDB			BYSTRYKH_HEMATOPOIESIS_STEM_CELL_QTL_TRANS (C2)	-	1.0E-03	
MSigDB			LIANG_HEMATOPOIESIS_STEM_CELL_NUMBER_QTL (C2)	-	1.2E-02	
MSigDB			GAL_LEUKEMIC_STEM_CELL_DN (C2)	-	4.2E-02	
MSigDB			LIANG_HEMATOPOIESIS_STEM_CELL_NUMBER_LARGE_VS_TINY_DN (C2)	-	4.4E-02	
Treatment Resistance		Drug Resistance	MSigDB	CREIGHTON_ENDOCRINE_THERAPY_RESISTANCE_3 (C2)	7.3E-09	-
			MSigDB	MASSARWEH_TAMOXIFEN_RESISTANCE_UP (C2)	2.9E-07	-
			MSigDB	KANG_CISPLATIN_RESISTANCE_UP (C2)	1.5E-03	-
			MSigDB	KANG_CISPLATIN_RESISTANCE_UP (C2)	1.5E-03	-
			GO	drug transporter activity	9.2E-03	-
			MSigDB	WHITESIDE_CISPLATIN_RESISTANCE_UP (C2)	1.3E-02	-
			MSigDB	CREIGHTON_ENDOCRINE_THERAPY_RESISTANCE_5 (C2)	1.3E-02	-
			MSigDB	WHITESIDE_CISPLATIN_RESISTANCE_UP (C2)	1.3E-02	-
		Cell cycle	GO	cell cycle	-	1.7E-11
			GO	mitotic cell cycle	-	9.3E-04
			GO	cell cycle process	-	9.3E-04
			GO	positive regulation of ubiquitin-protein ligase activity during mitotic cell cycle	-	9.3E-07
Transporter		Apoptosis	GO	negative regulation of ubiquitin-protein ligase activity during mitotic cell cycle	-	9.7E-04
			SMD	cancerModules Cell cycle (expression clusters)	-	1.4E-03
			Bioc	Cyclins and Cell Cycle Regulation	-	5.0E-03
			KEGG	Cell cycle	-	1.1E-02
			Bioc	Cell Cycle: G1/S Check Point	-	1.5E-02
			Bioc	Regulation of p27 Phosphorylation during Cell Cycle Progression	-	1.6E-02
			GO	M phase of mitotic cell cycle	-	2.1E-02
			GO	S phase of mitotic cell cycle	-	2.1E-02
		GO	interphase of mitotic cell cycle	-	2.3E-02	
		Bioc	Ubiquitylation in the Control of Cell Cycle	-	4.8E-02	
		GO	regulation of cell cycle	-	5.0E-02	
		DNA Damage	MSigDB	GO_TRANSITION_METAL_ION_TRANSMEMBRANE_TRANSPORTER_ACTIVITY	7.0E-03	-
			GO	drug transporter activity	9.2E-03	-
			MSigDB	GO_LIPID_TRANSPORTER_ACTIVITY (C5)	2.8E-02	-
			MSigDB	GO_DI_VALENT_INORGANIC_CATION_TRANSMEMBRANE_TRANSPORTER_ACTIVITY (C5)	3.3E-02	-
			MSigDB	REACTOME_METAL_ION_SLC_TRANSPORTERS (C2)	3.5E-02	-
MSigDB			REACTOME_GLUCOSE_AND_OTHER_SUGAR_SLC_TRANSPORTERS (C2)	4.4E-02	-	
GO			phosphatidylinositol transporter activity	4.8E-02	-	
GO			secondary active monocarboxylate transmembrane transporter activity	4.8E-02	-	
Apoptosis		MSigDB	CONCANNON_APOPTOSIS_BY_EPOXOMICIN_UP (C2)	1.2E-03	-	
	MSigDB	BROCKE_APOPTOSIS_REVERSED_BY_IL6 (C2)	2.4E-03	-		
	MSigDB	GO_ANTL_APOPTOSIS (C5)	2.6E-02	-		
	MSigDB	GO_REGULATION_OF_APOPTOSIS (C5)	4.3E-02	-		
	MSigDB	REACTOME_APOPTOSIS (C2)	-	9.3E-04		
	MSigDB	REACTOME_APOPTOSIS_INDUCED_DNA_FRAGMENTATION (C2)	-	1.2E-02		
	MSigDB	CONCANNON_APOPTOSIS_BY_EPOXOMICIN_DN (C2)	-	1.3E-02		
	MSigDB	DEBIASI_APOPTOSIS_BY_REOVIRUS_INFECTION_DN (C2)	-	1.7E-02		
DNA Damage	MSigDB	REACTOME_P53_INDEPENDENT_DNA_DAMAGE_RESPONSE (C2)	-	9.3E-04		
	GO	response to DNA damage stimulus	-	1.1E-03		
	SMD	cancerModules DNA damage response	-	1.2E-02		
	MSigDB	KYNG_DNA_DAMAGE_BY_4NQO (C2)	1.3E-02	3.0E-02		

P value up (down): hypergeometric test P value using top up (down) regulated genes in CFS versus cell line

Bioc biocarta, GO geneontology, MSigDB molecular signatures database, SMD stanford microarray database, KEGG kyoto encyclopedia of genes and genomes, C2 MSigDB curated gene sets, C5 MSigDB GO gene sets

**Fig. 3** **a** Analysis of tumor growth following injection of a high number ( $n = 3 \times 10^5$ ) of cells from CRC cell lines (5 CFS<sup>positive</sup> and 3 CFS<sup>negative</sup>, *left top panel*) or low number ( $n = 200$ ) of cells (2 CFS<sup>positive</sup> and 2 CFS<sup>negative</sup>, *right top panel*). **b** Comparison of tumor growth following injection of CRC cell line HCT116 and the corresponding CFS at 200 cells ( $*P = 0.034$  at 7 days and  $P = 0.012$  at 21 days), 500 cells ( $**P = 0.006$  at 21 days) and 1,000 cells



putative colorectal CSC markers (CD44, CD133, CD166, CD24, CD29, EPCAM, ALDH, OLFM4, LGR5) evaluated using arrays was not significantly different between 13 pairs of CFS cells and their corresponding parental cell lines (8 MSI: LIM2405, HCT8, HCT116, HCT15, TC-7 CO115, RKO, LS411; 5 MSS: V9P, HT29, SW620, Colo320, FET) (Fig. 4a). The expression of 6 putative colorectal CSC markers (CD44, CD166, CD133, CD24, EpCAM, CD29) was also compared using flow cytometry in the 25 parental CRC cell lines (Supplementary Table S2). Overall, the expression of these markers was highly variable and none was expressed exclusively in CFS<sup>positive</sup> cell lines. Moreover, their expression was highly variable over time, as shown for CD44 and CD166 expression in cell sub-populations sorted by FACS from LS174T and HCT116 parental cells (Fig. 4b). Finally, we quantified the expression of putative colorectal CSC markers (CD44, CD133, CD166, CD24, CD29, EPCAM) in 15 primary CRCs and 12 CRC tumor xenografts established from these primary CRCs and grown in nude mice. Most markers were expressed at significantly higher levels in CRC cell lines compared to primary CRCs and/or tumor xenografts (Fig. 4c). Overall, these results demonstrate that the CFS phenotype derived from CRC cell lines is only weakly related to the putative CSC phenotype from primary tumors. In line with this, a CFS assay showed similar results using sorted populations of CD166<sup>+</sup>CD44<sup>+</sup>EpCAM<sup>high</sup> or CD166<sup>-</sup>CD44<sup>-</sup>EpCAM<sup>low</sup> HCT116 cells (Fig. 4d).

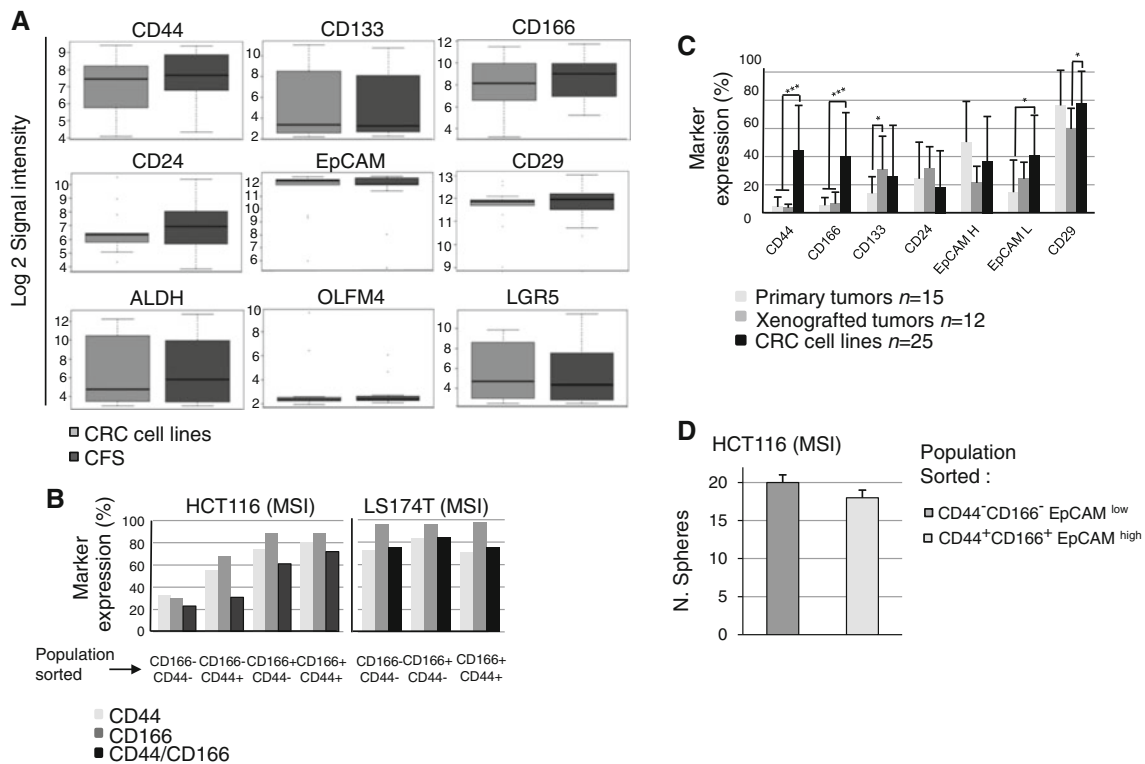
The ability of CRC cells to grow as CFS is strongly increased by prior treatment with 5-Fluorouracil

The ability of LoVo and HCT116 cell lines to grow as CFS in sphere-medium increased following treatment with the chemotherapeutic agent 5-Fluorouracil (5-FU) for 5 days at IC<sub>10</sub> % (5 μM for HCT116 and 7.5 μM for LoVo cells) (Fig. 5a, b; and data not shown). In contrast, 5-FU-resistant clones from the LIM1215 and LS174T CFS<sup>Negative</sup> cell lines remained unable to grow in sphere-medium after 5-FU treatment at IC<sub>10</sub> (Fig. 5a). CRC cell lines were also compared to their CFS counterparts for resistance to 5-FU. In 6 CRC cell lines tested (FET, HCT116, LS411, V9P, TC71, LIM2405), the CFS displayed greater resistance to 5-FU than their parental cells (Fig. 5c). Both CFS<sup>Positive</sup> and CFS<sup>Negative</sup> CRC cell lines displayed marked differences in resistance to this drug (data not shown). LIM2405 showed a strong predilection to grow as CFS in sphere-medium (data not shown), yet the adherent cells and CFS showed similar resistance to 5-FU (Fig. 5c, bottom and right panels).

CFS established from CRC cell lines share a gene signature that predicts disease relapse in CRC patients

Fifty-five genes that were differentially expressed between CFS and their corresponding CRC cell lines displayed a high level of up- or down-regulation (log<sub>2</sub>-fold change for CFS/parental cell line >0.8; Supplementary Table S1; Fig. 6). This 55-gene CFS expression signature predicted





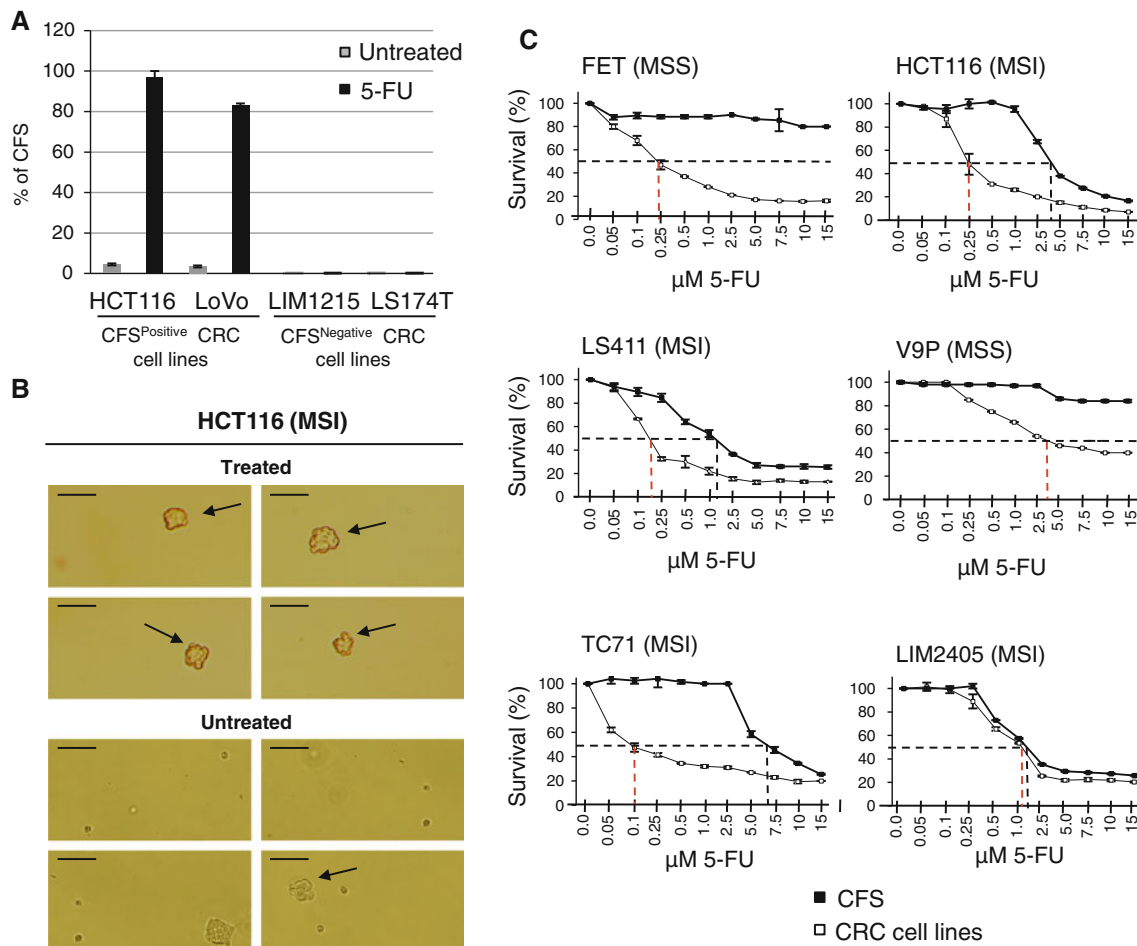
**Fig. 4** **a** Expression of putative colorectal CSC markers (CD44, CD133, CD166, CD24, CD29, EPCAM, ALDH, OLFM4, LGR5) were investigated using gene expression arrays in 13 CFS/parental cell lines (8 MSI: LIM2405, HCT8, HCT116, HCT15, TC-7 CO115, RKO, LS411; 5 MSS: V9P, HT29, SW620, Colo320, FET). **b** Expression of CD166 and CD44 markers was investigated by flow cytometry in different subpopulations of sorted cells from the HCT116 and LS174T parental cells. CD166<sup>-</sup>/CD44<sup>-</sup>, CD166<sup>+</sup>/CD44<sup>-</sup>, CD166<sup>+</sup>/CD44<sup>+</sup> and CD166<sup>-</sup>/CD44<sup>+</sup> (HCT116 only) cells

were sorted by FACS and re-inoculated in medium. After a few days of growth, cells from sorted subpopulations were analyzed again for marker expression. The CD expression pattern was compared to the cell sorted profile. **c** Expression of six markers in CRC cell lines ( $n = 25$ ), primary CRCs ( $n = 15$ ) and/or tumor xenografts established directly from the primary CRC ( $n = 12$ ). **d** CFS ability of HCT116 after sorting of two subpopulations of CD166<sup>+</sup>CD44<sup>+</sup> EpCAM<sup>high</sup> and CD166<sup>-</sup>CD44<sup>-</sup> EpCAM<sup>low</sup> cells

disease relapse in a previously described, retrospective series of stage II and III CRC patients [25]. Tumors were classified into two groups according to the expression level of the 55 genes (T-CFS<sup>High</sup> or T-CFS<sup>Low</sup>; for further details, see “Materials and methods” and Fig. 6a), where T-CFS<sup>High</sup> and T-CFS<sup>Low</sup> correspond to tumors displaying high or low levels of expression of the CFS signature, respectively. Patients with T-CFS<sup>High</sup> tumors showed significantly shorter disease-free survival (DFS) compared to T-CFS<sup>Low</sup> patients (Supplementary Figure S3A, bottom and left panel). To confirm these clinical findings, the same analysis was performed on another, independent cohort of stage II and III CRC patients [26]. A trend for similar clinical impact of the 55-gene CFS expression signature was observed (Supplementary Figure S2A, bottom and right panel). Following normalization of the data, the two patient series were combined to achieve greater statistical power. In the overall series and in univariate analysis, the survival of stage II and III CRC patients was associated with the expression level of the 55-gene CFS signature

(Fig. 7, left panel; Supplementary Figure S2A, top and left panel).

Since the 55-gene signature showed a trend for association with patient outcome in the second dataset, we sought to identify a novel prognostic signature by selecting genes that were individually associated with outcome in the first dataset. This allowed us to define an 8-gene CFS prognostic signature (Fig. 6b; Supplementary Table S1) that showed stronger associations with disease relapse in both the individual and combined patient cohorts (Fig. 7, right panel; Supplementary Figure S2B, top and right panel). To ensure this finding was related to the CFS 55-gene selection and not to the methodology, the same approach was repeated 1,000 times using signatures from 55 genes selected at random. This analysis did not increase the false positive rate in the second dataset (data not shown). In the overall series and in multivariate Cox analysis that included TNM stage, the survival of patients with stage II or III CRC was confirmed to be associated with the expression level of 8 genes from the CFS signature (Table 2).



**Fig. 5** **a** Percentage of CFS in sphere-medium from LoVo, HCT116, LIM1215 and LS174T cells before and after treatment with 5-Fluorouracil for 5 days at IC<sub>10</sub> % (7.5, 7.5, 5 and 20 μM, respectively). **b** Representative image of spheres (*arrows*) from HCT116 cells treated with 5-fluorouracil for 5 days at IC<sub>10</sub> % (*top*) and untreated

(*bottom*) after 3 days in sphere-medium (*size bar* 100 μm). **c** Cell survival curves for 6 CRC cell lines and their corresponding CFS following 5 days of exposure to increasing concentrations of 5-Fluorouracil. *Dotted lines* indicate the IC<sub>50</sub> value in each case

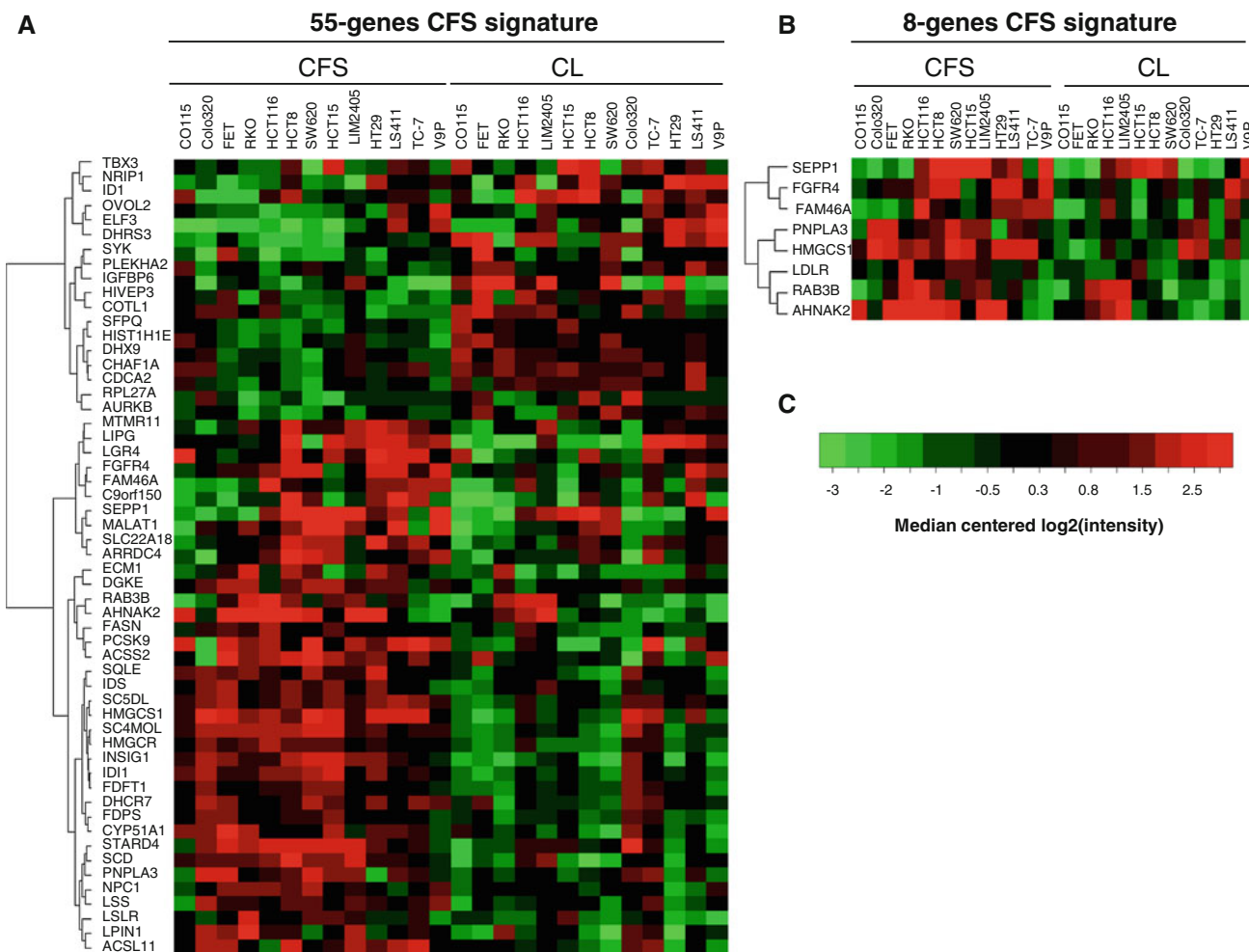
CFS can be used as models to identify new drugs that are more efficient at killing CFS than 5-fluorouracil

Because of the current clinical interest in CFS, the Institut Curie-CNRS chemical library (8,560 compounds) was screened using the HCT116 (MSI) and FET (MSS) CRC cell lines (Supplementary Fig. S3A). Fifteen new compounds were identified with the ability to kill both parental CRC cells and CFS from HCT116 and/or FET at low concentration (IC<sub>50</sub> % < 1 μM; Supplementary Fig. S3B). Of note, 8/15 (53 %) of these new compounds belong to the Nitrofurans' family previously reported to display anticancer and antioxidant properties (Supplementary Table S3, entries 1–8) [29, 30]. Since the library contains 240 of this class of compound (2.8 % of the total), Nitrofurans were significantly over-represented amongst the 15 drugs with high effectiveness against CFS ( $P = 2.4 \times 10^{-9}$ , Fisher's exact test). In contrast to results obtained

with 5-FU (see Fig. 5b), the ability of HCT116 and FET cell lines to grow in serum-free medium did not increase following treatment with three of the new compounds for 5 days at IC<sub>10</sub> % (Supplementary Fig. S3C).

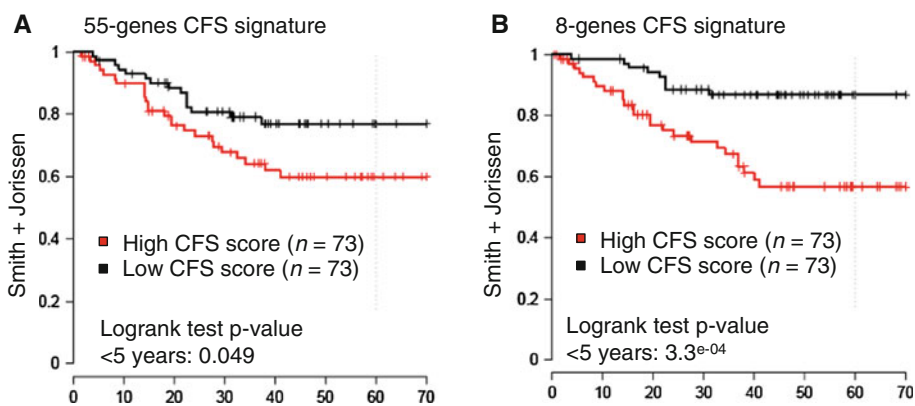
## Discussion

Not all CRC cell lines contain CFS. This result may reflect the true heterogeneity of CRC but may also be due to events that occurred *in vitro* during or after the establishment of CRC cell lines. All CRC cell lines tested here were highly tumorigenic in mouse xenograft assays. We did not observe increased tumorigenicity of CFS-positive compared to CFS-negative CRC cell lines, nor of CFS compared to their parental cell line counterpart. Surprisingly, at lower numbers of injected cells, CFS were less tumorigenic compared to their parental cell line. We have



**Fig. 6** Heatmaps of the 55-gene (a) and 8-gene signatures (b). Visualization of the expression level for each gene in each sample relative to the median gene intensity (c) across all samples. Green corresponds to down-regulation and red to up-regulation. CL parental CRC cell line

**Fig. 7** Kaplan–Meier analysis of disease-free survival for CRC patients classified according to expression levels for the 55-gene (left panel) or 8-gene (right panel) CFS signature



no obvious explanation for this result. In any case, CRC cell lines are not ideal models for evaluating tumorigenicity because of their high level of heterogeneity. Compared to primary colorectal tumors, aberrant expression of putative CSC markers was observed in the CRC cell lines investigated here, but was not different to that of CFS, as already

reported for HCT116 [16]. Moreover, the current transcriptome analyses revealed only weak correlations between CFS and the expression of stem cell gene markers. These results highlight the fact that CSC markers are not specific for CFS. The CFS phenotype is therefore quite different to that of putative CSCs from primary tumors. It

**Table 2** Association of the 8-gene CFS high/low score and TNM stage and prognosis in the combined dataset (Smith + Jorissen)

		Chi <sup>2</sup> test <i>P</i> value	CFS <sup>High</sup> ( <i>n</i> = 73)	CFS <sup>Low</sup> ( <i>n</i> = 73)							
TNM Stage	2 ( <i>n</i> = 82)	0.097	35 (49 %)	46 (63 %)	Cox univariate analysis		Cox multivariate analysis				
	3 ( <i>n</i> = 65)		38 (51 %)	27 (37 %)	<i>n</i> events	H.R.	95 % C.I.	<i>P</i> value	H.R.	95 % C.I.	<i>P</i> value
Jorissen + Smith	TNM stage	3	145	34	2.1	1–4.2	<i>0.037</i>	1.8	0.88–3.5	0.11	<i>4.10E–04</i>
8-genes signature	SC score	CFS <sup>Low</sup>	145	34	0.27	0.13–0.58	<i>8.0E–04</i>	0.29	0.14–0.64	<i>1.8E–03</i>	

H.R. Cox hazard ratio, 95 % C.I. 95 percent confidence interval of H.R.

Significant *p* value in italics (<0.05)

<sup>a</sup> Value: modality of the annotation associated to H.R.

would be interesting to perform similar experiments relating to the expression of CSC markers, tumorigenicity and gene signatures using primary tumors and their sphere counterparts.

Although clearly different to putative CSCs from primary tumors, an interesting question raised by this study is whether the CFS models established from cancer cell lines have clinical relevance. The differential expression of genes in CFS relative to their parental cells is partly a result of the presence of growth factors in one medium and not in the other. This could involve growth factor pathways as well as CSC-like genes. In support of this, it was recently shown that CFS cultures contain only a fraction of CSCs and therefore cannot be regarded as pure CSC models. With this in mind, our data highlight that CFS have retained interesting phenotypical characteristics, including increased resistance to 5-FU in standard culture conditions. Moreover, the 55-gene CFS signature identified here was common to all CRC cell lines and was predictive for disease relapse in CRC patients. Although validated in two independent CRC series, these findings require confirmation in additional studies using larger cohorts of patients. In summary, CFS models derived from CRC cell lines have interesting phenotypical features and may have clinical relevance for drug resistance and disease relapse, but are unlikely to serve as models for putative CSCs in primary CRC.

The genomic and mutational analyses of CRC cell lines and CFS performed here, including the study of microsatellite DNA repeats in MSI cell lines, help to explain the origin of CFS. The status of DNA microsatellites constitutes an accurate marker of the ‘history’ of each tumor cell. These genetic markers, including the non-coding microsatellite repeats (BAT26, BAT25, NR21, NR25, and NR27), were identical between CFS and their parental CRC cells. The CFS phenotype therefore corresponds to a cellular state achieved only by some clones under specific

culture conditions or exposure to drugs. Considering the transcriptome analyses, our results also indicate that only a fraction of cells from CRC cell lines have the capacity to rapidly adjust the expression of specific genes and hence to persist as CFS under challenging growth conditions. These data agree with other recent studies demonstrating widespread plasticity and dedifferentiation processes that affect some tumor cells under specific environmental conditions, particularly CRC cells [4, 31–35].

The expression signature of normal intestinal stem cells, also shared by cells with a stem-like cell phenotype within primary CRC, was recently found to be predictive of disease relapse in CRC patients [36]. However, another group reported the CSC gene signature in CRC may reflect the differentiation status of malignant tissue [37], rather than reflecting the number of CSCs as suggested in the former study. The results presented here relate only to the properties of CFS cultures derived from CRC cell lines, not from primary tumors. We speculate that the gene expression pattern observed in our CFS model corresponds to that of poorly differentiated ‘progenitor cells’, or dedifferentiated CRC cells, due to the altered cell culture conditions. This pattern could be expressed in an important fraction of cancer cells within the primary tumor. The risk of developing a recurrence of CRC might be associated with the expression of specific genes in these tumor cells and which are required for tumor regeneration following cancer therapy. Interestingly, 5 of the genes included in the limited 8-gene CFS signature are predicted to have indirect associations with signaling pathways. *LDLR* and *HMGCS1* participate in SREBP control of the lipid pathway by stimulating lipid synthesis, while *FGFR4* is a growth factor reported to be associated with poor prognosis and aggressive disease in many different cancer types including colon cancer [27]. Although still poorly described, *AHNAK2* and *FAM46A* are both upregulated in cisplatin-resistant gastric cancer cell lines [28].



The clinical and molecular heterogeneity of CRC is a major limitation of this type of study. Principal component analyses of the transcriptomic profiles showed that considerable heterogeneity remains between CFS displaying MSI or MSS. The response to chemotherapy is thought to be different between patients with MSI and MSS colon tumors, even when considering tumors with the same stage of disease [37]. The CFS gene signature identified here was shared by both MSI and MSS CRC. Nevertheless, we could not evaluate the clinical relevance of MSI-specific or MSS-specific CFS signatures for patients with these tumor subgroups because MSI status is not contained in the publicly available Affymetrix U133P2 datasets with Recurrence Free Survival annotations used for this study. This is a subject for future investigation.

In conclusion, the current findings support the existence of CFS subpopulations within CRC cell lines and provide a framework to explain the origin of these tumor cells. They also suggest that CRC cells acquire the CFS phenotype through mechanisms that are only weakly related to CSC, but which could nevertheless be important for the development of chemoresistance by CRC cells in vivo in primary tumors. The screening of 8,560 potential anticancer agents using an assay involving CFS populations derived from CRC cell lines may therefore be a useful approach to identify novel drugs for clinical application. The 15 new drugs identified in this study, including several new compounds belonging to the Nitrofurans' family, are currently being tested for toxicity and efficacy in pre-clinical studies using animal models.

**Acknowledgments** This work was supported by the 'Carte d'Identité des Tumeurs' (CIT) program (<http://cit.ligue-cancer.net>) from the Ligue Nationale Contre le Cancer and by grants from the 'Institut National du Cancer' (INCa) (to A.D.). The A.D. group has the label "La Ligue Contre le Cancer". A.C. is a recipient of an INCa fellowship (Institut National du Cancer). A.L. is a recipient of a MESR fellowship (Ministère de l'Enseignement Supérieur et de la Recherche).

**Conflict of interest** No potential conflicts of interests were disclosed.

## References

1. Todaro M, Alea MP, Di Stefano AB, Cammareri P, Vermeulen L, Iovino F, Tripodo C, Russo A, Gulotta G, Medema JP, Stassi G (2007) Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4. *Cell Stem Cell* 1(4):389–402. doi:10.1016/j.stem.2007.08.001
2. O'Brien CA, Pollett A, Gallinger S, Dick JE (2007) A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* 445(7123):106–110. doi:10.1038/nature05372
3. Ricci-Vitiani L, Lombardi DG, Pilozzi E, Biffoni M, Todaro M, Peschle C, De Maria R (2007) Identification and expansion of human colon-cancer-initiating cells. *Nature* 445(7123):111–115. doi:10.1038/nature05384
4. Reya T, Morrison SJ, Clarke MF, Weissman IL (2001) Stem cells, cancer, and cancer stem cells. *Nature* 414(6859):105–111. doi:10.1038/3510216735102167
5. Dean M, Fojo T, Bates S (2005) Tumour stem cells and drug resistance. *Nat Rev Cancer* 5(4):275–284. doi:10.1038/nrc1590
6. Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poepl A, Ling V, Beyene J, Canty AJ, Danska JS, Bohlander SK, Buske C, Minden MD, Golub TR, Jurisica I, Ebert BL, Dick JE (2011) Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat Med* 17(9):1086–1093. doi:10.1038/nm.2415
7. Al-Hajj M, Becker MW, Wicha M, Weissman I, Clarke MF (2004) Therapeutic implications of cancer stem cells. *Curr Opin Genet Dev* 14(1):43–47. doi:10.1016/j.gde.2003.11.007
8. Medema JP, Vermeulen L (2011) Microenvironmental regulation of stem cells in intestinal homeostasis and cancer. *Nature* 474(7351):318–326. doi:10.1038/nature10212
9. Liu Y, Bodmer WF (2006) Analysis of P53 mutations and their expression in 56 colorectal cancer cell lines. *Proc Natl Acad Sci USA* 103(4):976–981. doi:10.1073/pnas.0510146103
10. Volchenboum SL, Li C, Li S, Attiyeh EF, Reynolds CP, Maris JM, Look AT, George RE (2009) Comparison of primary neuroblastoma tumors and derivative early-passage cell lines using genome-wide single nucleotide polymorphism array analysis. *Cancer Res* 69(10):4143–4149. doi:10.1158/0008-5472.CAN-08-3112
11. Douglas EJ, Fiegler H, Rowan A, Halford S, Bicknell DC, Bodmer W, Tomlinson IP, Carter NP (2004) Array comparative hybridization analysis of colorectal cancer cell lines and primary carcinomas. *Cancer Res* 64(14):4817–4825. doi:10.1158/0008-5472
12. Willson JK, Bittner GN, Oberley TD, Meisner LF, Weese JL (1987) Cell culture of human colon adenomas and carcinomas. *Cancer Res* 47(10):2704–2713
13. Kondo T (2007) Stem cell-like cancer cells in cancer cell lines. *Cancer Biomark* 3(4–5):245–250
14. Yeung TM, Gandhi SC, Wilding JL, Muschel R, Bodmer WF (2010) Cancer stem cells from colorectal cancer-derived cell lines. *Proc Natl Acad Sci USA* 107(8):3722–3727. doi:10.1073/pnas.0915135107
15. Dittfeld C, Dietrich A, Peickert S, Hering S, Baumann M, Grade M, Ried T, Kunz-Schughart LA (2009) CD133 expression is not selective for tumor-initiating or radioresistant cell populations in the CRC cell lines HCT-116. *Radiother Oncol* 92(3):353–361. doi:10.1016/j.radonc.2009.06.034
16. Kai K, Nagano O, Sugihara E, Arima Y, Sampetean O, Ishimoto T, Nakanishi M, Ueno NT, Iwase H, Saya H (2009) Maintenance of HCT116 colon cancer cell line conforms to a stochastic model but not a cancer stem cell model. *Cancer Sci* 100(12):2275–2282. doi:10.1111/j.1349-7006.2009.01318.x
17. Reynolds BA, Weiss S (1992) Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* 255(5052):1707–1710
18. Lobo NA, Shimono Y, Qian D, Clarke MF (2007) The biology of cancer stem cells. *Annu Rev Cell Dev Biol* 23:675–699. doi:10.1146/annurev.cellbio.22.010305.104154
19. Fan X, Ouyang N, Teng H, Yao H (2011) Isolation and characterization of spheroid cells from the HT29 colon cancer cell line. *Int J Colorectal Dis* 26(10):1279–1285. doi:10.1007/s00384-011-1248-y
20. Vermeulen L, Todaro M, de Sousa Mello F, Sprick MR, Kemper K, Perez Alea M, Richel DJ, Stassi G, Medema JP (2008) Single-

- cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. *Proc Natl Acad Sci USA* 105(36):13427–13432. doi:[10.1073/pnas.0805706105](https://doi.org/10.1073/pnas.0805706105)
21. Jung P, Sato T, Merlos-Suarez A, Barriga FM, Iglesias M, Rossell D, Auer H, Gallardo M, Blasco MA, Sancho E, Clevers H, Batlle E (2011) Isolation and in vitro expansion of human colonic stem cells. *Nat Med* 17(10):1225–1227. doi:[10.1038/nm.2470](https://doi.org/10.1038/nm.2470)
  22. Lengauer C, Kinzler KW, Vogelstein B (1997) Genetic instability in colorectal cancers. *Nature* 386(6625):623–627. doi:[10.1038/386623a0](https://doi.org/10.1038/386623a0)
  23. Duval A, Hamelin R (2002) Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res* 62(9):2447–2454
  24. Pino MS, Chung DC (2010) The chromosomal instability pathway in colon cancer. *Gastroenterology* 138(6):2059–2072. doi:[10.1053/j.gastro.2009.12.065](https://doi.org/10.1053/j.gastro.2009.12.065)
  25. Smith JJ, Deane NG, Wu F, Merchant NB, Zhang B, Jiang A, Lu P, Johnson JC, Schmidt C, Bailey CE, Eschrich S, Kis C, Levy S, Washington MK, Heslin MJ, Coffey RJ, Yeatman TJ, Shyr Y, Beauchamp RD (2010) Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* 138(3):958–968. doi:[10.1053/j.gastro.2009.11.005](https://doi.org/10.1053/j.gastro.2009.11.005)
  26. Jorissen RN, Gibbs P, Christie M, Prakash S, Lipton L, Desai J, Kerr D, Aaltonen LA, Arango D, Kruhoffer M, Orntoft TF, Andersen CL, Gruidl M, Kamath VP, Eschrich S, Yeatman TJ, Sieber OM (2009) Metastasis-associated gene expression changes predict poor outcomes in patients with Dukes stage B and C colorectal cancer. *Clin Cancer Res* 15(24):7642–7651. doi:[10.1158/1078-0432.CCR-09-1431](https://doi.org/10.1158/1078-0432.CCR-09-1431)
  27. Spinola M, Leoni V, Pignatiello C, Conti B, Ravagnani F, Pastorino U, Dragani TA (2005) Functional FGFR4 Gly388Arg polymorphism predicts prognosis in lung adenocarcinoma patients. *J Clin Oncol* 23(29):7307–7311. doi:[10.1200/JCO.2005.17.350](https://doi.org/10.1200/JCO.2005.17.350)
  28. Kang HC, Kim IJ, Park JH, Shin Y, Ku JL, Jung MS, Yoo BC, Kim HK, Park JG (2004) Identification of genes with differential expression in acquired drug-resistant gastric cancer cells using high-density oligonucleotide microarrays. *Clin Cancer Res* 10(1 Pt 1):272–284
  29. Hayakawa I, Shioya R, Agatsuma T, Furukawa H, Sugano Y (2004) Thienopyridine and benzofuran derivatives as potent anti-tumor agents possessing different structure-activity relationships. *Bioorg Med Chem Lett* 14(13):3411–3414. doi:[10.1016/j.bmcl.2004.04.079](https://doi.org/10.1016/j.bmcl.2004.04.079)
  30. Baraldi PG, Romagnoli R, Giovanna Pavani M, del Carmen Nunez M, Bingham JP, Hartley JA (2002) Benzoyl and cinnamoyl nitrogen mustard derivatives of benzoheterocyclic analogues of the tallimustine: synthesis and antitumour activity. *Bioorg Med Chem* 10(5):1611–1618
  31. Gottesman MM (2002) Mechanisms of cancer drug resistance. *Annu Rev Med* 53:615–627. doi:[10.1146/annurev.med.53](https://doi.org/10.1146/annurev.med.53)
  32. Zhou S, Schuetz JD, Bunting KD, Colapietro AM, Sampath J, Morris JJ, Lagutina I, Grosveld GC, Osawa M, Nakauchi H, Sorrentino BP (2001) The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nat Med* 7(9):1028–1034. doi:[10.1038/nm0901-1028](https://doi.org/10.1038/nm0901-1028)
  33. Ito K, Hirao A, Arai F, Matsuoka S, Takubo K, Hamaguchi I, Nomiya K, Hosokawa K, Sakurada K, Nakagata N, Ikeda Y, Mak TW, Suda T (2004) Regulation of oxidative stress by ATM is required for self-renewal of haematopoietic stem cells. *Nature* 431(7011):997–1002. doi:[10.1038/nature02989](https://doi.org/10.1038/nature02989)
  34. Bao S, Wu Q, McLendon RE, Hao Y, Shi Q, Hjelmeland AB, Dewhirst MW, Bigner DD, Rich JN (2006) Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* 444(7120):756–760. doi:[10.1038/nature05236](https://doi.org/10.1038/nature05236)
  35. Diehn M, Cho RW, Lobo NA, Kalisky T, Dorie MJ, Kulp AN, Qian D, Lam JS, Ailles LE, Wong M, Joshua B, Kaplan MJ, Wapnir I, Dirbas FM, Somlo G, Garberoglio C, Paz B, Shen J, Lau SK, Quake SR, Brown JM, Weissman IL, Clarke MF (2009) Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* 458(7239):780–783. doi:[10.1038/nature07733](https://doi.org/10.1038/nature07733)
  36. Merlos-Suarez A, Barriga FM, Jung P, Iglesias M, Cespedes MV, Rossell D, Sevillano M, Hernando-Momblona X, da Silva-Diz V, Munoz P, Clevers H, Sancho E, Mangues R, Batlle E (2011) The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* 8(5):511–524. doi:[10.1016/j.stem.2011.02.020](https://doi.org/10.1016/j.stem.2011.02.020)
  37. Sinicrope FA, Foster NR, Thibodeau SN, Marsoni S, Monges G, Labianca R, Kim GP, Yothers G, Allegra C, Moore MJ, Gallinger S, Sargent DJ (2011) DNA mismatch repair status and colon cancer recurrence and survival in clinical trials of 5-fluorouracil-based adjuvant therapy. *J Natl Cancer Inst* 103(11):863–875. doi:[10.1093/jnci/djr153](https://doi.org/10.1093/jnci/djr153)
  38. Poupon MF, Arvelo F, Goguel AF, Bourgeois Y, Jacrot M, Hanania N, Arriagada R, Le Chevalier T (1993) Response of small-cell lung cancer xenografts to chemotherapy: multidrug resistance and direct clinical correlates. *J Natl Cancer Inst* 85(24):2023–2029
  39. Buhard O, Suraweera N, Lectard A, Duval A, Hamelin R (2004) Quasimonomorphic mononucleotide repeats for high-level microsatellite instability analysis. *Dis Markers* 20(4–5):251–257
  40. Gayet J, Zhou XP, Duval A, Rolland S, Hoang JM, Cottu P, Hamelin R (2001) Extensive characterization of genetic alterations in a series of human colorectal cancer cell lines. *Oncogene* 20(36):5025–5032. doi:[10.1038/sj.onc.1204611](https://doi.org/10.1038/sj.onc.1204611)
  41. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31(4):e15
  42. Letouze E, Sow A, Petel F, Rosati R, Figueiredo BC, Burnichon N, Gimenez-Roqueplo AP, Lalli E, de Reynies A (2012) Identity by descent mapping of founder mutations in cancer using high-resolution tumor SNP data. *PLoS ONE* 7(5):e35897. doi:[10.1371/journal.pone.0035897](https://doi.org/10.1371/journal.pone.0035897)
  43. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3:Article3. doi:[10.2202/1544-6115.1027](https://doi.org/10.2202/1544-6115.1027)