Sleeping Beauty genetic screen identifies miR-23b::BTBD7 gene interaction as crucial for colorectal cancer metastasis.

Eleonora Grisard[1*], Michela Coan[1,2*], Luigi Zandonà[1], Laura Cesaratto[1], Ilenia Rigo[1], Alice Paulitti[1], Eva Andreuzzi[1], Gian Luca Rampioni Vinciguerra[1,9], Evelina Poletto[1], Fabio Del Ben[3], Giulia Brisotto[3,4], Eva Biscontin[3], Matteo Turetta[5], Erik Dassi[6], Alex Mirnezami[7], Vincenzo Canzonieri[8,9], Andrea Vecchione[10], Gustavo Baldassarre[1], Maurizio Mongiat[1], Riccardo Spizzo[1,11] and Milena S. Nicoloso[1].

Author information

[1] Division of Molecular Oncology, Department of Translational Research, IRCCS CRO Centro di Riferimento Oncologico Aviano, Italy.
[2] Department of Life and Reproduction Sciences, University of Verona, Verona, Italy
[3] Immunopathology and Cancer Biomarkers, Department of Translational Research, IRCCS CRO Centro di Riferimento Oncologico Aviano, Italy.
[4] Veneto Institute of Oncology IOV-IRCCS, Comprehensive Cancer Centre, Department of Surgery, Oncology and Gastroenterology, University of Padova, Italy
[5] Pathology Department, University Hospital of Udine, Italy
[6] Laboratory of Translational Genomics, Centre for Integrative Biology, University of Trento, Italy
[7] Cancer Sciences, University Surgical Unit, University of Southampton, UK.
[8] Pathology Unit, IRCCS CRO Centro di Riferimento Oncologico Aviano, Italy.
[9] Department of Medical, Surgical and Health Sciences, University of Trieste Medical School, Trieste, Italy
[10] Faculty of Medicine and Psychology, Department of Clinical and Molecular Medicine, University of Rome "Sapienza", Santo Andrea Hospital, 00189 Rome, Italy

Supplementary Tables and Figures

Suppl. Table 1: list of SB insertions in TN4_20 as retrieved by linker-mediated PCR.

Position (hg19)	Closest Transcripts and Biotype (Genecode V19)	Description
Chr13 :94,910,954-910,955	GPC6 (protein-coding)	Intronic
Chr1 :212,405,256-212,405,257	RP11 15I11 (lincRNA)	Intronic
Chr19 :3,501,100-3,501,101	DOHH (protein-coding)	Promoter
Chr1 :235,641,751-235,641,752	B3GALNT2 (protein-coding)	Intronic
Chr1 :45,885,179-45,885,180	TESK2 (protein-coding)	Intronic
Chr5 :12,473,909-12,473,910	Upstream: CNND2 (protein-coding); Downstream: CT49 (lincRNA)	Intergenic (with LINE L1PREC2)
Chr14 :93,705,840-93,705,841	BTBD7 (protein-coding)	3'UTR

Suppl. Table 2: list of clinical information of patients belonging to The Tumour Genome Atlas Colon Rectal Project that have been used for microRNA gene expression analysis.

Metastasis	race	microsatellite	gender	age	pT	pN	M
M0	WHITE	MSI-L	FEMALE	71	T3	N0	M0
			MALE	78	T4b	N0	M0
			FEMALE	46	T3	N1	M0
			FEMALE	50	T3	N1	M0
			FEMALE	80	T3	N2	M0
			MALE	50	T3	N2	M0
			FEMALE	69	T3	N2	M0
			FEMALE	72	T3	N2	M0
M1			MALE	50	T3	N2	M1
			FEMALE	49	T3	N1	M1
			FEMALE	54	T3	N1	M1
			MALE	47	T4a	N1b	M1a
			FEMALE	57	T4	N0	M1
			FEMALE	82	T3	N2	M1
			MALE	38	T4	N2	M1
			FEMALE	72	T4a	N1	M1

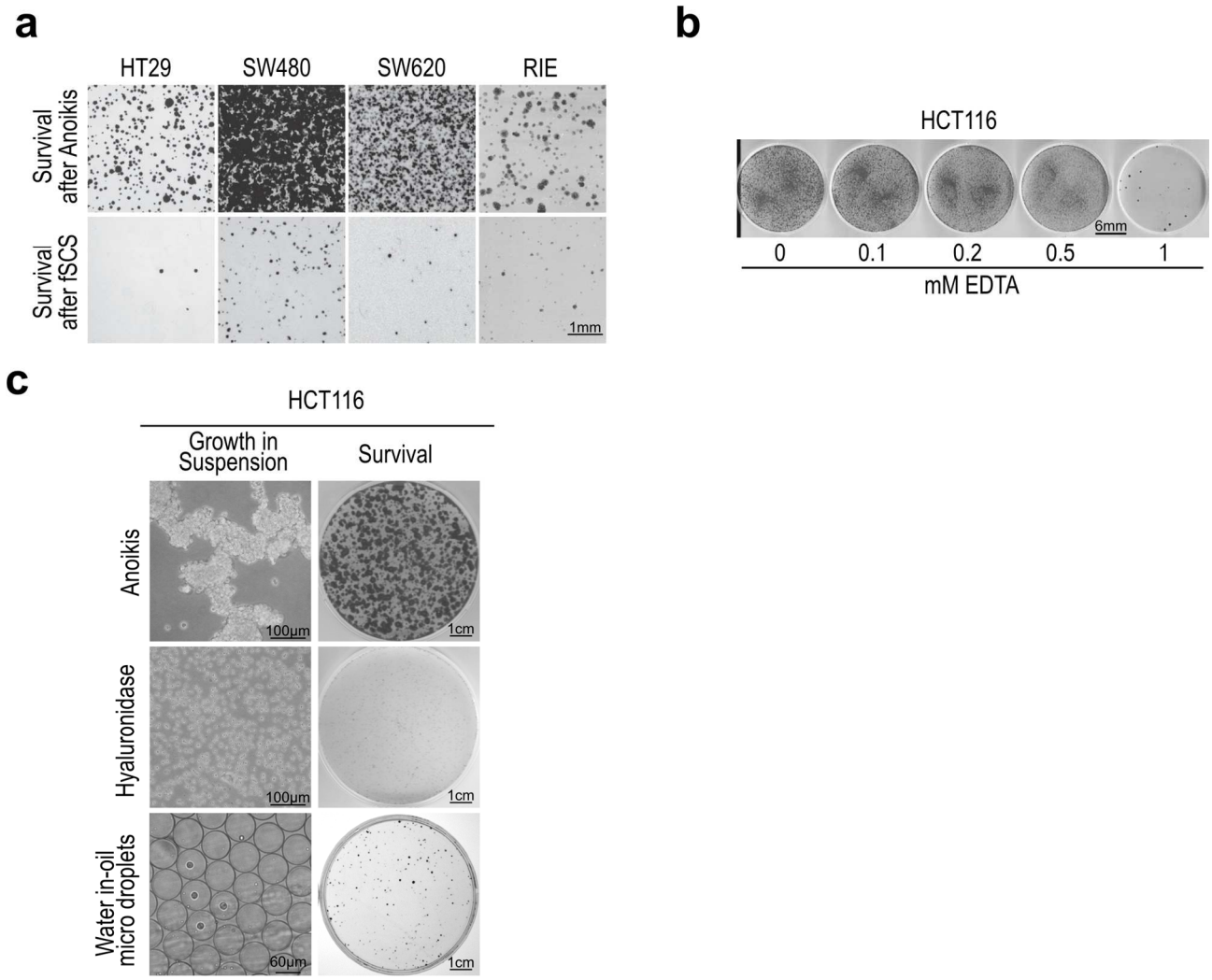
Suppl. Table 3: list of clinical information of patients belonging to CRO Aviano Biobank

		Normal	Tum-M0	Tum-M1	Met
	Number	10	10	11	9
TN	pT2 N0	-	6	0	2
	pT2 N+	-	0	0	1
	pT3 N0	-	3	1	2
	pT3 N+	-	1	8	4
	pT4 N+	-	0	2	0
Site	Colon NAS	-	9	4	7
	Sigma-Rectum	-	1	7	2
	Age (median)	61	66	71	71

Suppl. Table 4: List of primers used

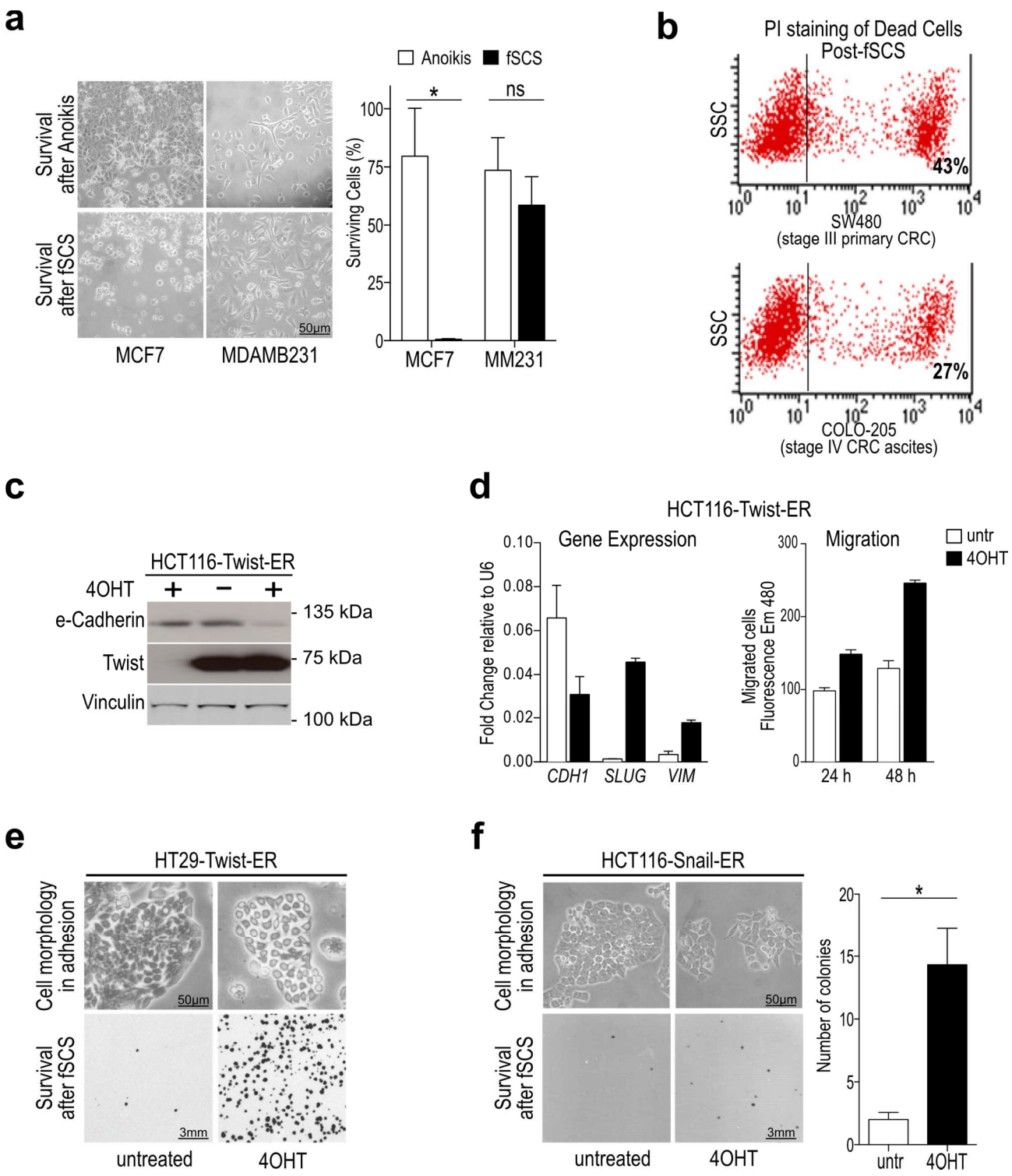
hEcad-5	TGCCCAGAAAATGAAAAAGG
hEcad-3	GTGTATGTGGCAATGCGTTC
hVim-5	GAGAACTTTGCCGTTGAAGC
hVim-3	GCTTCCTGTAGGTGGCAATC
HAS2-5	GCCTCATCTGTGGAGATGGT
HAS2-3	ATGCACTGAACACACCCAAA
hSlug-5	GGGGAGAAGCCTTTTTCTTG
hSlug-3	TCCTCATGTTTGTGCAGGAG
hTwist-5	GGAGTCCGCAGTCTTACGAG
hTwist-3	TCTGGAGGACCTGGTAGAGG
ZEB1-5	CAGGCAGATGAAGCAGGATG
ZEB1-3	GACCACTGGCTTCTGGTGTG
hOct4-5	AGCGATCAAGCAGCGACTAT
hOct4-3	AGAGTGGTGACGGAGACAGG
hNanog-5	CCCCAGCCTTTACTCTTCCTA
hNanog-3	CCAGGTTGAATTGTTCCAGGTC
AS-Bfa-Linker-AS	[PHOS]TAGTCCCTTAAGCGGAG[AMC3]
Nla-Linker-S	GTAATACGACTCACTATAGGGCTCCGCTTAAGGGACCATG
Nla-Linker-AS	[PHOS]GTCCCTTAAGCGGAGCC[AMC3]
IRDR-R-primary	GCTTGTGGAAGGCTACTCGAAATGTTTGACCC
IRDR-L-primary	CTGGAATTTTCCAAGCTGTTTAAAGGCACAGTCAAC
Linker-primary	GTAATACGACTCACTATAGGGC
IRDR-R-nested	CCACTGGGAATGTGATGAAAGAAATAAAAGC
IRDR-L-nested	GACTTGTGTCATGCACAAAGTAGATGTCC
Linker-nested	AGGGCTCCGCTTAAGGGAC
BTBD7A_5(CDS)	AGTCAAATGCCTGGTTACGG
BTBD7A_3(CDS)	TGTCTGGCACATTGGACATT
BTBD7upA_5(3'UTR)	GTTTCCAATTTGCCTTCTGC
BTBD7up A_3(3'UTR)	GGCTTTGAGGCTTTTCAGTG
BTBD7_luc_F	GCTCTAGAGC CCAGTTCCCATGGCCTAAAT
BTBD7_luc_R	GCTCTAGAGC GGTAGGTTCAAGAGCTACATGC
BTBD7_luc_mut-F	CTTTTTTAATGCTGTACTAAGATAGTCTGAGTAAGGTTGCTCTCAGA
BTBD7_luc_mut-R	TCTGAGAGCAACCTTACTCAGACTATCTTAGTACAGCATTAAAAAAG

Supplementary Figure S1.



Supplementary Figure S1. Forcing cells to a single cell suspension (fSCS) reduces cell survival compared to anoikis. (a) Representative images of fixed and stained colonies formed after *in vitro* anoikis or fSCS of human CRC cell lines (HT29, SW480, SW620) or normal Rat Intestinal Epithelial cells (RIE). **(b)** Representative images of fixed and stained colonies formed by HCT116 cells following 24h growth on ultralow attachment plates with serum free medium in the presence of increasing concentrations of EDTA (0.1 mM; 0.2 mM; 0.5 mM; 1 mM). **(c)** Representative images of growth in suspension (left) or of fixed and stained colonies (right) formed by HCT116 cells following anoikis (top panels), following 24h growth on ultralow attachment plates with serum free medium in the presence hyaluronidase (1,3 mg/mL) (middle panels) or following 24h inclusion into water-in-oil micro droplets as single cells.

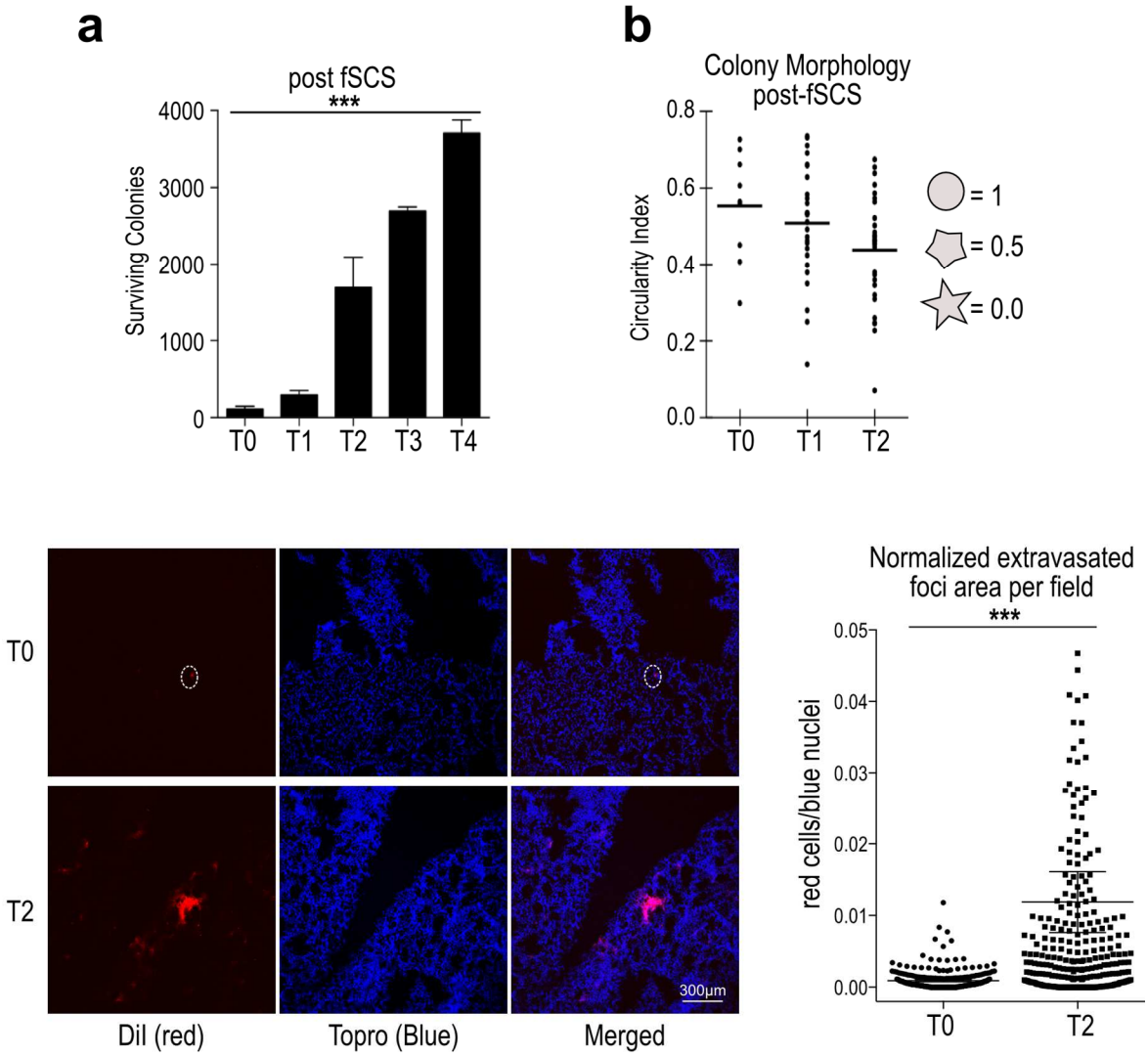
Supplementary Figure S2.



Supplementary Figure S2. Mesenchymal traits foster resistance to fSCS. (a) Representative images at 10x magnification of surviving MCF7 and MDAMB231 cells after 24h in anoikis (upper panels) or in fSCS (lower panel). Bars in right graph present mean +/- sem, expressed as percentage of surviving cells compared to plating control, counted at 10x magnification field from two independent experiments (*p<0.05, using paired t test with unequal variances). **(b)** Cell death by FACS analysis of

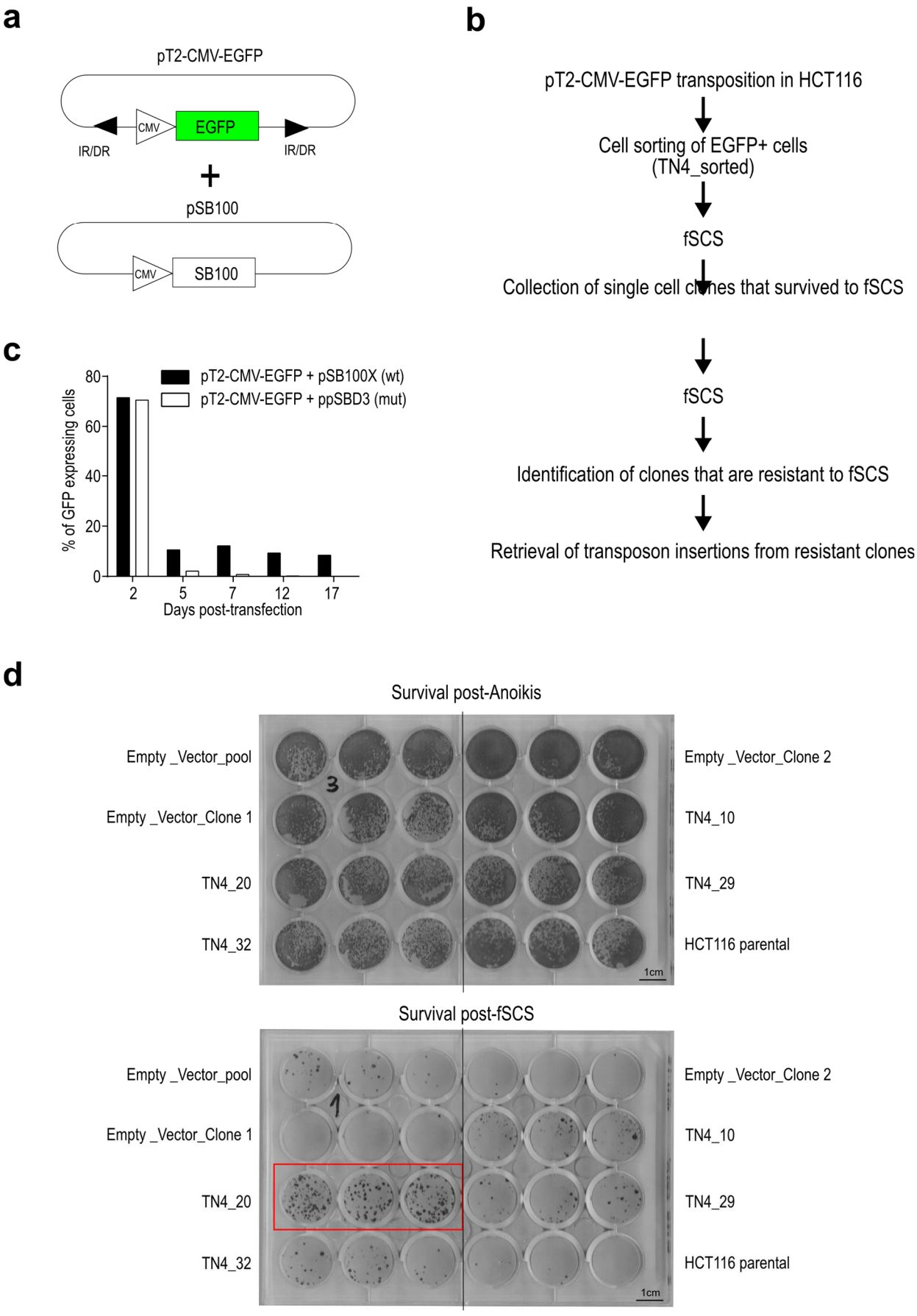
propidium iodide (PI) stained SW480 and COLO205 cells following fSCS assay (SW480 cells are derived from Dukes' stage B primary CRC, whereas COLO205 are derived from Dukes' stage D ascites). Reported percentages indicate the amount of PI positive/dead cells. **(c)** Western Blot analysis of e-Cadherin and Twist in HCT116-Twist-ER cells in presence or absence of 4-hydroxytamoxifen (4OHT). **(d)** Characterization of HCT116-Twist-ER inducible model. Left panel: qRT-analysis in HCT116-Twist- ER cells in presence or absence of 4OHT. Bars indicate mean +/- std of three independent biological replicates. Right panel: Transwell migration assay. Quantification of HCT116-Twist-ER cells motility in transwell migration assay after 24 or 48 hours in presence or absence of 4OHT. Bars indicate mean +/- std of 2 independent biological replicates. **(e)** Representative images at 10x magnification showing the morphology of HT29-Twist-ER cells in presence or absence of 4OHT. Lower panels: representative images of fixed and stained colonies formed by HT29-Twist-ER cells in presence or absence of 4OHT. **(f)** Left panel: representative images at 10x magnification showing the morphology of HCT116-Snail-ER cells in presence or absence of 4OHT. Lower panels: representative images of fixed and stained colonies formed by HCT116-Snail-ER cells in presence or absence of 4OHT. Right panel: count of surviving colonies post-fSCS formed by HCT116-Snail-ER cells in the presence or absence of 4OHT. Bars in graph represent mean +/- std of surviving colonies per field (10x) of one representative experiment (*$p<0.05$, using unpaired t test with unequal variances).

Supplementary Figure S3.



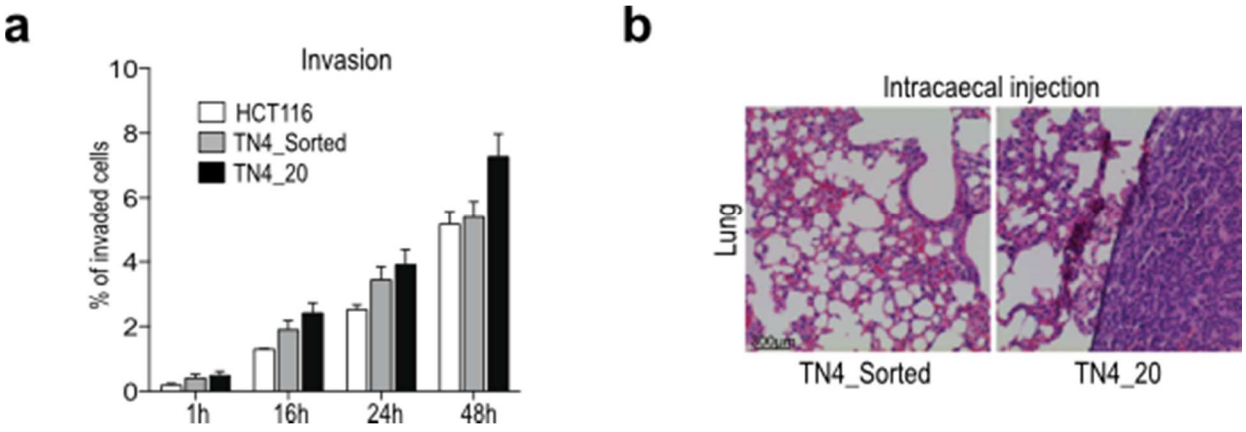
Supplementary Figure S3. Multiple rounds of fSCS enrich for cells showing resistance to fSCS, mesenchymal traits, and increased in vivo extravasation potential. (a) Quantification of the number of colonies generated by HCT116 cells subjected to multiple rounds of fSCS (T1, T2, T3 and T4) or that never underwent to fSCS (T0). Bars in graph represent mean +/- std from two independent biological experiments (***p<0.0005, by Kruskal Wallis non-parametric test). **(b)** Measurement of the morphology of surviving colonies after multiple rounds of fSCS. Circularity index 1= circular; 0.5= intermediate; 0.0= scattered colony. Dot plot graph indicates mean and standard deviation from one biological experiment. **(c)** *In vivo* extravasation assay. Left panel: representative confocal images of lung sections (10x magnification) from nude mice at 72h after intravenous injection of DiI labeled T0 and T2 cells (n=4 and n=5 for T0 and T2, respectively). (Red = DiI stained infiltrated cells; Blue = lung mice nuclei counterstained with Topro; white dashed circle points to T0 red cells). Right panels: Dot plot indicating mean with 95% CI of extravasated DiI labeled T0 or T2, measured as normalized red/blue area per field (***p<0.0005, using unpaired t test with unequal variances).

Supplementary Figure S4.



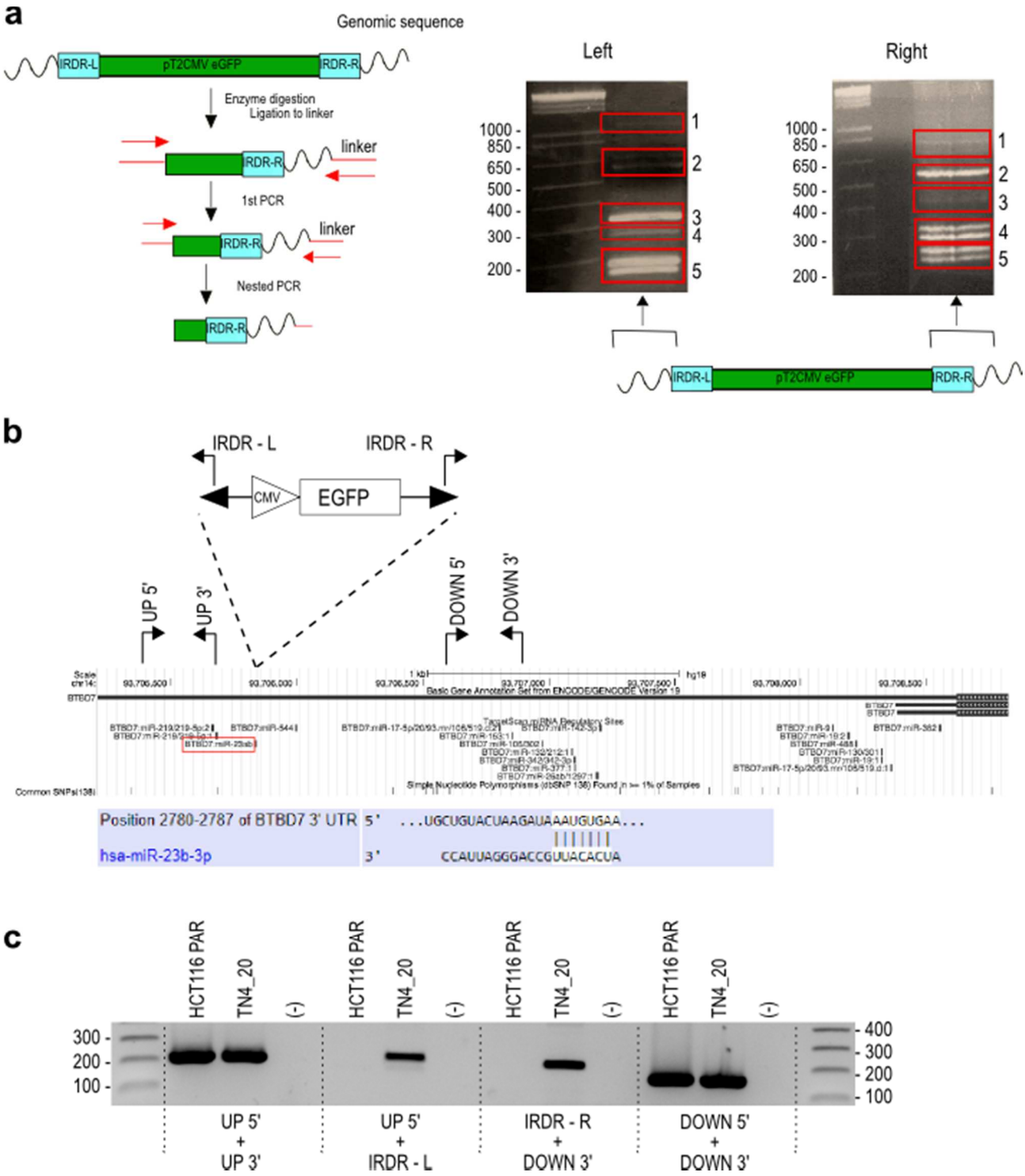
Supplementary Figure S4. Strategy of Sleeping Beauty forward genetic screen. (a) Schematic view of pT2-CMV-EGFP transposon and pSB100 transposase enzyme coding plasmids. **(b)** Scheme of the experimental procedure adopted to combine pT2-CMV-EGFP-based screening and fSCS assay in HCT116 cells. **(c)** FACS analysis of the percent of GFP positive HCT116 cells 2, 5, 7, 12 or 17 days after transient co-transfection of pT2-CMV-EGFP with wild type (pSB100x) or mutated (pSBD3) transposase enzyme. **(d)** Identification of clones resistant to fSCS upon combination of fSCS assay with the pT2-CMV-EGFP TN-based forward genetic screen. Representative images of fixed and stained colonies formed in 24-well plates by pT2-CMV-EGFP-transposed HCT116 clones after 2 rounds of fSCS (lower plate) or of Anoikis (upper plate, control condition). TN4_20 clone is the one showing the best capability to resist to fSCS. Each clone was tested in triplicate.

Supplementary Figure S5.



Supplementary Figure S5. TN4_20 clone shows increased invasive capabilities both in vitro and in vivo. (a) Invasion assay: quantification of HCT116, TN4_Sorted and TN4_20 cells motility in transwell migration assay. Bars indicate mean +/- std of 2 independent biological replicates. **(b)** Representative images at 10x magnification of H/E stained lungs of mice intracaecally injected with TN4_Sorted and TN4_20 cells.

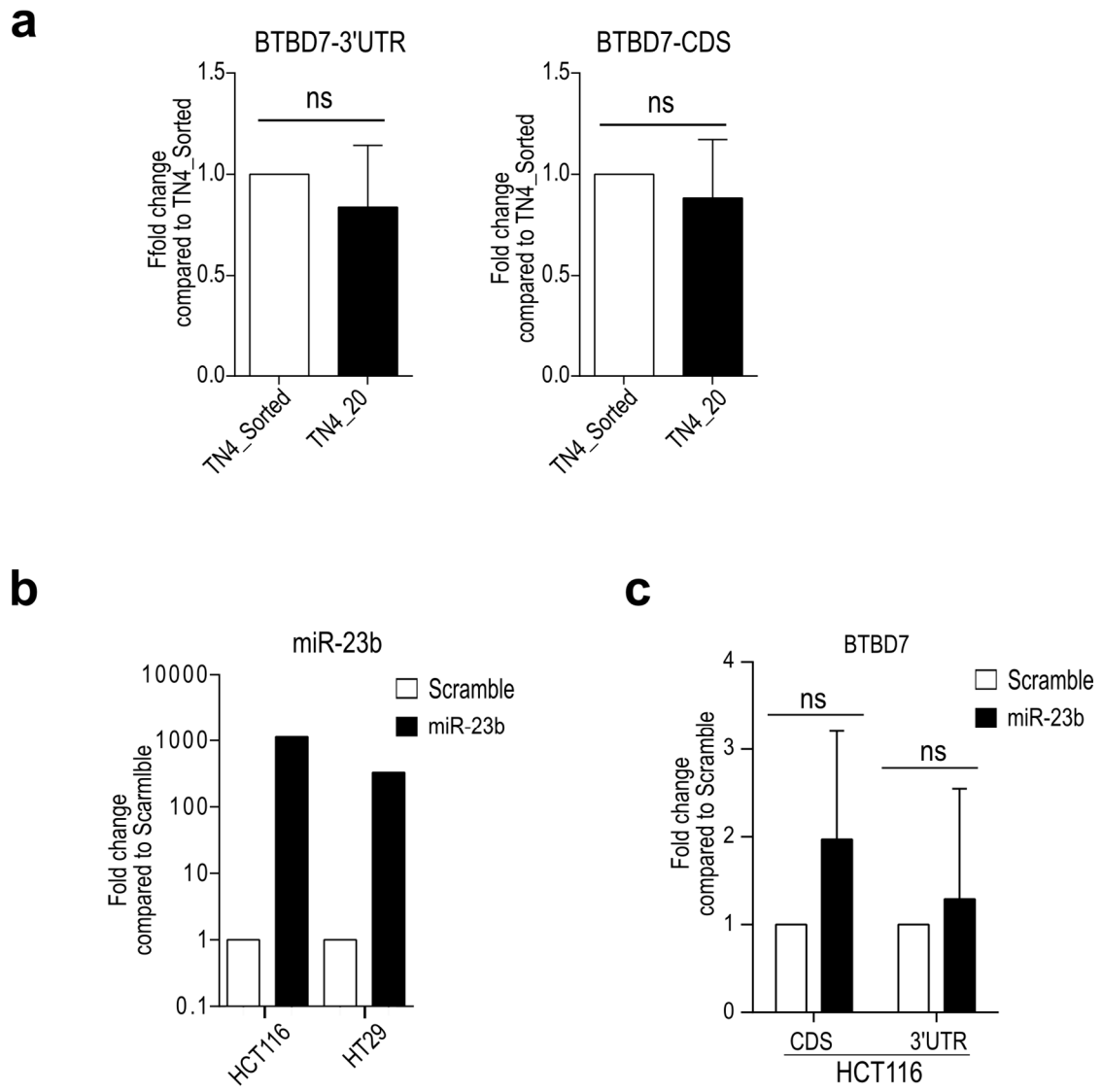
Supplementary Figure S6.



Supplementary Figure S6. Identification and confirmation of TN insertions in TN4_20 cell clone within the 3'UTR of BTBD7 gene close to miR-23b target site. (a) Left: Schematic depiction of Linker Mediated PCR. For more details, see experimental procedures. Green rectangle = pT2 CMV EGFP TN; Mirrored blue rectangles = TN inverted/direct repeats left (IR/DR-L) or right (IR/DR-R); Red

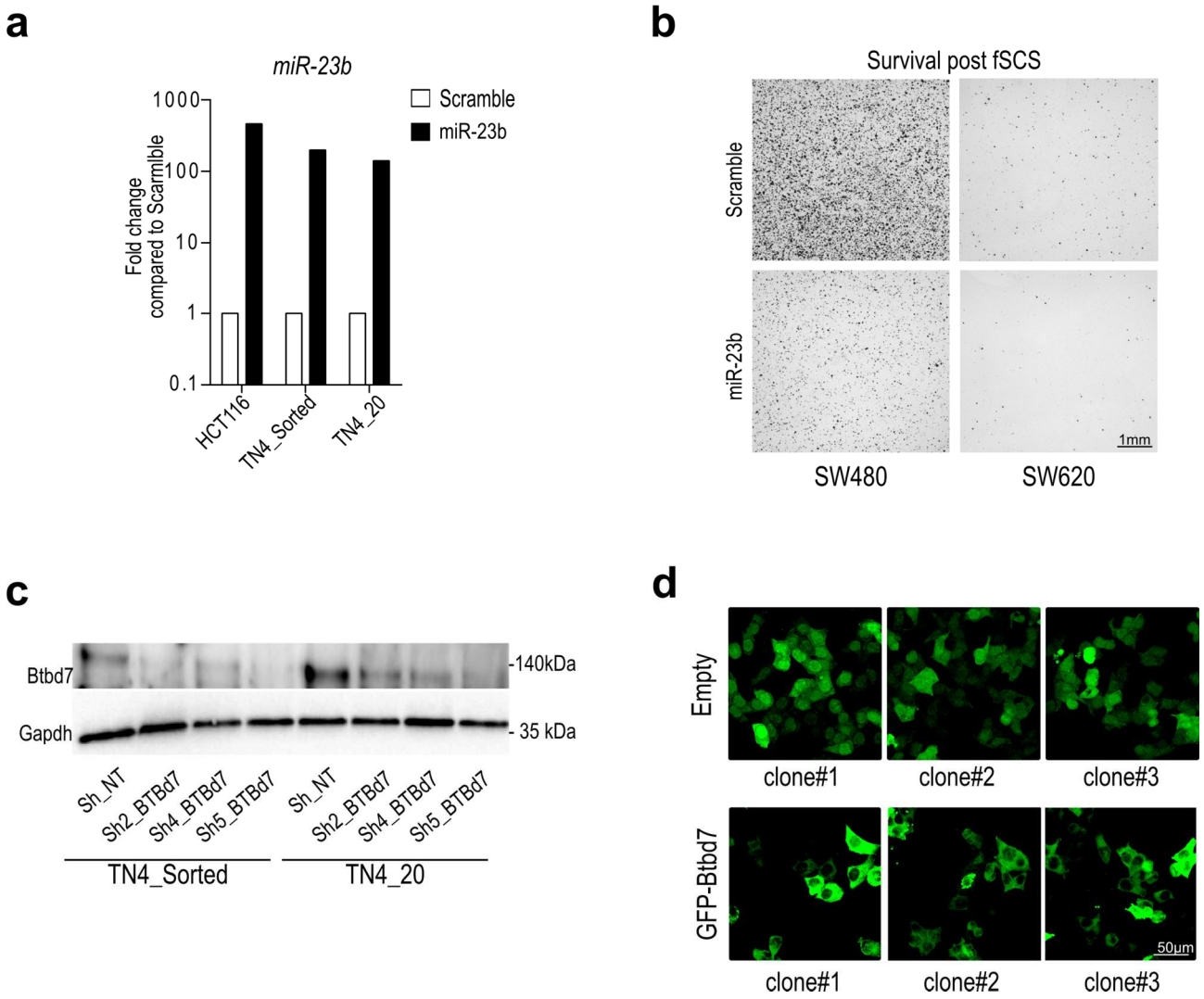
lines = Linkers ligated to genomic DNA for Linker Mediated PCR; Red arrows: primers complementary to Linkers. Right: agarose gel images of the multiple bands obtained by Linker Mediated PCR in correspondence of TN IR/DR-L (Left) or IR/DR-R (Right) as indicated by the scheme. For each IR/DR five bands with different molecular weight (indicated by red rectangles) were obtained and successively sequenced to retrieve TN positions (Suppl. Table S1). **(b)** Original screenshot from UCSC genome browser showing the region of BTBD7 3'UTR. The exact position of pT2-CMV-EGFP-TN insertion in the 3'UTR of BTBD7 gene as identified by Linker Mediated PCR is indicated. The position of 6 different primers used to confirm TN insertion in the 3'UTR of BTBD7 gene are indicated (UP5', UP3'; IRDR-L, IRDR-R; DOWN5', DOWN3') **(c)** Confirmation by PCR of the pT2-CMV-EGF-TN insertion in the 3'UTR of BTBD7 gene using the primers described above, which is present only in the TN4_20 clone and not in parental cells.

Supplementary Figure S7.



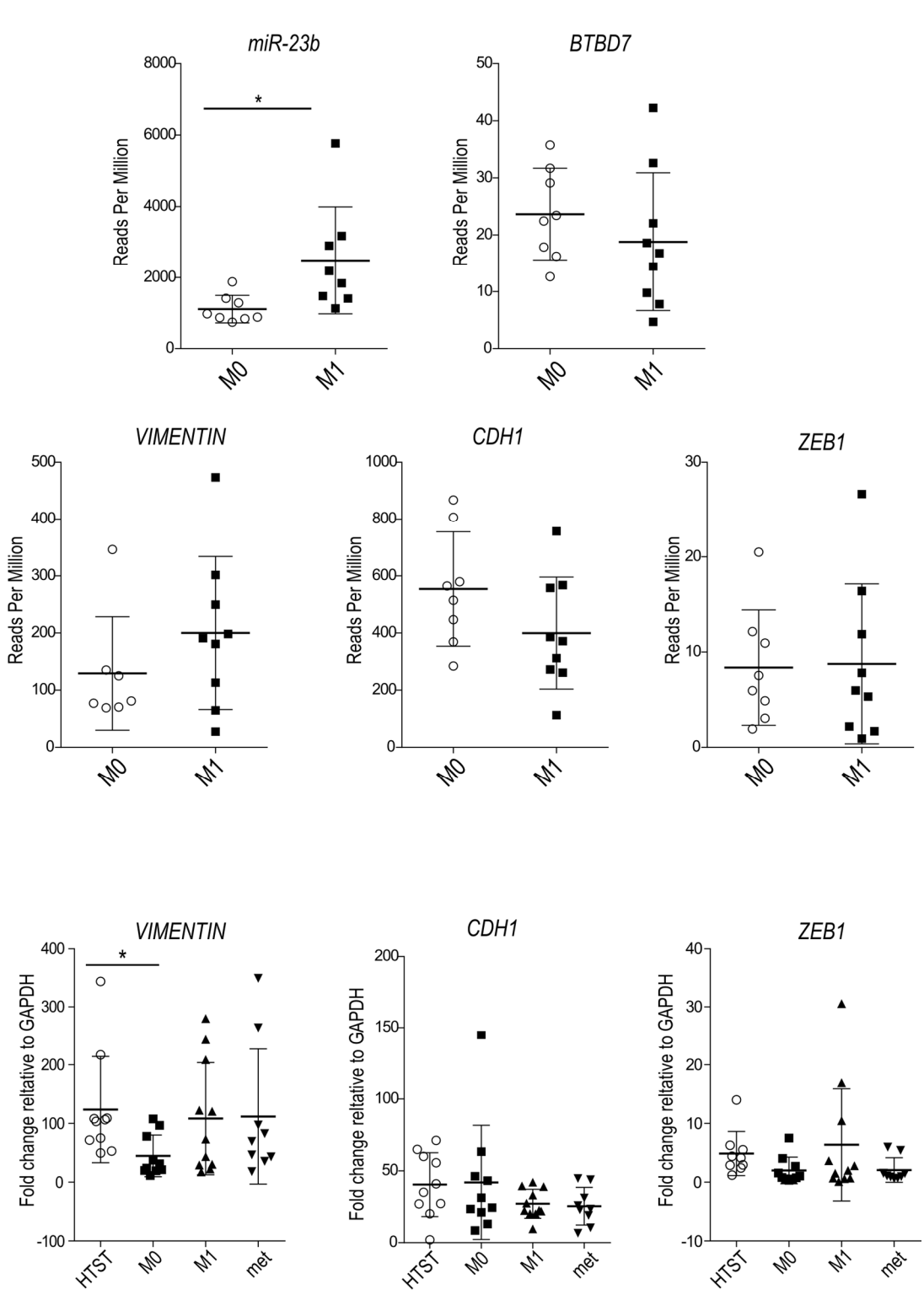
Supplementary Figure S7. TN insertion in TN4_20 clone and miR-23b transfection do not affect BTBD7 mRNA gene expression levels. (a) qRT-analysis of BTBD7 3'UTR (left) or CDS (right) expression in TN4_Sorted or TN4_20 cells. Bars indicate mean +/- std of two independent biological replicates (ns, not significant, using paired t test with unequal variances). **(b)** Representative qRT-analysis of miR-23b expression in HCT116 and HT29 cells transiently transfected with miR-23b precursor compared to scramble. **(c)** qRT-analysis of BTBD7 3'UTR or CDS expression in HCT116 cells transiently transfected with miR-23b precursor compared to Scramble. Bars indicate mean +/- std of two independent biological replicates (ns, not significant, using paired t test with unequal variances).

Supplementary Figure S8.



Supplementary Figure S8. Effect of miR-23b transfection or BTBD7 silencing in several cell models. (a) Representative qRT-analysis of miR-23b expression in HCT116, TN4_Sorted and TN4_20 cells transiently transfected with miR-23b precursor compared to Scramble. **(b)** Representative images of fixed and stained post-fSCS colonies formed in 10 cm plates by SW480 and SW620 after transfection of miR-23b precursor or Scramble. **(c)** Western Blot analysis of Btbd7 expression in TN4_Sorted and TN4_20 cells after transduction of cells with three different Sh_RNAs targeting BTBD7 (Sh2_BTBD7; Sh4_BTBD7; Sh5_BTBD7). **(d)** Representative confocal images of HCT116 GFP-Btbd7 clones #1, #2, #3 or GFP-Empty clones #1, #2, #3.

Supplementary Figure S9.



Supplementary Figure S9. Gene expression of selected genes in a subset of TCGA CRC samples and from CRO-Biobank CRC samples. (a) RNA-seq analysis of primary CRC without (M0; n=8) or with metastasis (M1; n=8) from The Cancer Genome Atlas database. Data are expressed as RPM (Reads per Million). Dot plot indicate mean +/- std (*p<0.05; using unpaired t test with unequal variances). **(b)** qRT-analysis of *ZEB1*, *CDH1* and *VIMENTIN* expression in HTST (Healthy Tissue

Surrounding Tumor) (n=10), in primary CRC samples without (M0; n=10), or with metastasis at diagnosis (M1; n=11), and in liver metastatic samples (met; n=9). Bars indicate mean +/- std (*p<0.05, using unpaired t test assuming unequal variances).