Stem Cell Reports
Article

ISSCR

OPEN ACCESS

Resistance to Taxanes in Triple-Negative Breast Cancer Associates with the Dynamics of a CD49f+ Tumor-Initiating Population

Jorge Gómez-Miragaya,[1,10] Marta Palafox,[1,10,11] Laia Paré,[2] Guillermo Yoldi,[1] Irene Ferrer,[1,12] Sergi Vila,[1,13] Patricia Galván,[2,3] Pasquale Pellegrini,[1,14] Hector Pérez-Montoyo,[1,15] Ana Igea,[4] Purificación Muñoz,[1] Manel Esteller,[1,5,6] Angel R. Nebreda,[4,6] Ander Urruticoechea,[7,16] Idoia Morilla,[7] Sonia Pernas,[7] Fina Climent,[8] María Teresa Soler-Monso,[8] Ana Petit,[8] Violeta Serra,[9] Aleix Prat,[2,3] and Eva González-Suárez[1,*]

[1]Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), Avinguda de la Gran Via, 199 – 203, L'Hospitalet de Llobregat, 08908 Barcelona, Spain
[2]Translational Genomics and Targeted Therapeutics in Solid Tumors, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), 08036 Barcelona, Spain
[3]Translational Genomics Group, Vall d'Hebron Institute of Oncology (VHIO), 08035 Barcelona, Spain
[4]Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology, 08028 Barcelona, Spain
[5]Unitat de Bioquímica i Biologia Molecular, Departament de Ciències Fisiològiques II, Universitat de Barcelona-IDIBELL, 08908 Barcelona, Spain
[6]Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain
[7]Breast Cancer Unit, Catalan Institute of Oncology, IDIBELL, 08908 Barcelona, Spain
[8]Pathology Department, University Hospital of Bellvitge, IDIBELL, 08908 Barcelona, Spain
[9]Experimental Therapeutics Group, Vall d'Hebron Institute of Oncology (VHIO), 08035 Barcelona, Spain
[10]Co-first author
[11]Present address: Experimental Therapeutics Group, Vall d'Hebron Institute of Oncology (VHIO), 08035 Barcelona, Spain
[12]Present address: Lung Cancer Clinical Research Unit, CNIO, 28029 Madrid, Spain
[13]Present address: August Pi i Sunyer Institute for Biomedical Research (IDIBAPS), Hospital Clínic - CIBEREHD, University of Barcelona Medical School, 08036 Barcelona, Spain
[14]Present address: Gladstone Institutes, San Francisco, CA 94107, USA
[15]Present address: Ability Pharmaceuticals, SL, 08290 Barcelona, Spain
[16]Present address: Oncologic Center (Onkologikoa), San Sebastian, 20014 Gipuzkoa, Spain
*Correspondence: egsuarez@idibell.cat
http://dx.doi.org/10.1016/j.stemcr.2017.03.026

SUMMARY

Taxanes are a mainstay of treatment for breast cancer, but resistance often develops followed by metastatic disease and mortality. Aiming to reveal the mechanisms underlying taxane resistance, we used breast cancer patient-derived orthoxenografts (PDX). Mimicking clinical behavior, triple-negative breast tumors (TNBCs) from PDX models were more sensitive to docetaxel than luminal tumors, but they progressively acquired resistance upon continuous drug administration. Mechanistically, we found that a CD49f+ chemoresistant population with tumor-initiating ability is present in sensitive tumors and expands during the acquisition of drug resistance. In the absence of the drug, the resistant CD49f+ population shrinks and taxane sensitivity is restored. We describe a transcriptional signature of resistance, predictive of recurrent disease after chemotherapy in TNBC. Together, these findings identify a CD49f+ population enriched in tumor-initiating ability and chemoresistance properties and evidence a drug holiday effect on the acquired resistance to docetaxel in triple-negative breast cancer.

INTRODUCTION

Triple-negative breast cancer (TNBC) is a heterogeneous disease with divergent profiles of chemosensitivity and prognosis (Perou et al., 2000; Prat et al., 2010; Shah et al., 2012; Yu et al., 2013). Standard chemotherapy with anthracyclines and taxanes is the mainstay treatment. A subset of TNBCs shows increased chemosensitivity compared with other breast cancer subtypes; however, for a significant number of patients, overall prognosis is poorer, with high risk of early relapse. Once metastases appear the patient median survival is drastically reduced (Andre and Zielinski, 2012). Despite enormous efforts, the cause of resistance to chemotherapy agents, including taxanes, is unclear (Bonnefoi et al., 2011). There remains an urgent unmet need to identify the population of patients that will benefit from taxanes, on one hand, and to determine the mechanisms of resistance, on the other.

There is increasing evidence that in a variety of neoplasia, including breast cancer, only a subset of cancer cells are capable of reconstituting the tumor after transplantation. These cells called cancer stem cells (CSCs) or tumor-initiating cells (TICs), have the ability to self-renew and regenerate tumor heterogeneity (Al-Hajj et al., 2003) and show intrinsic resistance to conventional chemotherapies, leading to recurrence or metastasis. In fact, breast tumors from patients who received neoadjuvant chemotherapy are substantially enriched for CSCs compared with tumors of untreated patients (Yu et al., 2007), suggesting that anticancer agents kill the bulk of tumor cells, but spare the CSCs (Dean et al., 2005). In breast cancer, a variety of markers (CD44, CD24, EpCAM, CD49f, CD133/2, CD10, and ALDH activity) have been shown to identify CSCs (Al-Hajj et al., 2003; Bachelard-Cascales et al., 2008; Li et al., 2008; Lim et al., 2009; Stingl et al., 2006). However, it is still unclear whether all these markers are appropriate

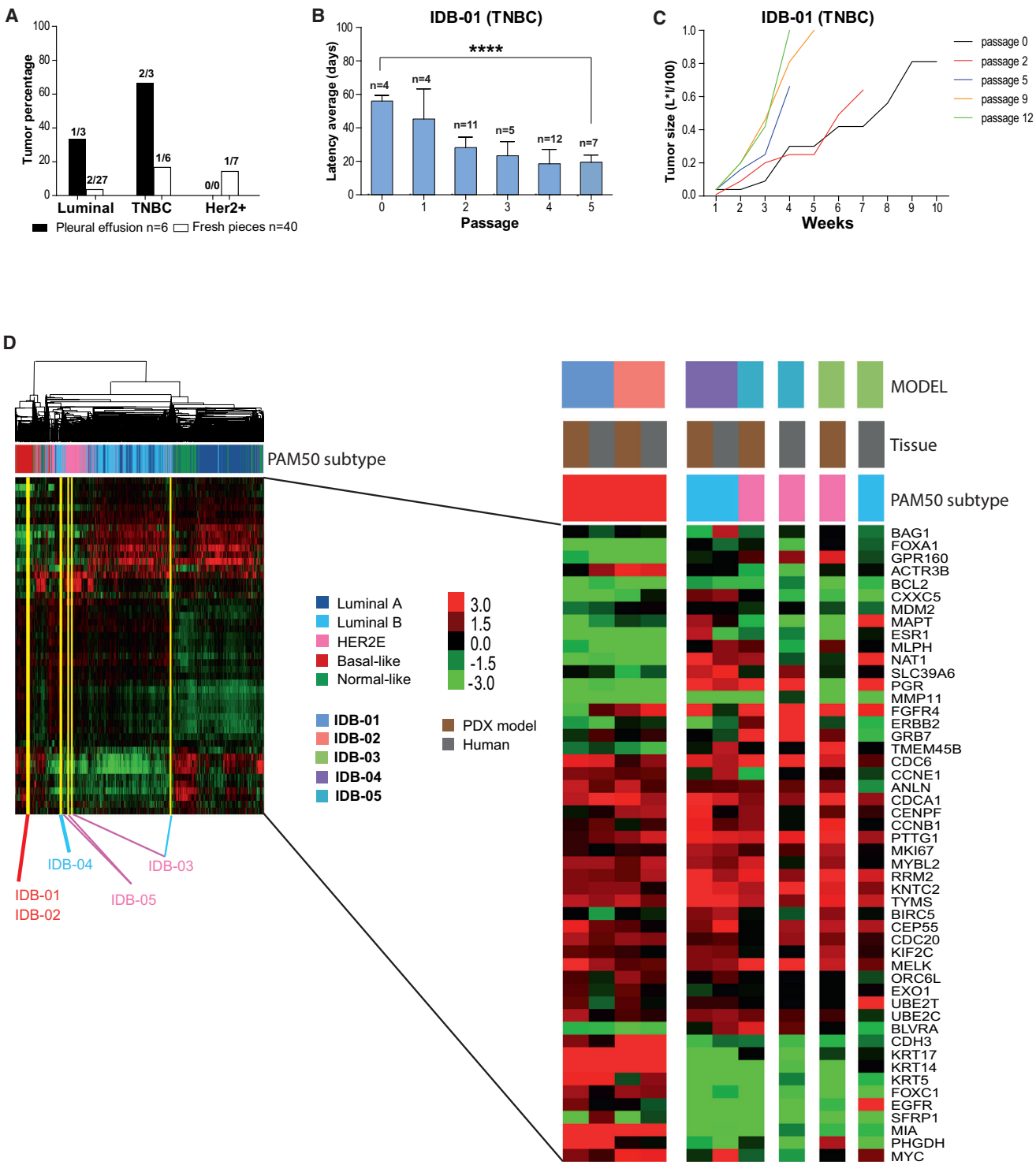


Figure 1. Generation and Characterization of PDX Models of Human Breast Cancer

(A) Percentage of palpable tumors that engrafted relative to total number of independent patient samples, classified by subtype and source. The total number of original patient samples is indicated and mice that did not survive for at least 60 days after surgery were excluded.

(B) Tumor latency in IDB-01 at the indicated passages. Total number of tumors (n), mean, SD, and t test p values are shown. ****$p < 0.0001$.

(legend continued on next page)

for the different breast cancer subtypes, and further studies are necessary to identify the population of TICs and their functionality in each type of tumors.

The lack of appropriate tools and models has hindered our efforts to gain insight into the mechanisms of drug resistance. The best approach to investigate acquired resistance requires analysis of primary or metastatic samples collected before and after recurrence, but these paired sensitive/resistant samples are often difficult to obtain. To advance our knowledge in clinical breast cancer and the molecular mechanisms of resistance, we have generated breast cancer patient-derived orthoxenografts (PDXs), which allow the amplification and perpetuation of human tumors by serial passages. Our panel of breast cancer PDXs recapitulates the heterogeneity of the clinical disease and constitutes a unique tool for studying the biological mechanisms of clinical response to taxanes and acquisition of resistance. We demonstrate that a CD49f+ cell population with tumor-initiating ability and increased resistance to taxanes is present in the initially sensitive TNBC tumors and expands during continued exposure to the drug in vivo, contributing to taxane resistance and tumor recurrence. Remarkably, the transcriptional differences observed between the CD49f+ population of sensitive and resistant tumors accompany and may contribute to the acquisition of chemoresistance. Finally, we demonstrate that docetaxel sensitivity is recovered in the absence of the drug and associates with changes in the CD49f+ population.

RESULTS

PDX Models Resemble Human Tumors of Origin in Early Passages

PDX were generated as described (DeRose et al., 2011; Zhang et al., 2013; Table S1). Increased tumor rates and shorter latency to tumor formation was observed in samples derived from pleural effusions compared with tumor pieces. The TNBC engrafted better than luminal tumors and all palpable tumors derived from grade 3 human samples (Figures 1A, S1A, and S1B). Of the mammary glands, 52% with no palpable tumor contained human mammary epithelium, mostly normal ducts and grade 1 intraductal carcinoma indicating engraftment of these low-grade lesions (Figures S1C and S1D; Table S1). Tumor lines (Table 1, yellow in Table S1) were maintained by consecutive rounds of transplantation, and include two TNBC models derived from pleural effusions, the second one a BRCA1 mutant (IDB-01, IDB-02); two luminal/HER2-negative models (IDB-03 and IDB-04) derived from tumor pieces and pleural effusion, respectively; and one (IDB-05) derived from a tumor piece of a triple-positive (ER+ PR+ HER2+) breast cancer. In most models, shorter latency and faster tumor growth were observed in late passages (Figures 1B, 1C, S1E, and S1F). Thus, as demonstrated previously (DeRose et al., 2011; Dobrolecki et al., 2016; Zhang et al., 2013), establishment of PDX models was associated with increased tumor aggressiveness and poor prognosis.

Expression analyses of markers used in the clinical setting for histopathological tumor classification and selection of treatment (ER, PR, HER2, CK5/6, CK18, and p53), in parental human tumors and PDX tumors at early (0–1) and late passages (4–8) demonstrate that PDX retain most human characteristics in the early passages, but occasional changes are observed in some models (Table 1; Figures S1G and S2A). ER and PR mRNA and protein expression was detected in tumors from all passages of the luminal models IDB-04 and IDB-05 (Figures S2A and S1G), but only IDB-05 required estrogen/progesterone pellets to grow (Figure S1H). IDB-03, ER+ and PR+ in the patient, lost ER and PR expression in the PDX and a population of p53+ cells was enriched (Table 1; Figure S2A and S1G). After surgically resection of tumors, most models developed local relapses and metastases to clinical relevant sites (Table 1; Figure S2B).

Next, we performed intrinsic subtyping of our 5 PDX models and their corresponding human tumors of origin using the PAM50 subtype predictor (Parker et al., 2009), and clustered these samples with 1,834 breast tumor samples representing all subtypes (Prat et al., 2015b). Mimicking the intrinsic subtypes of their corresponding human tumors, the two TNBC models were identified as basal-like, IDB-04 (HR+/HER2–) as luminal B, and the HER2+ IDB-05 as HER2 enriched (HER2-E). Interestingly, the human tumor of origin for IDB-03 was identified as luminal B but the PDX was identified as HER2-E by PAM50 without HER2 overexpression (Figure 1D). As reported in similar PDX collections (Dobrolecki et al., 2016), our mouse grafts retain initial human tumor characteristics, but some models change during serial passages in mice, which may reflect evolution of the clinical disease.

(C) Tumor growth in IDB-01, calculated as L × I (mm × mm)/100 versus time (weeks). Each line represents a representative tumor.
(D) Unsupervised clustering using the PAM50 genes across the PDX models, human tumors of origin, and 1,834 human breast cancer clinical samples (Prat et al., 2015b). The type of sample and the subtype call of each sample are shown. Each square represents the relative transcript abundance. All PDX tumors were from passage 5.
See also Table S1; Figures S1 and S2.

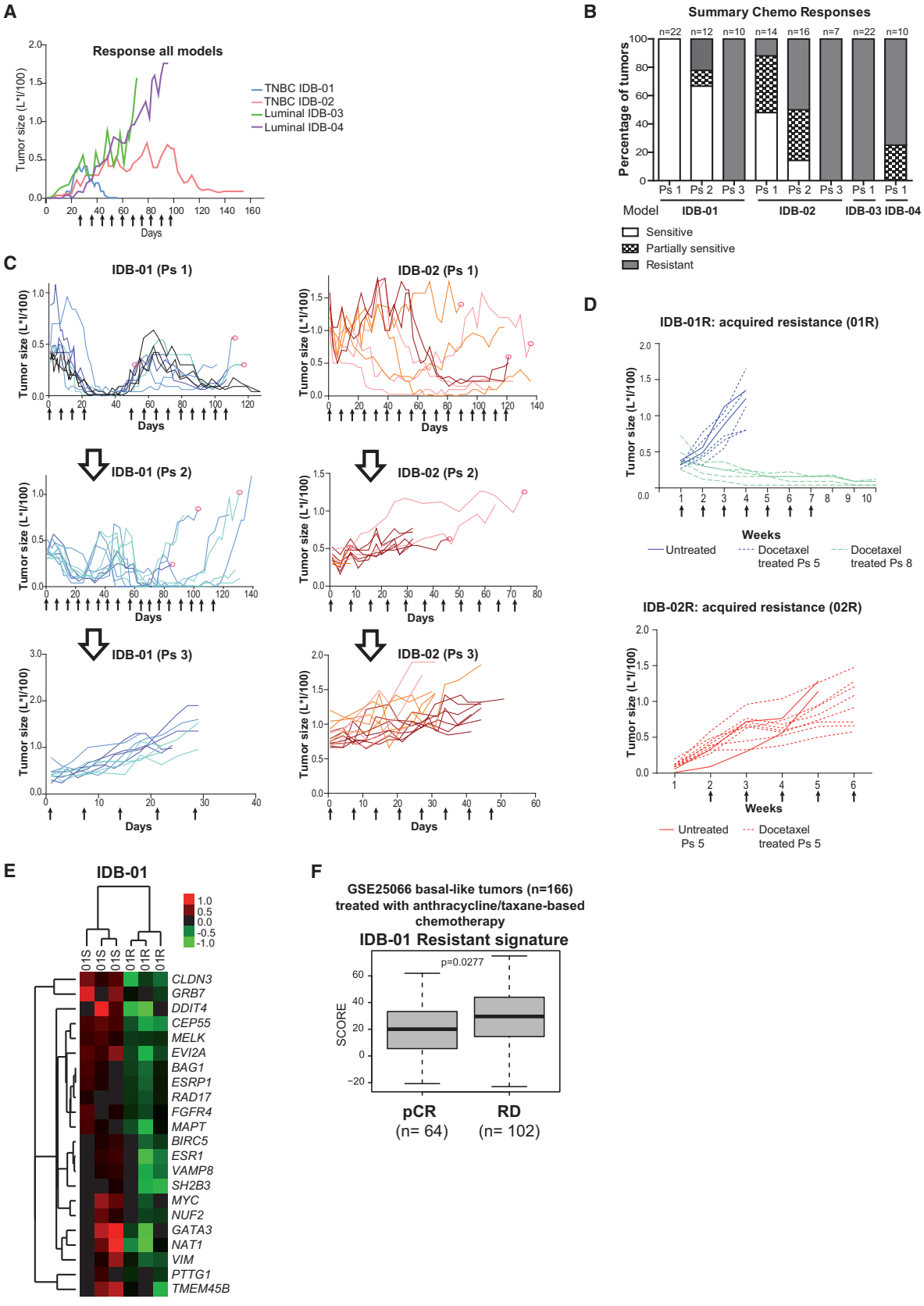
Table 1. Main Characteristics of Human Tumor of Origin and Mouse Grafts in Five Established IDB Models

	Model	IDB-01 (TNBC)	IDB-02 (TNBC)	IDB-03 (Luminal)	IDB-04 (Luminal)	IDB-05 (HER2+)
Mouse	phenotype early passage	ER-PR-HER2-CK5/6+ CK18+ p53-	ER-PR-HER2-CK5/6+ CK18+ p53- BRCA1 mut	ER+ PR+ HER2-CK5/6- CK18+ p53+ BRCA2 mut	ER+ PR+ HER2-CK5/6- CK18+ p53-	ER+ PR+ HER2+ CK5/6-CK18+ p53-
	phenotype late passage	no change	no change	loss of ER and PR	no change	no change
	passage	13	8	16	7	9
	latency (p5) (days)	19	42	18	63	27
	growth without hormone pellets	yes	yes	yes	yes	no
	Local relapse (%)	17.24 (n = 116)	6.94 (n = 144)	20.79 (n = 178)	23.40 (n = 94)	8.54 (n = 94)
	axillary metastasis (%)	10	10	46	0	8
	lymph ND metastasis (%)	14	33	41.2	0	50
	lung metastasis (%)	20	20	10	0	20
	metastasis to other sites	ND	Yes (brain, axillary, subcutaneous)	Yes (bone, kidney)	Yes (liver)	ND
Human	subtype	TNBC grade 3	TNBC grade 3	luminal grade 3	luminal grade 3	HER2+ grade 3
	IHC	ER-PR-HER2-CK5/6+ CK18+ p53-	ER-PR-HER2-CK5/6+ CK18+ p53- BRCA1mut	ER+ PR+ HER2-CK5/6- CK18+ p53+	ER+ PR+ HER2-CK5/6- CK18+ p53-	ER+ PR+ HER2+ CK5/6-CK18+ p53-
	source	pleural	pleural	tumor pieces	pleural	tumor pieces
	treatment	FEC, docetaxel, capecitabine	FEC, docetaxel	not treated	paclitaxel, carboplatin, capecitabine	not treated

ER, estrogen receptor; PR, progesterone receptor; CK, cytokeratin, FEC, triple treatment composed of 5-fluorouracil, epirubicin and cyclophosphamide. Frequency of tumor relapse per mammary gland (local relapse) and metastasis is indicated in each model. Only mice that survived for at least 60 days after primary tumor excision with no relapse/metastasis were considered as relapse/metastasis free. All metastases were confirmed by pathologists. ND, not determined. See also Table S1.

Basal-like PDX Are Initially Sensitive to Docetaxel but Acquire Resistance after Continuous Exposure to the Drug In Vivo

Next, we tested the sensitivity of orthotopic mouse models to docetaxel, one of the most commonly used chemotherapeutics in breast cancer and other solid tumors (Figure S3A). According to docetaxel response, tumors were classified as sensitive when the treatment induced complete tumor regression; partially sensitive when the treatment interfered with tumor growth inducing complete regression in some tumors but not in others; and resistant when tumors continued growing despite docetaxel treatment. In line with these criteria, luminal tumors from IDB-03 and IDB-04 were resistant to docetaxel, the TNBC IDB-01 model was sensitive (IDB-01S), and the TNBC IDB-02 was partially sensitive to the drug (Figures 2A, 2B, and S3B). Despite the initial pathological complete response, all IDB-01 tumors started growing again 30–60 days after treatment interruption. In the second round of treatment, more doses of docetaxel were required to eliminate tumors and a more heterogeneous response between individual tumors was observed (partially sensitive tumors). This behavior was accentuated during consecutive docetaxel treatments and the tumors became resistant in passage 3 (Figures 2B and 2C). Resistance was retained for at least two passages in the absence of docetaxel, as IDB-01-resistant tumors (IDB-01R, passage 5) grew at comparable growth rates irrespective of docetaxel treatment (Figure 2D). Importantly, tumors growing without the selective pressure of docetaxel partially regained sensitivity after five passages (IDB-01R, passage 8), which demonstrates that taxane resistance can be reverted



(legend on next page)

(Figure 2D). No differences in latency to tumor formation or tumor growth were observed between IDB-01S (sensitive tumors of origin) and the resistant ones, IDB-01R, derived from them (Figures S3C and S3D). Gene expression analyses of IDB-01S and IDB-01R tumors identified a signature (22 downregulated genes in IDB-01R) that was predictive of residual disease after anthracycline/taxane-based therapy in 166 patients with basal-like disease (GSE25066) (Hatzis et al., 2011) and poor survival in the TCGA dataset, highlighting the clinical relevance of our sensitive and resistant PDX pairs (Figures 2E, 2F, and S3E). In IDB-02 an already heterogeneous response was observed after the first doses, and docetaxel treatment could not be interrupted in most mice (Figures 2B, 2C, and S3B). Tumors started growing very fast after interruption of the treatment and became resistant in passage 2. A third passage and additional docetaxel treatments did not change tumor growth, demonstrating that tumors had acquired resistance to docetaxel, which was retained for at least two passages (Figures 2B–2D). IDB-02R resistant tumors showed similar latency as IDB-02S sensitive tumors but grew significantly faster (Figures S3C and S3D). These results demonstrate that our triple-negative PDX tumors are more sensitive to docetaxel than the luminal ones. In the clinic, a better response to chemotherapy is observed in TNBC compared with luminal tumors, and in some studies taxanes have been shown to be superior to anthracyclines in this subtype (Kim et al., 2010; Martin et al., 2011). Moreover, initially sensitive PDX tumors gradually became less responsive to docetaxel and acquired resistance after continuous exposure to the drug, mimicking the clinical scenario.

Docetaxel Acquired Resistance Is Accompanied by an Increase in the CD49f+ Population

It has been shown that chemotherapy efficiently eliminates the bulk tumor cells but spares the CSC population (Li et al., 2008). Thus, we analyzed the expression of markers previously shown to identify CSCs in our PDX tumor collection including paired sensitive and resistant tumors from IDB-01 and IDB-02 (Figure S4A-B). Variability in marker expression was detected between models with the same histological and molecular subtype. Docetaxel-resistant luminal tumors (IDB-03 and IDB-04) showed the highest percentages of EpCAM, CD49f, and CD24 cells, but the CD133 population was scarce. IDB-03 contained an abundant CD44+ population and ALDH activity, and is the only one expressing CD10. A CD133+ population was found in basal-like and HER2+ PDX. The CD44+ CD24– population, shown to identify human breast CSCs (Al-Hajj et al., 2003), was only detected in the TNBC IDB-02 (Figures S4B and S4C).

No significant changes in the expression of CD44, CD24, CD133, or CD10 were found between sensitive and resistant TNBC paired samples, neither in IDB-01 nor in IDB-02. The CD44+ CD24– population remained barely detectable in the chemoresistant models, and the ALDH+ population, based on ALDH enzymatic activity, was also comparable between paired sensitive and resistant tumors (Figure 3A). In contrast, the frequency of CD49f+ cells significantly increased in TNBC-resistant tumors compared with paired sensitive ones in both models. A significant increase in the frequency of EpCAM+ cells was also observed in IDB-01R compared with IDB-01S tumors (Figure 3A). Resistant tumors from IDB-01 and IDB-02 showed significantly higher mRNA expression levels of *CD49f* (*ITGA6*) but not *EpCAM*, than the corresponding sensitive tumors (Figure 3B).

We next sought to investigate the clinical relevance of our findings by analyzing different clinical datasets. In basal-like tumors from the EORTC 10994/BIG-1-00 clinical trial (Bonnefoi et al., 2011), higher expression of *CD49f* and *EpCAM* was associated with a non-pathological

Figure 2. TNBC PDX Tumors were Sensitive to Docetaxel and Acquired Resistance after Continuous Treatment, whereas the Luminal Tumors were Resistant

(A) Representative kinetics of tumor growth during docetaxel treatment. Docetaxel treatment (20 mg/kg i.p., once per week) started when tumors reached 6 × 6 mm. Each line illustrates a representative tumor (passages 4–14).

(B) Percentage of sensitive, partially sensitive or resistant tumors of each model to docetaxel. Total number of tumors (n) and passage are indicated.

(C) Representative kinetics of tumor growth during acquisition of resistance to docetaxel in the basal-like IDB-01 and IDB-02. Each line represents one tumor and each color represents an independent sensitive tumor of origin. Ps, passage treated with docetaxel. Red circles indicate the tumors that were transplanted.

(D) Representative kinetics of tumor growth during docetaxel treatment after acquisition of resistance to taxanes. Each line represents one tumor. IDB-01R tumors were analyzed after growing for two and five passages, respectively, in the absence of docetaxel.

(E) Supervised expression analysis of the genes found differentially expressed between IDB-01R and IDB-01S tumors. Each square represents the relative transcript abundance.

(F) Association of IDB-01 resistant signature with chemotherapy response in 166 patients with basal-like breast cancer (Hatzis et al., 2011). Response was measured as pathological complete response (pCR) or residual disease (RD). Mean values, box and whiskers (min to max) and t test p values are shown.

(A, C, and D) Arrows represent docetaxel doses. See also Figure S3.

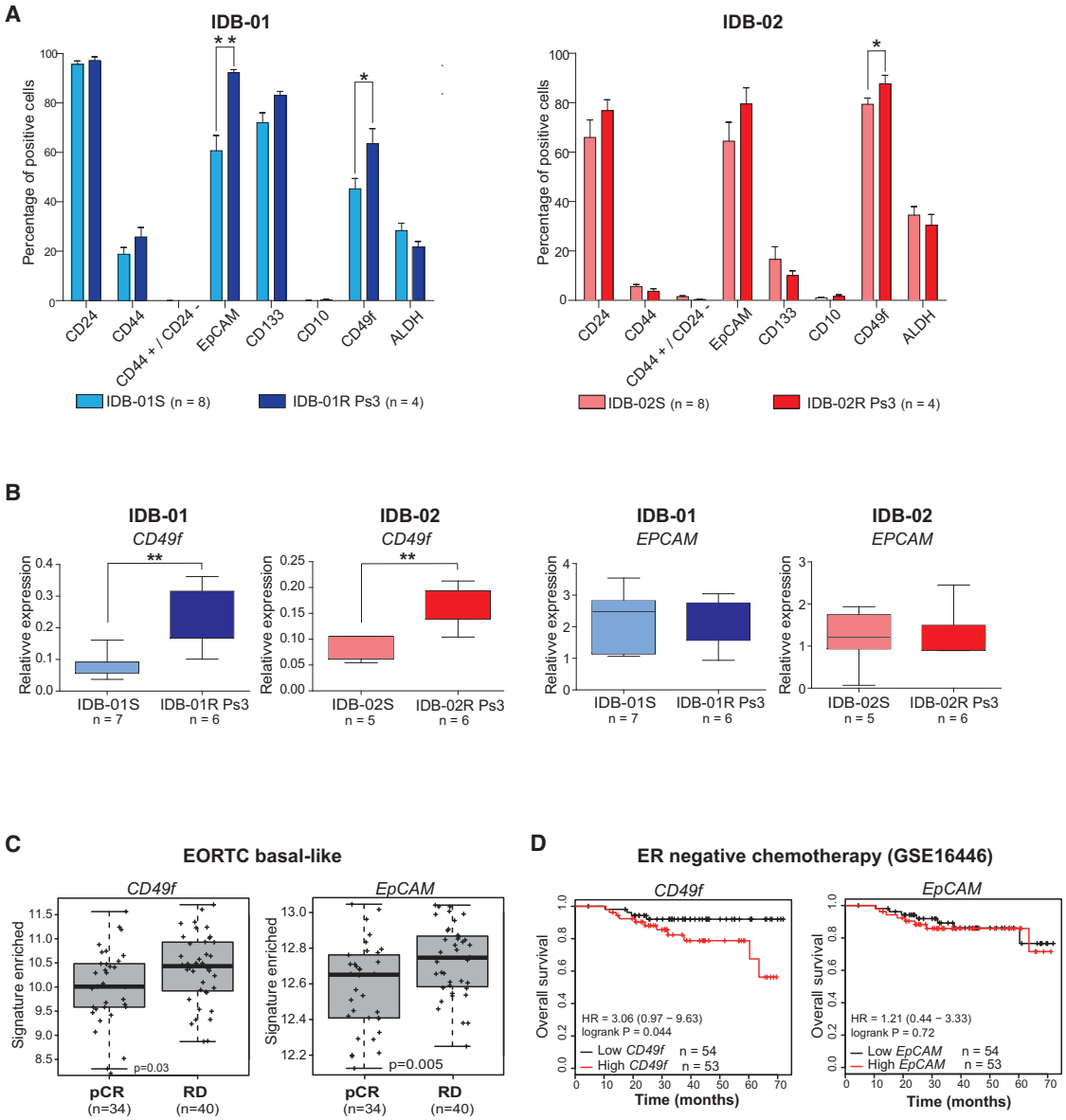


Figure 3. The CD49f+ Population Is Enriched after Acquisition of Resistance to Docetaxel

(A) Frequency of indicated markers within the H2Kd– population in IDB-01 and IDB-02, sensitive and resistant tumors analyzed by flow cytometry at passage 3, at least 5 days after the last docetaxel treatment. Total number of tumors analyzed (n) mean values, SDs and t test p values are shown. *0.01 < p < 0.05; **0.001 < p < 0.01.

(B) Box and whiskers (min to max) graph showing expression levels of *CD49f* and *EpCAM* mRNA relative to *PPiA* in additional sensitive and resistant tumors measured by qRT-PCR. Determinations were done in triplicate and means are used. t test p values for significant differences are shown. **0.001 < p < 0.01.

(C) Box and whiskers (min to max) graph showing association of *CD49f* and *EpCAM* with chemotherapy response in 74 patients with basal-like breast cancer EORTC (Bonnefoi et al., 2011). Response was measured as pathological complete response (pCR) or residual disease (RD). t test p values are shown.

(D) Kaplan-Meier analysis of overall survival of ER-tumors all treated with chemotherapy using *CD49f* and *EpCAM* mRNA expression in the clinical dataset (GSE16446) from the TOP TRIAL (Desmedt et al., 2011).

See also Figures S3 and S4.

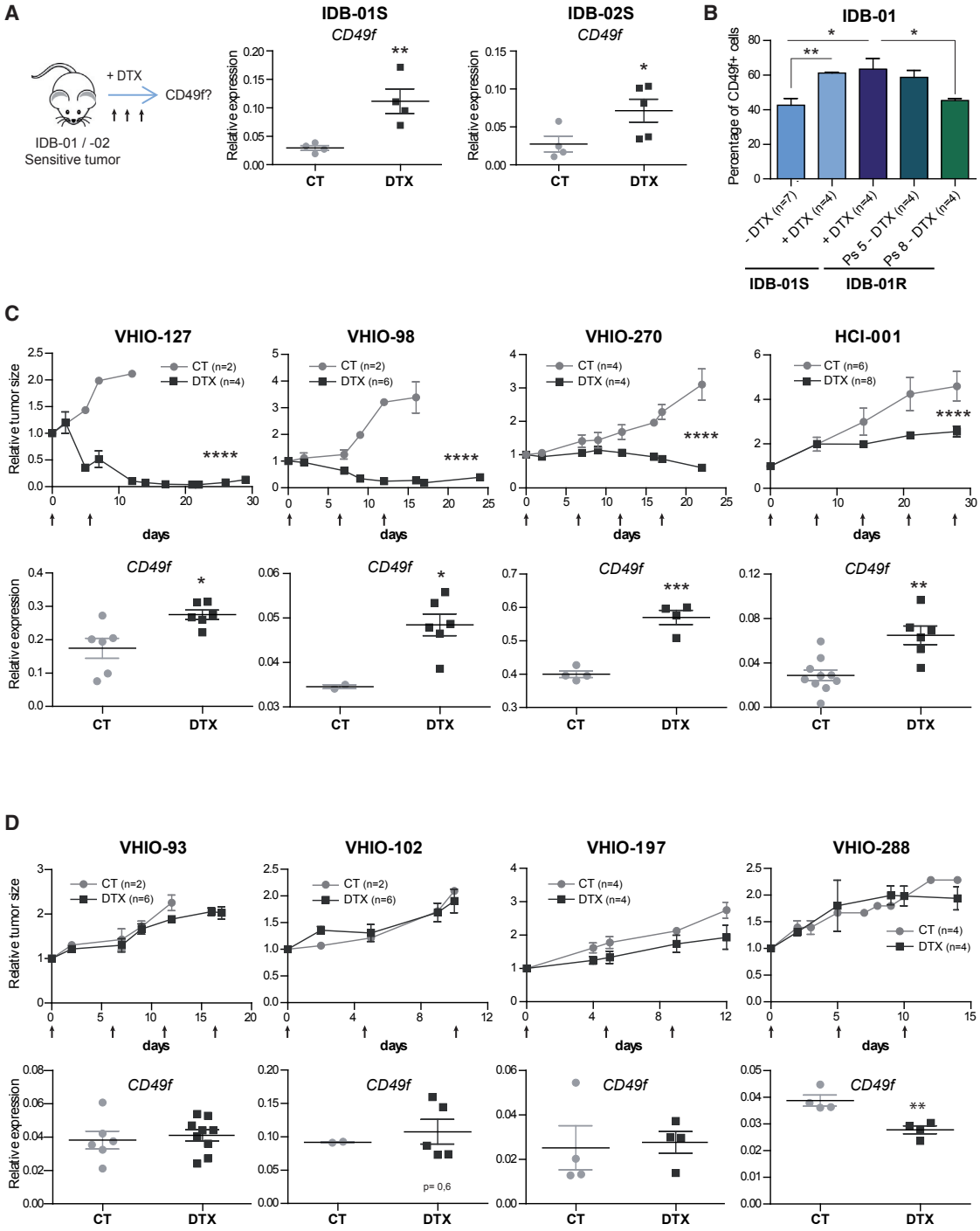


Figure 4. CD49f Expression Increases in Residual Disease of Most TNBC PDX Tumors after Treatment with Docetaxel, but Not on Resistant Tumors

(A) Scheme of short-term docetaxel treatment and *CD49f* mRNA expression levels in sensitive tumors from IDB-01S and IDB-02S after short-term treatment with docetaxel (DTX) and in untreated controls (CT). Each dot represents one tumor. *0.01 < p < 0.05; **0.001 < p < 0.01.

(B) Frequency of CD49f+ cells within H2Kd– in IDB-01S tumors after two to three doses of docetaxel and in IDB-01R tumors that have been treated with docetaxel (at least 5 days after last treatment) or growing in the absence of docetaxel for two and five passages. *0.01 < p < 0.05; **0.001 < p < 0.01.

(legend continued on next page)

complete response (non-pCR) after chemotherapy (Figure 3C). Using GOBO, the Gene expression-based Outcome for Breast cancer Online tool (Ringner et al., 2011), high expression levels of *EpCAM* and *CD49f* combined predicted a reduction in distal metastasis-free survival in basal-like tumors (Figure S3F). Associations with poor overall survival were obtained for *CD49f*, but not *EpCAM*, in other ER-negative or basal-like tumor samples after chemotherapy treatment (Clarke et al., 2013; Desmedt et al., 2011) (Figures 3D, S3G, and S3H). These results demonstrate that, whereas CD44+ CD24– and ALDH activity are not altered, the percentage of the CD49f+ population significantly increases during the acquisition of resistance to docetaxel in basal-like breast cancer.

A Chemoresistant CD49f+ Population Is Present in Most TNBC Tumors

We hypothesized that a chemoresistant CD49f+ population is present in the original sensitive tumors. To test this hypothesis we analyzed *CD49f* mRNA expression in IDB-01S and IDB-02S tumors after two to three doses of docetaxel treatment when tumors were shrinking, and found a significant increase in *CD49f* mRNA expression in the residual disease of both PDX tumors (Figure 4A). Next, we evaluated by flow cytometry the percentage of cells expressing CD49f in residual disease and found that the frequency of CD49f+ cells in residual disease of IDB-01S after docetaxel treatment increases by 20%; these levels are comparable with those of resistant IDB-01R tumors, indicating that the surviving population is enriched in CD49f+ cells (Figure 4B). Importantly, in IDB-01R tumors that regained sensitivity to taxanes after growing in the absence of docetaxel (passage 8), the frequency of the CD49f+ population decreases again to basal levels, similar to those found in sensitive tumors of origin (Figure 4B).

To evaluate whether a chemoresistant CD49f+ population could be found in other TNBC tumors, we analyzed *CD49f* expression after short-term in vivo treatment with docetaxel in 12 additional TNBC PDX tumors derived from patient samples (Bruna et al., 2016; DeRose et al., 2011). Four of these PDX tumors were resistant to docetaxel (no differences in tumor growth after docetaxel treatment), and eight showed different grades of sensitivity to the drug (tumors either shrank or showed tumor growth stabilization after two to four doses of docetaxel). After docetaxel treatment, an increase in *CD49f* mRNA expression levels was observed in residual disease of five out of the eight TNBC-sensitive tumors treated, whereas in resistant tumors *CD49f* expression remained unaltered (Figures 4C, 4D, and S5A). No changes in the expression of the most common partners of CD49f, CD29 (*ITGB1*) and CD104 (*ITGB4*), were observed between sensitive, resistant and residual disease in TNBC tumors (Figure S5B-D). The increase in *CD49f* expression in residual tumors suggests that CD49f+ chemoresistant cells are present in docetaxel-sensitive tumors and get enriched in residual disease.

In addition, we analyzed *CD49f* mRNA expression in five independent TNBC cell lines after 72 h of treatment with increasing concentrations of docetaxel. Different cell lines showed different grades of sensitivity to taxanes but, in four out of the five cell lines tested, a significant increase in *CD49f* mRNA expression was found in cells that survive docetaxel treatment compared with the untreated ones (Figure 5A). No changes in *CD49f* expression were observed at shorter time points with negligible cell death, suggesting that docetaxel does not induce *CD49f* expression and that the observed increase in residual disease, most probably represents the survival of a pre-existing CD49f+ population (Figure 5B). Higher levels of *CD49f* mRNA after paclitaxel treatment were also observed in some cell lines (Figure S5E). No changes in docetaxel sensitivity were observed in MDA-MB-436 cells upon stable reduction of CD49f expression with two independent short hairpin RNA constructs, ruling out a functional role for CD49f itself in chemoresistance of these cells (Figures 5C–5E and S5F).

Together these results demonstrate that higher expression of CD49f was observed in residual disease after docetaxel treatment for most TNBC-sensitive models (seven out of ten PDX models and four out of five cell lines), suggesting that despite the heterogeneity of the TNBC subtype a chemoresistant CD49f+ population is present in most TNBC.

CD49f+/hi Cells Show Enhanced Tumor-Initiating Ability and Resistance to Docetaxel

Next, we asked whether chemoresistant CD49f+ cells showed a higher tumor-initiating potential than CD49f– cells and could be responsible for tumor recurrence. Using fluorescence-activated cell sorting (FACS), we sorted the higher and lower quartile of tumor cells based on CD49f expression from IDB-01S and IDB-02S tumors and functionally tested their tumor-initiating potential (Figure 6A).

(C and D) Docetaxel-sensitive tumors (C) and docetaxel-resistant tumors (D). Top panels: tumor size of the indicated PDX tumors treated with docetaxel (20 mg/kg, arrows) and corresponding controls relative to the size at the first day of treatment. n = total number of tumors. ****$p < 0.0001$. Bottom panels: *CD49f* mRNA expression levels in PDX tumors after short-term treatment with docetaxel and in untreated controls. Each dot represents one tumor. *$0.01 < p < 0.05$; **$0.001 < p < 0.01$; ***$0.001 < p < 0.0001$.
(A–D) Mean values, SEM, and t test p values are shown in all cases.
See also Figure S5.

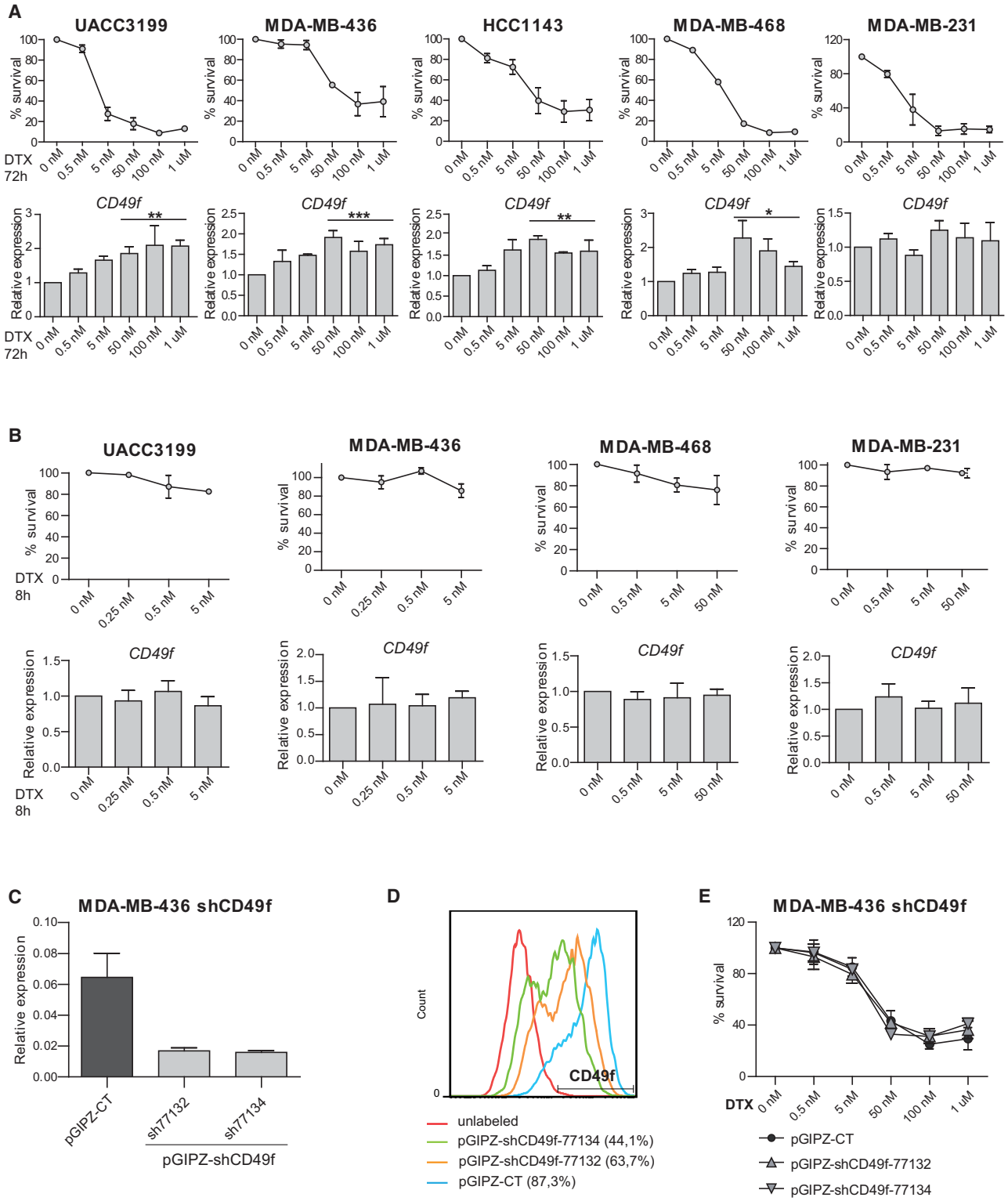


Figure 5. CD49f Expression Increases in Surviving TNBC Cells after Treatment with Docetaxel

(A and B) Top panels: percentage of surviving cells treated with docetaxel for 72 h (A) or 8 h (B). Bottom panels: *CD49f* mRNA expression levels in the indicated TNBC cell lines treated with docetaxel relative to untreated controls. *0.01 < p < 0.05; **0.001 < p < 0.01; ***0.001 < p < 0.0001.

(legend continued on next page)

Indeed, only CD49f+/hi cells, but not the CD49f− in the IDB-01S model, were able to give rise to tumors when re-implanted in mice (Figure 6B). Tumors derived from the IDB-01S-CD49f+/hi cells contained a more abundant CD49f+ population, but also CD49f– cells demonstrating that the tumor-initiating CD49f+/hi cells were able to give rise to non-TICs CD49f− cells (Figure 6C). Docetaxel attenuated growth in tumors derived from IDB-01S-CD49f+/hi cells, but tumors were still palpable after ten doses of docetaxel, in contrast to sensitive tumors of origin IDB-01S that were not detectable after four doses (Figure 6D). Thus, IDB-01S-CD49f+/hi derived tumors are more resistant to docetaxel than the original IDB-01S tumors.

In IDB-02, where tumors were partially sensitive to docetaxel and contained a higher proportion of CD49f+ cells, both CD49f+/hi and CD49f− cells gave rise to tumors. However, limiting dilution assays and extreme limiting dilution analyses (ELDA) revealed that the CD49f+/hi population showed a 5-fold increase in tumor-initiating ability compared with the CD49f− cells (Figure 6E). In addition, the CD49f+/hi cells gave rise to tumors with shorter latency than CD49f− cells (Figure 6F). CD49f+ cells were more abundant in CD49f+/hi than in tumors derived from CD49f−, but tumors from both groups contained CD49f+ and CD49f− cells (Figure 6G), demonstrating that CD49f− cells can also give rise to CD49f+ cells. Again, IDB-02S-CD49f−/lo-derived tumors were more sensitive to docetaxel than the ones derived from IDB-02S-CD49f+ cells (Figure 6H).

Unsupervised gene expression profiling of FACS-sorted CD49f+/hi and CD49f− cells from IDB-01S and -01R, using 105 breast cancer-selected genes, revealed two main clusters which broadly represents the CD49f+ and CD49f− populations (Figures 6I and S6A). Compared with CD49f− cells, CD49f+ cells showed downregulation of keratins, claudins and *CDH3*, and upregulation of *SFRP1*, *MIA* and proliferation-related genes (*UBE2C*, *CDC6* and *CDC20*) (Figure 6I). Further gene expression analyses revealed significant transcriptome differences between CD49f+/hi cells from resistant and sensitive tumors, including enhanced decrease in tight junction proteins, claudins, and *CDH3*, which may suggest a more claudin-low phenotype (Prat et al., 2010). Downregulation of tumor suppressors (e.g., *PTEN* and *RAB25*) is also observed in CD49f+ cells from resistant tumors (Figure S6A). Interestingly, CD49f+/hi cells showed increased proliferation by gene expression analysis than CD49f− cells, especially within sensitive tumors (Figure S6B). Among the two CD49f+ signatures, the IDB-01R/CD49f+ signature was found to predict residual disease following anthracycline/taxane-based therapy in breast tumors (GSE25066), concordant with our preclinical observations (Figure S6C). On the other hand, the IDB-01S/CD49f+ signature was found to predict pathological complete response (pCR) following anthracycline/taxane-based therapy, likely due to the large difference in proliferation between CD49f+ and CD49f− cells in IDB-01S tumors (Figures S6B and S6C) (Hatzis et al., 2011). The IDB-01S/CD49f+ signature was associated with lower recurrence-free survival in an additional dataset of breast cancer patients (Prat et al., 2010) (Figure S6D). Together, these results demonstrate that sensitive tumors of origin contain a tumorigenic and docetaxel-resistant CD49f+ population that changes and expands during the acquisition of taxane resistance; whereas in the absence of the drug, the CD49f+ chemoresistant population shrinks and taxane sensitivity is restored.

DISCUSSION

Patient-derived xenograft (PDX) models have emerged as an important intermediate tool between basic research and clinical trials to expedite the translation of basic research findings into effective therapies for patients. We have generated a panel of PDX models that recapitulates the heterogeneity of human breast tumors. Initial collections of breast PDX were reported to remain phenotypically identical to human tumors during serial passages (DeRose et al., 2011; Zhang et al., 2013). However, in agreement with our findings, there is increasing evidence that tumors in PDX are not "static" and can evolve, as observed in patients (Eirew et al., 2015).

Our PDX models constitute a unique tool to investigate resistance in cancer as they mimic clinical responses: TNBC tumors are more sensitive to chemotherapy than the luminal tumors, confirming previous clinical results (Berry et al., 2006; Colleoni et al., 2004; Guarneri et al., 2006; Martin et al., 2011), and even initially sensitive tumors develop resistance upon continuous exposure to taxanes. Both basal-like tumors (IDB-01 and IDB-02) derived from metastatic samples that were heavily exposed to multiple treatments including taxanes showed minimal clinical response. Strikingly, sensitivity to docetaxel was restored upon xenografting and was retained for months. Moreover, we observed that in PDX tumors with acquired resistance, sensitivity is partially restored when maintained in the

(C and D) *CD49f* mRNA expression levels (C) and CD49f protein expression measured by flow cytometry (D) in cells stably infected with two independent shCD49f knockdown constructs and control vector (pGIPZ).
(E) Percentage of surviving shCD49f-infected and control pGIPZ-infected cells treated with indicated doses of docetaxel for 72 hr.
RT-PCR Determinations were done in triplicate and means are used in the calculations. Mean values of three independent experiments, SEM, and t test p values for the higher concentrations are shown. See also Figure S5.

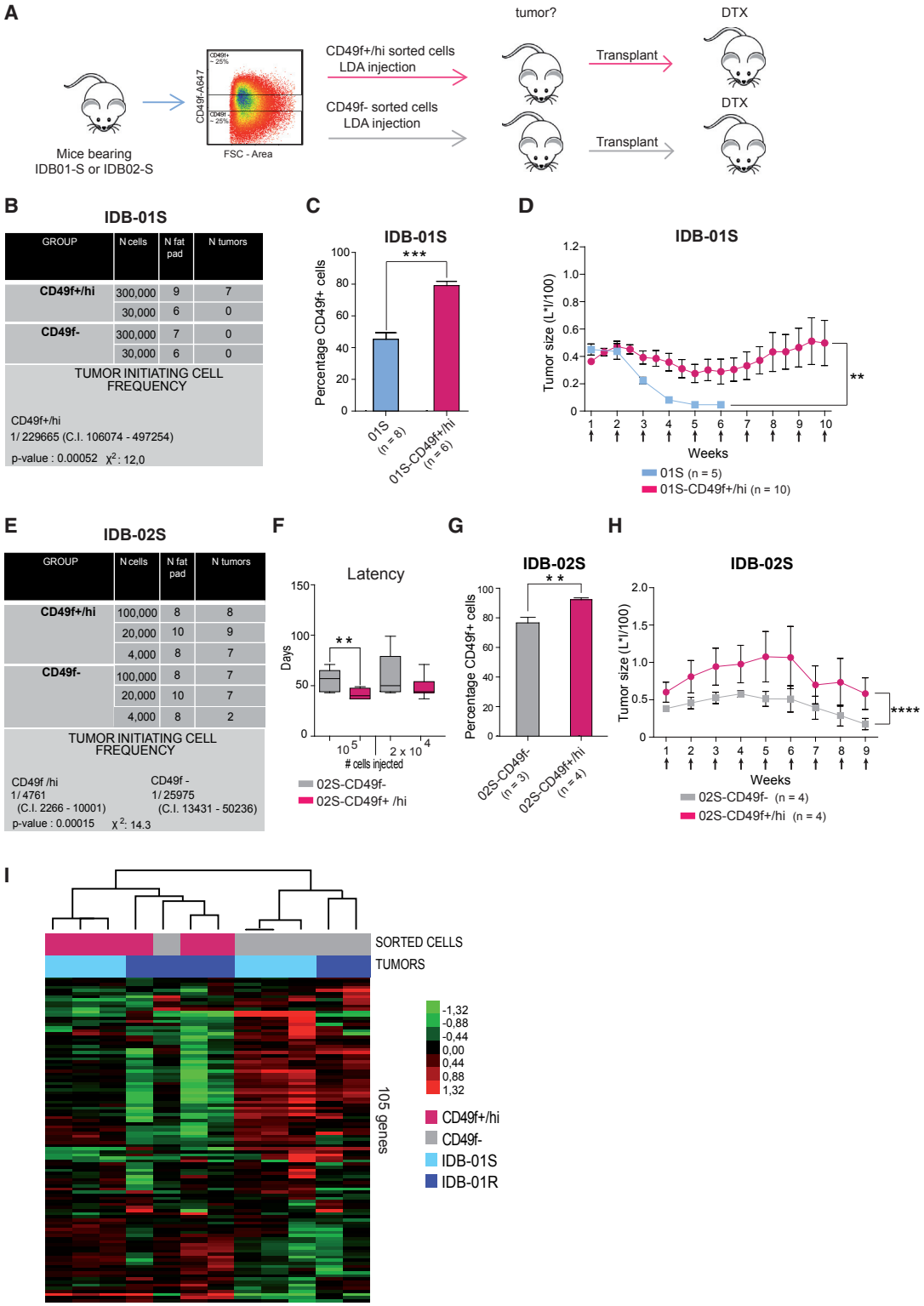


Figure 6. CD49f+ Population Is Enriched in Tumor-Initiating Cells with Increased Resistance to Docetaxel

(A) Scheme of functional experiments.

(B and E) Table showing limiting dilution assay of CD49f+/hi and CD49f− tumor cells from IDB-01 (B) and IDB-02 (E) cells. Tumor-initiating cell frequency (with confidence intervals) for each group was calculated by ELDA; chi-square values and associated probabilities are shown.

(legend continued on next page)

absence of the drug. This regain of sensitivity, the so-called "drug holiday," has been described for targeted therapies in melanoma (Das Thakur et al., 2013; Sun et al., 2014). We now demonstrate that the same is true for cytotoxics such as docetaxel, with important implications for clinical decisions and drug scheduling, as resistant metastatic disease may benefit from intermittent docetaxel treatment.

Our data demonstrate that a pre-existing and chemoresistant CD49f+ subpopulation is present in most sensitive TNBC, expands during long-term therapy, and has the ability to generate novel tumors contributing to recurrence and acquisition of chemoresistance (as shown in the graphical abstract), and importantly that this population shrinks again in the absence of taxanes, restoring drug sensitivity. Previous reports have also shown the increased tumor-initiating ability of CD49f+ cells in breast and other solid tumors (Haraguchi et al., 2013; Lo et al., 2012; Meyer et al., 2010; Vassilopoulos et al., 2014). These findings do not imply that the CD49f+ cells are the CSC in TNBC, but demonstrate that the CD49f+ population is associated with taxane resistance.

Aiming to further characterize the chemoresistant CD49f+ population, an unbiased approach was undertaken. Gene expression analysis revealed important differences, not only between CD49f+ and CD49f− cells, but also between CD49f+ cells from sensitive and resistant tumors. These changes may suggest that the chemoresistant CD49f+ population has expanded during the exposure to docetaxel, and can provide novel therapeutic targets for the metastatic chemoresistant basal-like tumors. Given the heterogeneity of the TNBC subtype, the significant increase in CD49f expression observed in residual or stabilized disease of most TNBC cell lines and PDX models is remarkable and indicates that modulation of CD49f positivity as a biomarker of taxane resistance is not a peculiarity of a single PDX model but a general event in TNBC, which can be exploited for clinical benefit.

These findings can be clinically validated in the neoadjuvant setting, evaluating whether an enrichment of the CD49f population is observed in residual disease following taxane-based chemotherapy. However, as the rates of pCR in TNBC are high (30%–40%), a dynamic study of early changes in the CD49f population after the first cycles of taxane treatment and occurrence of pCR could be a better approach. The clinical utility of the biomarker could be tested in a prospective clinical trial in the neoadjuvant setting where patients are randomized based on the biomarker modulation to change treatment or continue with taxane-based therapy. Improvement of clinical outcomes (pCR rates or survival) should be the final objective. The effect of novel drugs can be evaluated in the subgroup of chemoresistant CD49f-enriched TNBC. In clinical series the presence of CD49f+ in breast cancer is associated with a poor clinical outcome (Friedrichs et al., 1995; Ye et al., 2015). Moreover, within several CSC markers (CD44, CD24, ALDH1A3, and CD49f) analyzed by IHC in breast cancer samples, only CD49f retained prognostic value in a multivariate analyses in ER– disease (Ali et al., 2011). Our results provide a functional rationale for the poor outcome associated with CD49f expression in hormone receptor-negative breast cancer. Further studies will reveal whether this population can be manipulated in order to unveil the ever-elusive status of tumor drug resistance and recurrence.

EXPERIMENTAL PROCEDURES

Patient Characteristics and Generation of PDX

IDB PDX were generated by orthotopic transplantation of primary tumor pieces obtained directly after surgery or cancer cells isolated from pleural effusions and transplanted into the fat pad of immunodeficient mice, as described previously (DeRose et al., 2011). The clinical characteristics from original patient samples, the number and strain of recipient mice, and the outcome of the implant are indicated in Table S1 (IDB-01-05 models). All experimental procedures were performed according to Spanish regulations. Informed consent was obtained from all subjects and the study received approval from the institutional Ethics Committee. Additional models were generated following similar procedures (Bruna et al., 2016). All research involving animals was performed at the IDIBELL animal facility in compliance with protocols approved by the IDIBELL Committee on Animal Care and following national and European Union regulations.

Breast Cancer Cell Isolation, Flow Cytometry, and Sorting

Single cells were isolated from tumors as described previously (Smalley, 2010). Single cells were resuspended and blocked with

(C and G) Frequency of CD49f+ cells in tumors derived from indicated cells. Mean values, SEM, and significant t test p values are shown. **0.001 < p < 0.01; ***0.001 < p < 0.0001.

(D and H) Kinetics of tumor growth during docetaxel treatment in tumors derived from indicated cells. Mean values, SEM, and t test p values are shown. **0.001 < p < 0.01; ****p <0.0001.

(F) Latency of tumors derived from the injection of the indicated number of IDB-02S-CD49f+/hi and 02S-CD49f− tumor cells. Mean values, SEM and significant t test p values are shown. **0.001 < p < 0.01.

(I) Unsupervised analysis of all CD49f sorted samples from IDB-01S and -01R tumors using 105 breast cancer-related genes. The type of sample and tumor are shown below the array tree. Each square represents the relative transcript abundance.

See also Figure S6.

PBS 2% fetal bovine serum (FBS), 2 mM EDTA, and immunoglobulin G blocking reagent for 10 min on ice. Then they were labeled with antibodies against CD24-PE (555428), CD44-APC (559942), EpCAM-FITC (347197), CD10-PECy5 (555376), and CD49f-A647 (562473) (all from BD Pharmingen), CD133/1-PE (130-098-826 from Miltenyi Biotec), and CD49f-APC (FAB13501A from R&D Systems). Mouse cells were excluded in flow cytometry using H2Kd-PECy7 (116622 from BioLegend). Gating was based on "Fluorescence Minus One" controls. Single cells were assessed for their ALDH activity using the ALDEFLUOR assay system (01700 from STEMCELL Technologies) following the manufacturer's procedures. A population of 10,000 living cells was captured in all FACS experiments. FACS analysis and sorting was performed using Gallios and MoFlo (Beckman Coulter) flow cytometers, respectively. Data was analyzed using the FlowJo software (see Figure S4).

Therapeutic and Limiting Dilution Assays

Docetaxel (Hospira/Actavis, 20 mg/kg) was administered intraperitoneally once per week (unless reported otherwise), followed 24 hr later by Fortecortin (Dexametasona, 0.132 mg/kg, Merck). The treatment scheme of resistant variants generation is shown in Figure S3. For orthotopic ELDA, isolated tumor cells were mixed 1:1 with Matrigel Basement Membrane (BD Biosciences) and orthotopically implanted in the inguinal mammary gland of non-obese diabetic/severe combined immunodeficiency females. Tumor development was monitored once per week for a maximum of 25 weeks. In all assays the tumor-initiating potential was defined as the ability to form palpable, growing tumors of ≥ 2 mm diameter.

Culture and Treatment of Human Breast Cancer Cells

All cell lines were purchased from the American Type Culture Collection (Rockville, MD), except for UACC3199 which was obtained from the Arizona Cancer Center (Tucson, AZ). All cells but HCC1143, which was cultured in RPMI 1640, were maintained in DMEM high glucose, containing 10% FBS (Gibco), L-glutamate (Gibco), and penicillin/streptomycin (PAA Laboratories) at 37°C in 5% CO_2. At 60%–70% confluence the indicated concentrations of docetaxel or paclitaxel were added. Cells were collected at the indicated time points and counted with trypan blue to exclude dead ones. All cell lines were routinely tested for mycoplasma, and were shown to be free of contamination.

Gene Expression-Based Analyses

A minimum of ~100 ng of total RNA was used to measure the expression of 105 breast cancer-related genes and five house-keeping genes using the nCounter platform (Nanostring Technologies). Data was log base 2 transformed and normalized using five housekeeping genes (*ACTB, MRPL19, PSMC4, RPLP0* and *SF3A1*). The list of 105 genes includes genes from the following three signatures: PAM50 intrinsic subtype predictor (n = 50) (Parker et al., 2009), claudin-low subtype predictor (n = 43) (Prat et al., 2010), 13-VEGF/hypoxia signature (n = 13) (Hu et al., 2009), and eight individual genes that have been found to play an important role in breast cancer (e.g., CD24). Raw gene expression data and signatures can be found in Table S2. All tumors were assigned to an intrinsic molecular subtype of breast cancer (luminal A, luminal B, HER2-enriched, basal-like, and claudin-low) and the normal-like group using the previously reported PAM50 subtype and the claudin-low subtype predictors (Parker et al., 2009; Prat et al., 2010, 2015b).

Gene Expression-Based Signatures

Genes differentially expressed between the two groups were identified using a two-class unpaired Significance Analysis of Microarrays (SAM) (Tusher et al., 2001) and a false discovery rate of <5%. The final signature of up- and/or downregulated genes was then summarized as a single "enrichment/activity score" by multiplying the SAM score of each gene by its expression value in the tested sample and then summing all the values of each sample. Each signature was evaluated in GSE25066, a microarray-based dataset of patients treated with neoadjuvant anthracycline/taxane-based chemotherapy (Hatzis et al., 2011) and the Perou-extended dataset GSE18229 (Prat et al., 2010). This microarray dataset was normalized as described previously (Prat et al., 2015a). Raw data can be found in Table S2.

Statistical Analyses

All data are expressed as mean ± SEM. Statistical comparison was performed by Student's t test using GraphPad Prism version 5.04. $p \leq 0.05$ was considered statistically significant. The statistical significance of difference between groups is expressed by asterisks: *$0.01 < p < 0.05$; **$0.001 < p < 0.01$; ***$0.001 < p < 0.0001$; ****$p < 0.0001$.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2017.03.026.

AUTHOR CONTRIBUTIONS

J.G.M., M.P., and L.P., collection and assembly of data, data analysis and interpretation, writing, and final approval of manuscript; J.G.M., identified the enrichment in the CD49f+ population and performed functional assays and analyses of residual disease; M.P., generated the PDX models and docetaxel-resistant variants; I.F., S.V., P.P., H.P.M., A.I., G.Y., P.G., A.U., T.S.M., J.C., I.M., M.E., A.N., V.S., A.P., and S.P., collection and assembly of data and final approval of manuscript; P.M., data analysis and interpretation and final approval of manuscript; A.P. and E.G.S., conception and design, financial support, collection and assembly of data, data analysis and interpretation, writing, and final approval of manuscript.

ACKNOWLEDGMENTS

We thank A. Villanueva, C. Saura, C. Cruz, A. Fernández, E. Nadal, M. Martin, and R. Iggo for reagents and helpful advice, A. Welm and Y. DeRose for sharing PDX models, V. Peg, X. Serres, J. Balmaña, J. Pérez, and C. Hierro for providing samples, S. Hernández-Ortega, R. Gil, L. Barberá, and G. Boigues, the Pathology Department of the University Hospital of Bellvitge, the IDIBELL animal facility, histology service and UB-SCT, for technical support, and members of the laboratory for useful discussions and reading of the manuscript. This work

was supported by grants to E. González Suárez by the Spanish Ministry of Economy and Competitivity MINECO and from the Health Institute Carlos III (ISCIII) SAF2008-01975, SAF2011-22893, SAF2014-55997, PIE13/00022, co-funded by FEDER funds/European Regional Development Fund (ERDF – a way to build Europe), by a Career Catalyst Grant from the Susan G Komen Foundation CCR13262449 and by funds to A. Prat from ISCIII-PI13/01718, by a Career Catalyst Grant from the Susan G. Komen Foundation, by Banco Bilbao Vizcaya Argentaria (BBVA) Foundation, and by a Sociedad Española de Oncología Médica (SEOM) grant. The PDX from VHIO were supported by a "GHD-pink" research support via the FERO Foundation to V. Serra. V.S. is recipient of ISCIII grants (PI13/01714 and CP14/0028). M.P., J.G.M., and P.P. were recipients of FPU/FPI grants from the MINECO.

Received: October 13, 2016
Revised: March 31, 2017
Accepted: March 31, 2017
Published: April 27, 2017

REFERENCES

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc. Natl. Acad. Sci. USA *100*, 3983–3988.

Ali, H.R., Dawson, S.J., Blows, F.M., Provenzano, E., Pharoah, P.D., and Caldas, C. (2011). Cancer stem cell markers in breast cancer: pathological, clinical and prognostic significance. Breast Cancer Res. *13*, R118.

Andre, F., and Zielinski, C.C. (2012). Optimal strategies for the treatment of metastatic triple-negative breast cancer with currently approved agents. Ann. Oncol. *23*, vi46–51.

Bachelard-Cascales, E., Chapellier, M., Delay, E., Pochon, G., Voeltzel, T., Puisieux, A., Caron de Fromentel, C., and Maguer-Satta, V. (2008). The CD10 enzyme is a key player to identify and regulate human mammary stem cells. Stem Cells *28*, 1081–1088.

Berry, D.A., Cirrincione, C., Henderson, I.C., Citron, M.L., Budman, D.R., Goldstein, L.J., Martino, S., Perez, E.A., Muss, H.B., Norton, L., et al. (2006). Estrogen-receptor status and outcomes of modern chemotherapy for patients with node-positive breast cancer. JAMA *295*, 1658–1667.

Bonnefoi, H., Piccart, M., Bogaerts, J., Mauriac, L., Fumoleau, P., Brain, E., Petit, T., Rouanet, P., Jassem, J., Blot, E., et al. (2011). TP53 status for prediction of sensitivity to taxane versus non-taxane neoadjuvant chemotherapy in breast cancer (EORTC 10994/BIG 1-00): a randomised phase 3 trial. Lancet Oncol. *12*, 527–539.

Bruna, A., Rueda, O.M., Greenwood, W., Batra, A.S., Callari, M., Batra, R.N., Pogrebniak, K., Sandoval, J., Cassidy, J.W., Tufegdzic-Vidakovic, A., et al. (2016). A biobank of breast cancer explants with preserved intra-tumor heterogeneity to screen anticancer compounds. Cell *167*, 260–274.e22.

Clarke, C., Madden, S.F., Doolan, P., Aherne, S.T., Joyce, H., O'Driscoll, L., Gallagher, W.M., Hennessy, B.T., Moriarty, M., Crown, J., et al. (2013). Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. Carcinogenesis *34*, 2300–2308.

Colleoni, M., Viale, G., Zahrieh, D., Pruneri, G., Gentilini, O., Veronesi, P., Gelber, R.D., Curigliano, G., Torrisi, R., Luini, A., et al. (2004). Chemotherapy is more effective in patients with breast cancer not expressing steroid hormone receptors: a study of preoperative treatment. Clin. Cancer Res. *10*, 6622–6628.

Das Thakur, M., Salangsang, F., Landman, A.S., Sellers, W.R., Pryer, N.K., Levesque, M.P., Dummer, R., McMahon, M., and Stuart, D.D. (2013). Modelling vemurafenib resistance in melanoma reveals a strategy to forestall drug resistance. Nature *494*, 251–255.

Dean, M., Fojo, T., and Bates, S. (2005). Tumour stem cells and drug resistance. Nat. Rev. Cancer *5*, 275–284.

DeRose, Y.S., Wang, G., Lin, Y.C., Bernard, P.S., Buys, S.S., Ebbert, M.T., Factor, R., Matsen, C., Milash, B.A., Nelson, E., et al. (2011). Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nat. Med. *17*, 1514–1520.

Desmedt, C., Di Leo, A., de Azambuja, E., Larsimont, D., Haibe-Kains, B., Selleslags, J., Delaloge, S., Duhem, C., Kains, J.P., Carly, B., et al. (2011). Multifactorial approach to predicting resistance to anthracyclines. J. Clin. Oncol. *29*, 1578–1586.

Dobrolecki, L.E., Airhart, S.D., Alferez, D.G., Aparicio, S., Behbod, F., Bentires-Alj, M., Brisken, C., Bult, C.J., Cai, S., Clarke, R.B., et al. (2016). Patient-derived xenograft (PDX) models in basic and translational breast cancer research. Cancer Metastasis Rev. *35*, 547–573.

Eirew, P., Steif, A., Khattra, J., Ha, G., Yap, D., Farahani, H., Gelmon, K., Chia, S., Mar, C., Wan, A., et al. (2015). Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. Nature *518*, 422–426.

Friedrichs, K., Ruiz, P., Franke, F., Gille, I., Terpe, H.J., and Imhof, B.A. (1995). High expression level of alpha 6 integrin in human breast carcinoma is correlated with reduced survival. Cancer Res. *55*, 901–906.

Guarneri, V., Broglio, K., Kau, S.W., Cristofanilli, M., Buzdar, A.U., Valero, V., Buchholz, T., Meric, F., Middleton, L., Hortobagyi, G.N., and Gonzalez-Angulo, A.M. (2006). Prognostic value of pathologic complete response after primary chemotherapy in relation to hormone receptor status and other factors. J. Clin. Oncol. *24*, 1037–1044.

Haraguchi, N., Ishii, H., Mimori, K., Ohta, K., Uemura, M., Nishimura, J., Hata, T., Takemasa, I., Mizushima, T., Yamamoto, H., et al. (2013). CD49f-positive cell population efficiently enriches colon cancer-initiating cells. Int. J. Oncol. *43*, 425–430.

Hatzis, C., Pusztai, L., Valero, V., Booser, D.J., Esserman, L., Lluch, A., Vidaurre, T., Holmes, F., Souchon, E., Wang, H., et al. (2011). A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. JAMA *305*, 1873–1881.

Hu, Z., Fan, C., Livasy, C., He, X., Oh, D.S., Ewend, M.G., Carey, L.A., Subramanian, S., West, R., Ikpatt, F., et al. (2009). A compact VEGF signature associated with distant metastases and poor outcomes. BMC Med. *7*, 9.

Kim, S.I., Sohn, J., Koo, J.S., Park, S.H., Park, H.S., and Park, B.W. (2010). Molecular subtypes and tumor response to neoadjuvant

chemotherapy in patients with locally advanced breast cancer. Oncology *79*, 324–330.

Li, X., Lewis, M.T., Huang, J., Gutierrez, C., Osborne, C.K., Wu, M.F., Hilsenbeck, S.G., Pavlick, A., Zhang, X., Chamness, G.C., et al. (2008). Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. J. Natl. Cancer Inst. *100*, 672–679.

Lim, E., Vaillant, F., Wu, D., Forrest, N.C., Pal, B., Hart, A.H., Asselin-Labat, M.L., Gyorki, D.E., Ward, T., Partanen, A., et al. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. *15*, 907–913.

Lo, P.K., Kanojia, D., Liu, X., Singh, U.P., Berger, F.G., Wang, Q., and Chen, H. (2012). CD49f and CD61 identify Her2/neu-induced mammary tumor-initiating cells that are potentially derived from luminal progenitors and maintained by the integrin-TGFbeta signaling. Oncogene *31*, 2614–2626.

Martin, M., Romero, A., Cheang, M.C., López García-Asenjo, J.A., García-Saenz, J.A., Oliva, B., Román, J.M., He, X., Casado, A., de la Torre, J., et al. (2011). Genomic predictors of response to doxorubicin versus docetaxel in primary breast cancer. Breast Cancer Res. Treat. *128*, 127–136.

Meyer, M.J., Fleming, J., Lin, A.F., Hussnain, S.A., Ginsburg, E., and Vonderhaar, B.K. (2010). CD44posCD49fhiCD133/2hi defines xenograft-initiating cells in estrogen receptor-negative breast cancer. Cancer Res. *70*, 4624–4633.

Parker, J.S., Mullins, M., Cheang, M.C., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z., et al. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. J. Clin. Oncol. *27*, 1160–1167.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al. (2000). Molecular portraits of human breast tumours. Nature *406*, 747–752.

Prat, A., Parker, J.S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J.I., He, X., and Perou, C.M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res. *12*, R68.

Prat, A., Fan, C., Fernandez, A., Hoadley, K.A., Martinello, R., Vidal, M., Viladot, M., Pineda, E., Arance, A., Munoz, M., et al. (2015a). Response and survival of breast cancer intrinsic subtypes following multi-agent neoadjuvant chemotherapy. BMC Med. *13*, 303.

Prat, A., Pineda, E., Adamo, B., Galvan, P., Fernandez, A., Gaba, L., Diez, M., Viladot, M., Arance, A., and Munoz, M. (2015b). Clinical implications of the intrinsic molecular subtypes of breast cancer. Breast *24* (*Suppl 2*), S26–S35.

Ringner, M., Fredlund, E., Hakkinen, J., Borg, A., and Staaf, J. (2011). GOBO: gene expression-based outcome for breast cancer online. PLoS One *6*, e17911.

Shah, S.P., Roth, A., Goya, R., Oloumi, A., Ha, G., Zhao, Y., Turashvili, G., Ding, J., Tse, K., Haffari, G., et al. (2012). The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature *486*, 395–399.

Smalley, M.J. (2010). Isolation, culture and analysis of mouse mammary epithelial cells. Methods Mol. Biol. *633*, 139–170.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. Nature *439*, 993–997.

Sun, C., Wang, L., Huang, S., Heynen, G.J., Prahallad, A., Robert, C., Haanen, J., Blank, C., Wesseling, J., Willems, S.M., et al. (2014). Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. Nature *508*, 118–122.

Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA *98*, 5116–5121.

Vassilopoulos, A., Chisholm, C., Lahusen, T., Zheng, H., and Deng, C.X. (2014). A critical role of CD29 and CD49f in mediating metastasis for cancer-initiating cells isolated from a Brca1-associated mouse model of breast cancer. Oncogene *33*, 5477–5482.

Ye, F., Qiu, Y., Li, L., Yang, L., Cheng, F., Zhang, H., Wei, B., Zhang, Z., Sun, L., and Bu, H. (2015). The presence of EpCAM(-)/CD49f(+) cells in breast cancer is associated with a poor clinical outcome. J. Breast Cancer *18*, 242–248.

Yu, F., Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J., and Song, E. (2007). let-7 regulates self renewal and tumorigenicity of breast cancer cells. Cell *131*, 1109–1123.

Yu, K.D., Zhu, R., Zhan, M., Rodriguez, A.A., Yang, W., Wong, S., Makris, A., Lehmann, B.D., Chen, X., Mayer, I., et al. (2013). Identification of prognosis-relevant subgroups in patients with chemoresistant triple-negative breast cancer. Clin. Cancer Res. *19*, 2723–2733.

Zhang, X., Claerhout, S., Prat, A., Dobrolecki, L.E., Petrovic, I., Lai, Q., Landis, M.D., Wiechmann, L., Schiff, R., Giuliano, M., et al. (2013). A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. Cancer Res. *73*, 4885–4897.

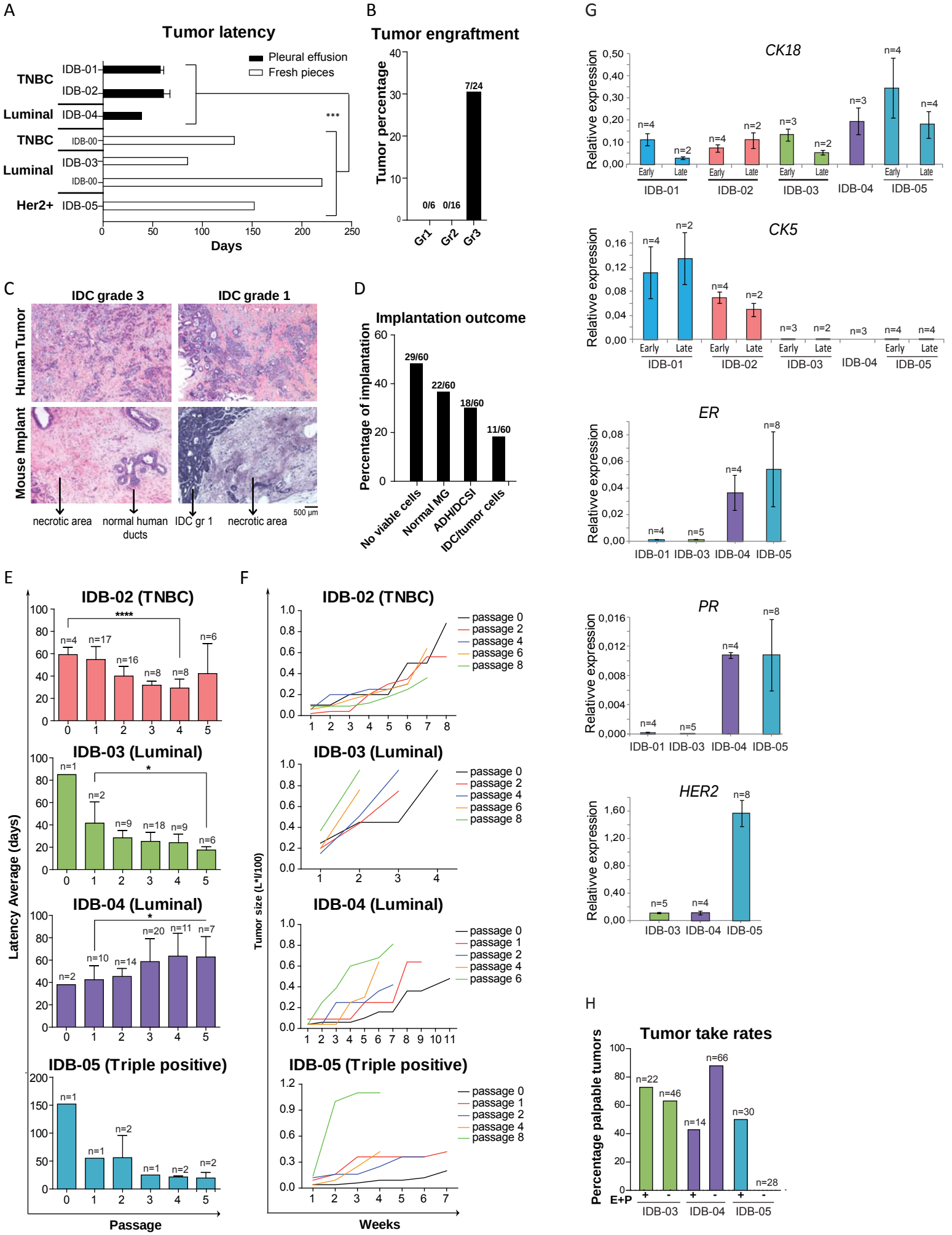
Stem Cell Reports, Volume *8*

Supplemental Information

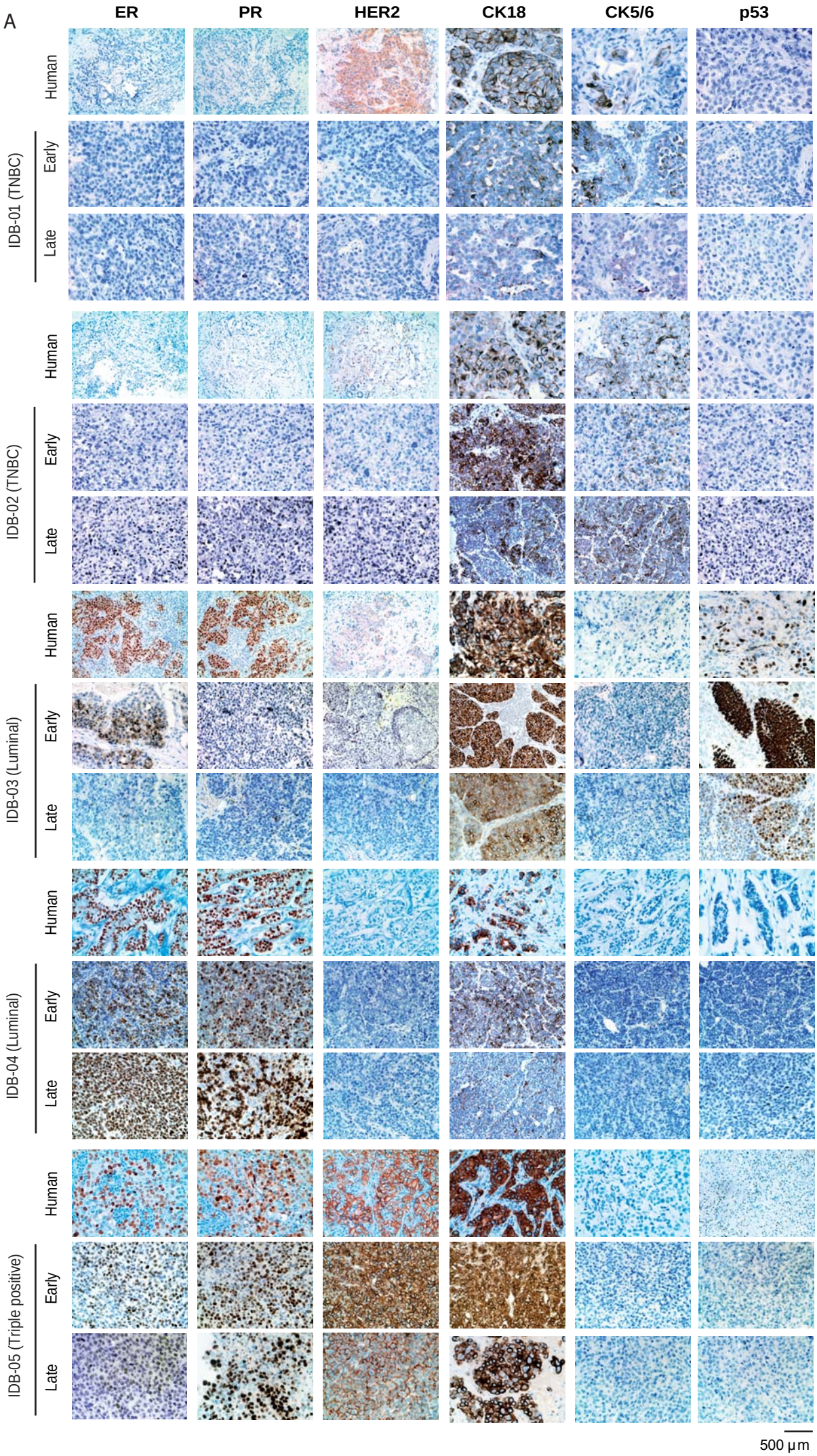
Resistance to Taxanes in Triple-Negative Breast Cancer Associates with the Dynamics of a CD49f+ Tumor-Initiating Population

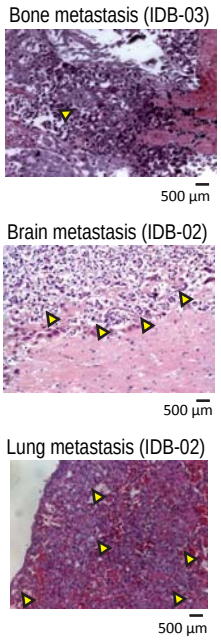
Jorge Gómez-Miragaya, Marta Palafox, Laia Paré, Guillermo Yoldi, Irene Ferrer, Sergi Vila, Patricia Galván, Pasquale Pellegrini, Hector Pérez-Montoyo, Ana Igea, Purificación Muñoz, Manel Esteller, Angel R. Nebreda, Ander Urruticoechea, Idoia Morilla, Sonia Pernas, Fina Climent, María Teresa Soler-Monso, Ana Petit, Violeta Serra, Aleix Prat, and Eva González-Suárez

Supplementary Figure S1

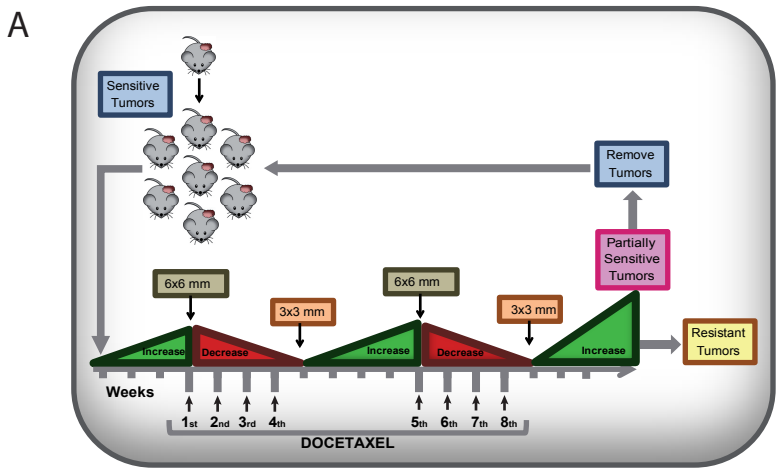


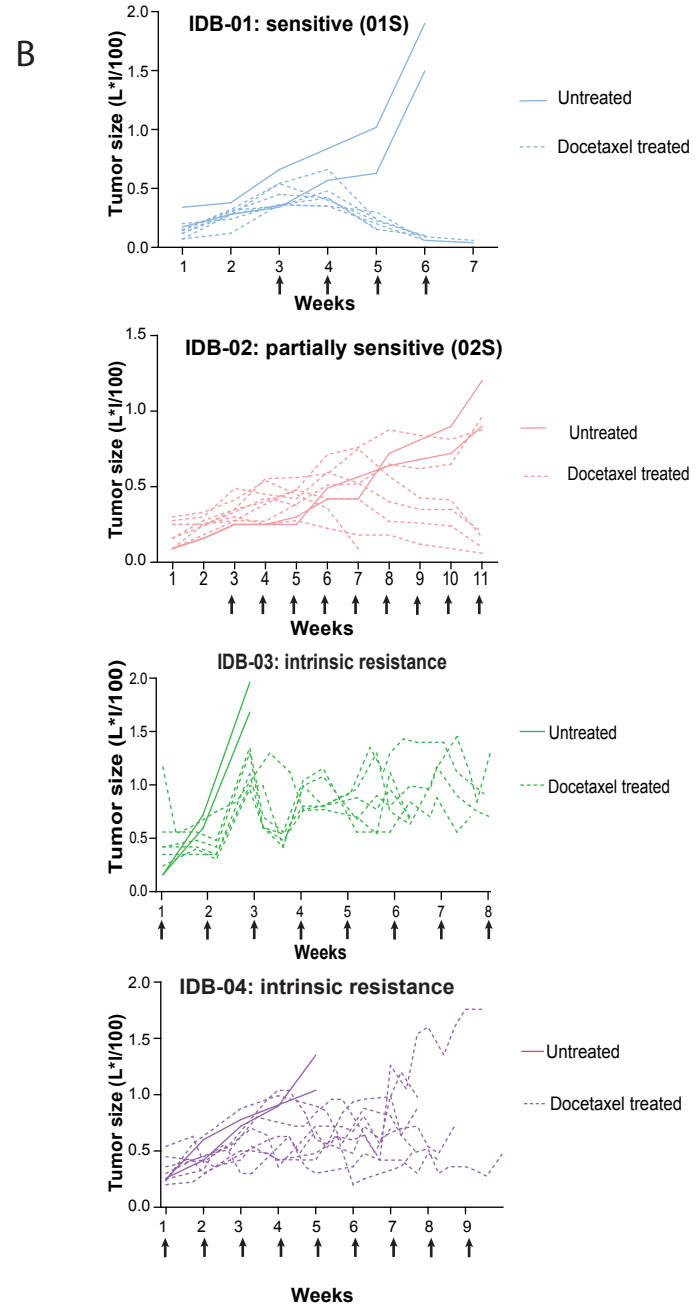
Supplemental Figure S2

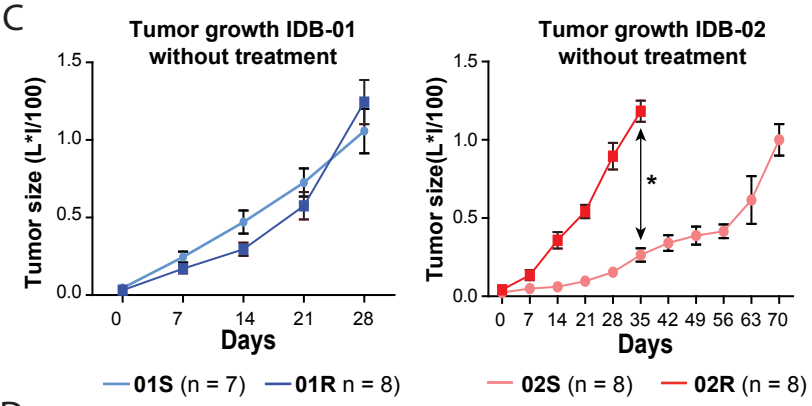


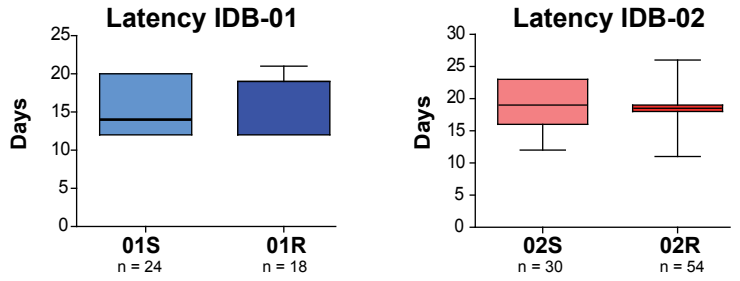


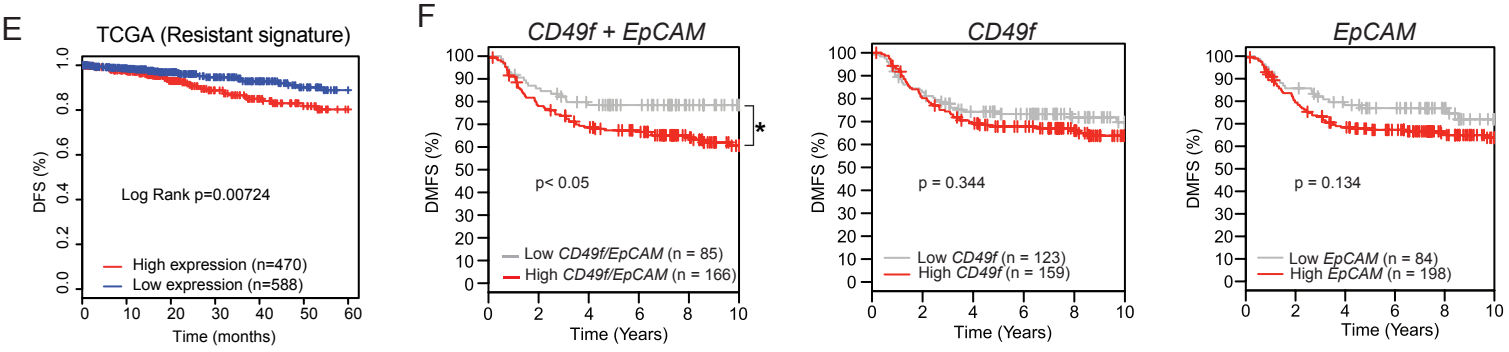
Supplemental Figure S3

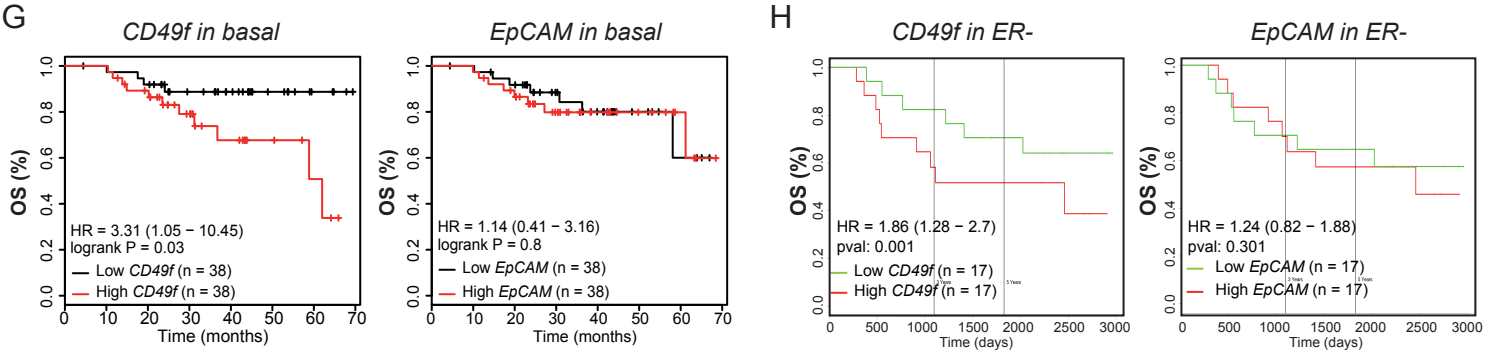






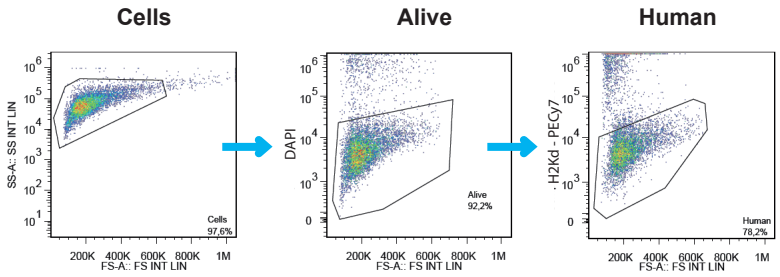




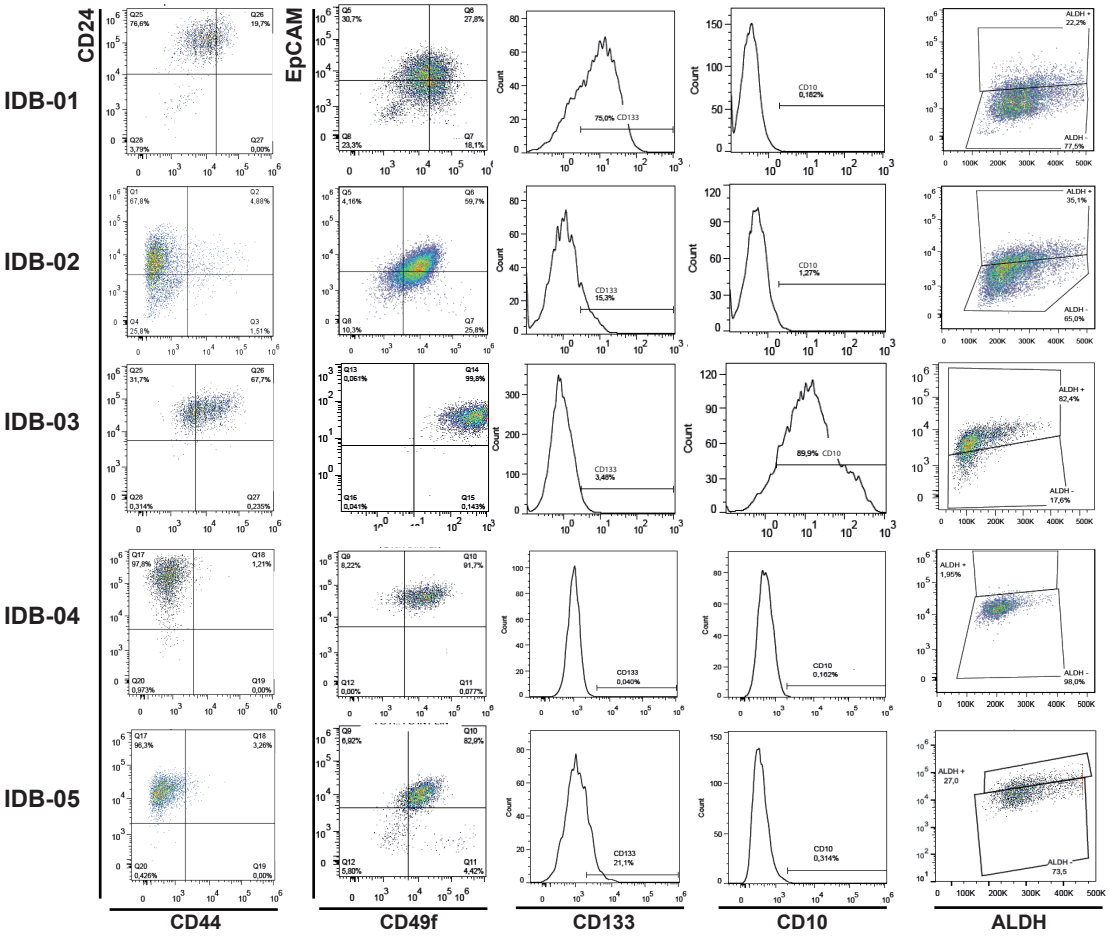


Supplemental Figure S4

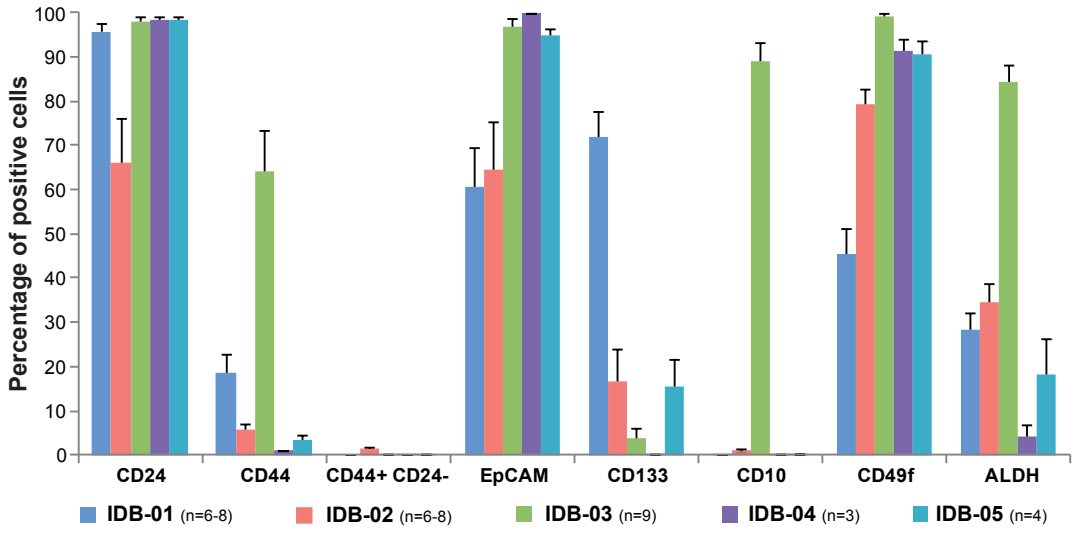
A



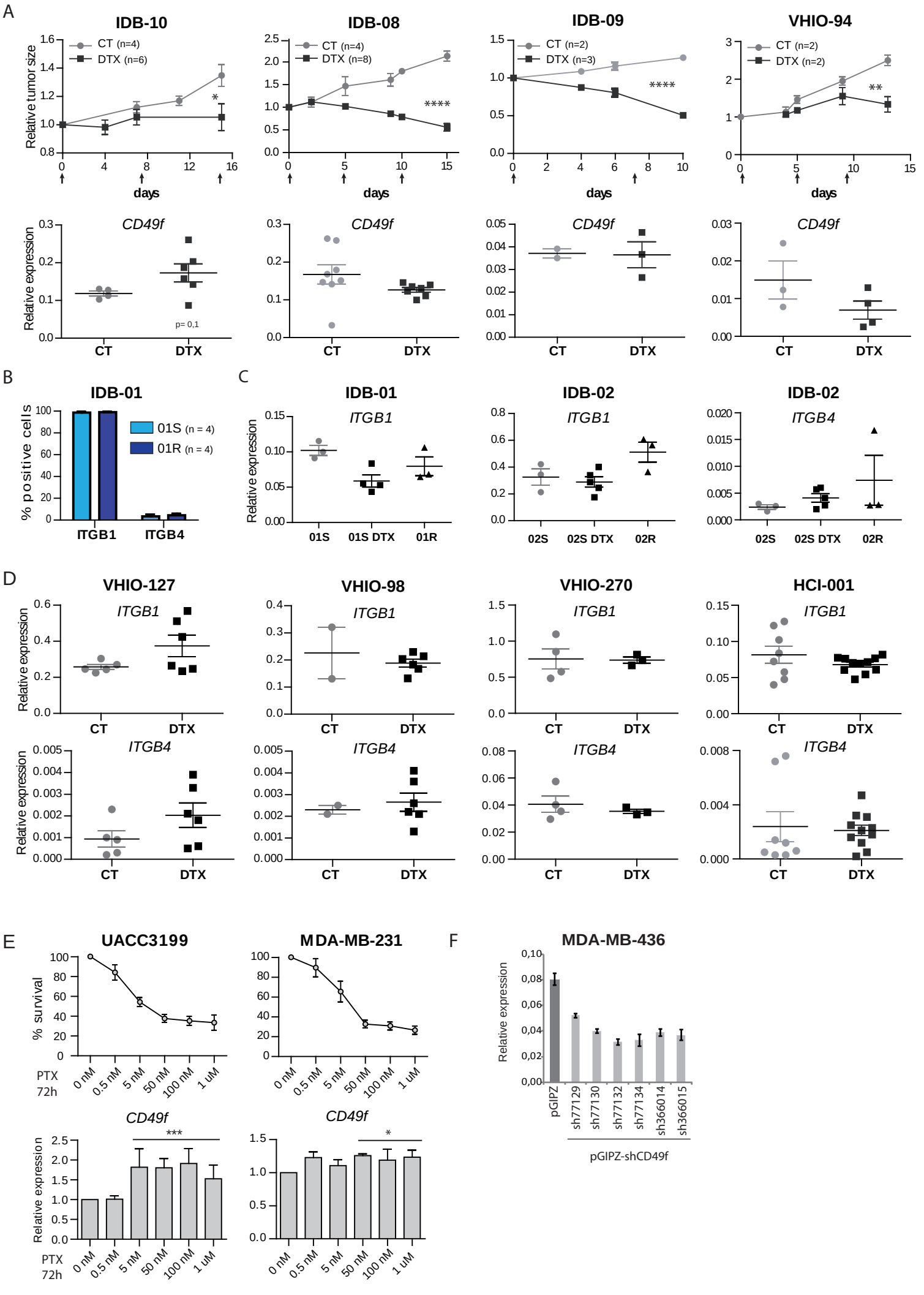
B



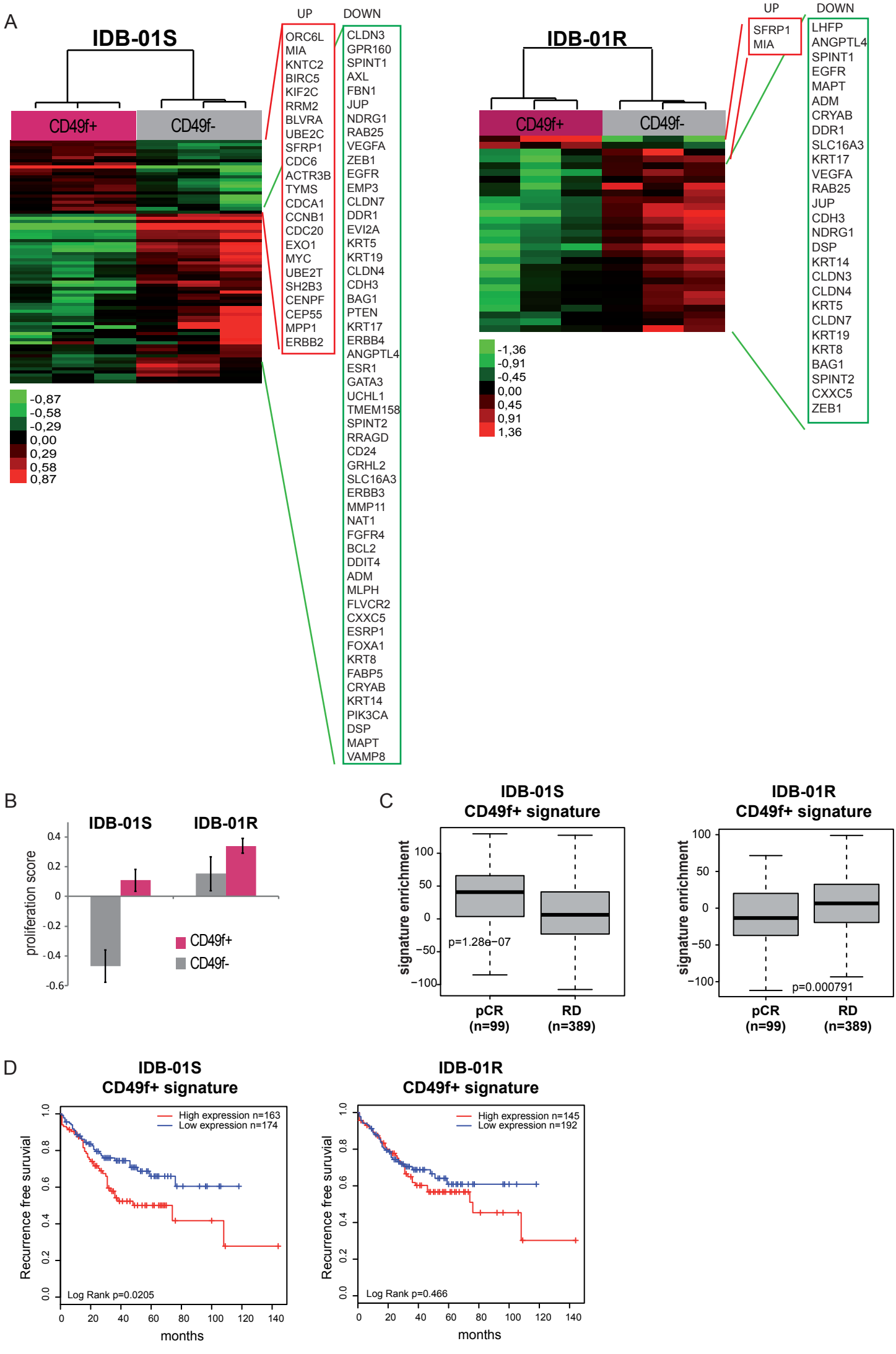
C



Supplementary Figure S5



Supplemental Figure S6



Supplemental Figure legends

Supplemental Figure S1. Generation of breast cancer PDX models (related to Figure 1 and Table 1)
A. Time in days until palpable lesions are detected in mice at first passage according to the source. All engrafted tumors were grade 3 and their subtype is shown. Mean values, SEM and p-values are shown (***, $0.001 < p < 0.0001$).
B. Percentage of palpable tumors engrafted relative to the total number of independent patient samples, depicted in pathological grade classification. Mice that did not survive for at least 60 days after surgery were excluded from the analyses. Total number of samples is indicated (n). Gr1 = grade 1; Gr2 = grade 2; Gr3 = grade 3.
C. Representative images of H&E staining of the original human tumors and the mouse mammary glands where they were implanted. No palpable lesions were detected in these mice and accordingly extensive necrotic areas were observed. However, viable human epithelial structures, mainly normal ducts, ADH (atypical ductal hyperplasia), DCIS (ductal carcinoma in situ) and occasionally IDC (invasive ductal carcinoma) or viable tumor cells were detected. Note that the morphology of IDC grade 1 is maintained in the mouse implant.
D. Percentage of mouse mammary glands showing human structures but not tumor growth. Quantification of indicated human structures in mouse mammary glands where no palpable lesions were detected is represented. MG = mammary glands. Only mice that survived for at least 60 days after tumor implantation were considered. Total number of MG analyzed is indicated (n).
E. Tumor latency (days) in each mouse tumor model at the indicated passages. Total number of tumors considered (n), mean, standard deviation and t-test p values for significant differences are shown (*, $0.01 < p < 0.05$; ****, $p < 0.0001$).
F. Tumor growth in mice, calculated as L x I (mm x mm)/100 versus time (weeks). Each line represents a representative tumor and each color represents a different passage as indicated. Week 1 is the time at which palpable tumors were first detected.
G. mRNA expression levels of indicated genes relative to *PPiA* at indicated passages: early (1^{st} to 5^{th}) and late (more than 10^{th}) measured by qRT-PCR. Total number of tumors analyzed is indicated (n). Determinations were done in triplicate and mean values were used for the calculations. Mean values for "n" independent tumors and standard deviations are shown.
H. Percentage of palpable tumors relative to the total number of tumors implanted with (+) or without (-) exogenous 17β-estradiol and progesterone pellets in luminal models. Total number of implanted mammary glands is indicated (n).

Supplemental Figure S2. Mouse PDX grafts resemble human tumors of origin at early passages but changes are observed at late passages in some models (related to Figure 1 and Table 1)
A. Representative images of estrogen receptor (ER), progesterone receptor (PR), HER2, cytokeratins (CK18, CK5/6) and p53 protein expression in human tumors of origin and tumors growing in mice of the indicated models at early (1^{st} to 2^{nd}) and later passages (4^{th} to 8^{th}) as determined by immunohistochemistry.
B. Representative H&E pictures of macroscopic metastasis.

Supplemental Figure S3. Docetaxel treatment in PDX models and association of CD49f and EpCAM expression with survival (related to Figure 2 and 3)
A. Schematic representation of docetaxel resistant tumors generation. Several tumors from each model were implanted in both mammary inguinal glands of 8 recipient NOD/Scid females. Tumor-bearing animals were individually identified and when tumors reached 6x6 of size, mice were treated with docetaxel (20 mg/kg i.p.) once a week as indicated. If the tumor volume decreases below 3x3 size treatment was interrupted and re-initiated when tumor volume increased again over the size of 6x6. After 10 to 12 doses, mice were sacrificed, tumors were excised and re-implanted in mice (second passage) to repeat the process until tumors become resistant.
B. Representative kinetics of tumor growth during docetaxel treatment in sensitive TNBC IDB-01S and IDB-02S tumors and resistant luminal tumors IDB-03 and IDB-04. Tumor size (v = Lxl/100) is shown over time (weeks). Each line represents one tumor and arrows indicate docetaxel treatments.
C. Tumor growth in the initially sensitive (IDB-01S and IDB-02S) and after resistance to taxanes was acquired (IDB-01R and IDB-02R). Total number of tumors analyzed (n), mean values and SEM are shown. (*, $0.01 < p < 0.05$).

D. Box and whiskers graph (min to max) showing time in days until palpable tumors are detected (tumor latency) in the initially sensitive (IDB-01S and IDB-02S) and after resistance to taxanes was acquired (IDB-01R and IDB-02R). Total number of tumors is shown (n).
E. Kaplan-Meier analyses of disease free survival (DFS) of the resistant signature identified in the TCGA database for breast tumors.
F. Kaplan-Meier analyses of distal metastasis free survival of basal-like tumors using *CD49f* and *EpCAM* mRNA expression independently and combined in the online GOBO database (Ringner et al., 2011).
G. Kaplan-Meier analyses of overall survival of basal-like tumors using *CD49f* and *EpCAM* mRNA expression in the clinical data set (GSE16446) from the TOP TRIAL (Desmedt et al., 2011) .
H. Kaplan-Meier analyses of overall survival of basal-like tumors using *CD49f* and *EpCAM* mRNA expression in the clinical data set (GSE42568) (Clarke et al., 2013) .

Supplemental Figure S4. Heterogeneous expression of CSC markers in PDX models and FACs gating (related to Figure 3)
A. Gating scheme. Analyzed cells were first gated as DAPI negatives (live cells) and H2Kd- (human cells).
B. Representative dot-plots and histograms of indicated markers in H2kD- human cells from each tumor model. The axes were established according to fluorescence minus one (FMO) controls or ALDH negative control.
C. Frequency of indicated markers within the human H2Kd- population analyzed by flow cytometry on established PDX. Total number of tumors analyzed (n), mean values and SEM are shown. Passage 7-14 was analyzed for all models

Supplemental Figure S5. No changes in ITGB1 and ITGB4 in docetaxel resistant tumors or residual disease (related to Figure 4 and 5)
A. Top panels: Tumor size of the indicated PDX tumors treated with docetaxel (20 mg/kg arrows) and corresponding controls relative to the size at the first day of treatment. Total number of tumors (n), mean, SEM and t-test p values are shown (*, $0.01 < p < 0.05$; **, $0.001 < p < 0.01$; ****, $p < 0.0001$). Bottom panels: *CD49f* mRNA expression levels relative to *PPiA* in tumors treated with docetaxel and untreated controls measured by qRT-PCR. Determinations were done in triplicates and means are used.
B. Frequency of ITGB1+ and ITGB4+ cells in H2Kd- human cells in IDB-01S and IDB-01R tumors measured by flow cytometry. Total number of tumors analyzed (n), mean values and SEM are shown.
C. *ITGB1* and *ITGB4* mRNA expression levels relative to *PPiA* in sensitive untreated tumors, residual disease after docetaxel treatment and tumors with acquired resistance in IDB-01 and IDB-02 models measured by qRT-PCR. No expression of *ITGB4* was found in IDB-01 tumors. Determinations were done in triplicates. Means and SEM are shown.
D. *ITGB1* and *ITGB4* mRNA expression levels relative to *PPiA* in tumors treated with docetaxel and in untreated controls of indicated PDX models measured by qRT-PCR. Determinations were done in triplicates.Means and SEM are shown.
E. Top panels. Percentage of surviving cells treated with indicated doses of paclitaxel for 72h. Bottom panels: *CD49f* mRNA expression levels in cells treated with paclitaxel relative to untreated controls. Mean values of 3 independent experiments, SEM and t-test p values for the higher concentrations are shown. (*, $0.01 < p < 0.05$; ***, $0.001 < p < 0.0001$)
F. Bars show *CD49f* mRNA expression levels relative to *PPiA* in the indicated TNBC cell lines stably infected with six independent shCD49f knock-down constructs and control vector (pGIPZ) measured by qRT-PCR. Determinations were done in triplicates and means are used.

Supplemental Figure S6. Molecular characterization of CD49f+ and CD49f- cell populations (related to Figure 6)
A. Supervised expression analysis of the genes found differentially expressed between CD49f+ and CD49f- cells within IDB-01S and -01R tumors. Each square represents the relative transcript abundance.
B. Expression of the PAM50 proliferation score across CD49f+ and CD49f- cells within IDB-01S and -01R tumors. Mean values and SEM are shown
C. Association of IDB-01S-CD49f+ or 01R-CD49f+ signatures with chemotherapy response in 508 patients with breast cancer (GSE25066). Response was measured as pathological complete response or residual disease. Total number of tumors analyzed (n), box and whiskers (min to max) graphs and mean values are shown.

D. Kaplan-Meier analysis of overall survival and distant metastasis free survival using sensitive and resistant CD49f+ signatures in the Perou extended database (Prat et al., 2010).

Supplemental Items

Supplemental Table S1. Characteristics of human tumors of origin and recipient mice (related to Table 1, Figure 1 and 2)

Source, histopathological status and clinical data available of the 61 human samples implanted; number, strain, survival of recipient mice and outcome after H&E examination of recipient mammary glands are shown. Samples excluded from the analysis are shown in red. Samples highlighted in yellow indicate established tumor models. IDC: invasive ductal carcinoma; ADH: atypical ductal hyperplasia; DCIS: ductal carcinoma in situ; QT: chemotherapy; HTP: hormonotherapy, n.d.: no data, n.e. not evaluable. MG: mammary gland.

Supplemental Table S2. Gene expression signatures and raw data (related to Figure 2, Figure 6 and S6)

Supplemental Experimental Procedures

Generation of PDX

A total of 61 samples from breast cancer patients, 54 from fresh primary tumor pieces obtained directly after surgery and 7 from cancer cells isolated from pleural effusions were transplanted into the fat pad of 90 immunodeficient mice. Three strains of immunodeficient mice have been used. Nude mice (Athymic Nude - Foxn1nu Harlan), NOD/ Scid (NOD.CB17-Prkdcscid/J; JAX via Charles River) and Scid/Beige (CB17.Cg-PrkdcscidLystbg-J/Crl, JAX via CR) mice (Carroll and Bosma, 1991). Mice were maintained in specific pathogen-free animal housing (IDIBELL). NOD/Scid and Scid/Beige mice were bred in our animal facility. Fresh primary human breast tumor fragments and cells isolated from pleural effusion were obtained from patients at the time of surgery or thoracocentesis, with informed written patient consent. Fragments of 30 to 60 mm^3 were implanted into cleared fat pad from 4^{th} mammary glands of 3-weeks-old NOD/Scid or Nude females and pellets of 17β-estrogen (0,1 mg) (Innovative Research of America) were implanted into the intraescapular fad pat; in the case of tumor cells isolated from pleural effusions, $3x10^6$ tumor cells were injected in 4th mammary glands of Nude, Scid/Beige or NOD/Scid females. Pathological characteristics of implanted samples, number and strain of host mice and engraftment outcome are shown in Supplementary Table S1. Orthotopic tumors appeared at the graft site 30 to 152 days after implantation and they were subsequently transplanted from mouse to mouse without clearing epithelia. In each passage, samples were collected, cryopreserved in DMSO-fetal bovine serum solution (1:10) as stock, or directly at -80ºC for gene expression analysis, and fixed with phosphate buffered saline (PBS) 10% formol, for histological studies.

Mice that died within 60 days after tumor implantation without developing tumor or samples without pathological data available were excluded from the analyses (15 samples, in red in Table S1). Tumor growth (engraftment) was analyzed in 46 patients´ samples implanted in 71 immunodeficient mice. In the first passage, palpable tumors from 7 patients were obtained; all of them derived from grade 3 tumors. We were able to successfully maintain 5 tumor lines by consecutive rounds of transplantation (yellow in table S1, indicated as IDB-01-IDB-05). To evaluate the influence of hormones ER+PR+ tumors were maintained in mice with or without 17β-estrogen (0,1 mg) progesterone (32,5 mg) pellets (Innovative Research of America) implanted into their intraescapular fad pat.

The additional models included (IDB-08-IDB-10) were generated following similar procedures, after implantation of fresh tumor pieces in the cleared fat pad of NSG (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ, JAX via Charles River) mice. All the *in house* generated PDX models were maintained by serial transplantation in the intact fat pad of Nod/Scid mice, except for IDB-08 which was maintained in NSG mice.

Model HCI001 (TNBC) was donated by A Welm and Y DeRose (DeRose et al., 2011) and was maintained in passage by serial transplantation over time in the fat pad of NSG mice. Models VHIO-93, -94, -98, -102, -127, -197, -270 and -288 (all of them TNBC, VHIO-127 is a BRCA1 mutant) were generated by subcutaneous implantation of primary tumor pieces (except for VHIO-127 and VHIO-288 which derived from metastasis) on the back of Nude mice as described previously (Bruna et al., 2016; Garcia-Garcia et al., 2012). The models from VHIO were collected and implanted in the intact fact pad of Nod/Scid mice in our animal facility to perform docetaxel treatments.

Tumor growth

Every mice was monitored for tumor incidence by palpation and visual inspection and for weight variations. Mice were sacrificed before tumors reached a diameter of 1,5 cm, or when 20% loss of their initial weight or deterioration of health was observed. Individual tumor size was calculated as L*l/100, with "L" being the largest diameter and "l" the smallest. Growth curves were established as a function of time. Tumor latency was recorded for all palpable tumors and mice. To emulate the clinical procedure tumors were surgically removed before they reached 1,5 cm of diameter and mice were left alive to determine the incidence of relapse and metastasis. We randomly sectioned the axillary mammary gland, lymph nodes and lungs of mice that survived longer than 60 days after primary tumor excision.

Histology and immunohistochemistry

Samples from patient or mouse tumors, and mammary fat pads of host mice that did not develope tumors were fixed in formalin immediately after resection and embedded into paraffin. Lungs, brain, liver, kidneys and other organs were collected following the same protocol. Bones were treated with 10% formic acid before were formalin fixation and paraffin embedding. For light microscopic examination 3-µm-thick sections were

stained with hematoxylin and eosin (H&E). Selected lungs from all models and brains from Model B were sectioned every 75-100 µm, stained with H&E and scored for metastasis. For immunohistochemical (IHC) studies 3-µm-thick sections were embedded in paraffin and incubated with antibodies against ER (1:30; DAKO, clon 1D5, IR657), PR (diluted; DAKO, clon 636, IR68), HER2 (1:350; DAKO, SK001), CK18 (diluted; DAKO, clon DC10, IR 618), CK5/6 (1:100; Zymed, clon D5/16B4, MAB1620) and p53 (1:50; Biogenex, AM195). Pre-treatment with citrate pH6 was done on slides stained for ER, PR, HER2, CK18 and p53 and with citrate pH9 on CK5/6 stained slides. Antibodies were detected using biotin-conjugated secondary antibodies (HRP; DAKO). The antigen-antibody complex was conjugated with streptavidin horseradish peroxidase and visualized with diaminobenzidine (Kit DAKO LSAB). Sections were counterstained with hematoxylin and appropriate positive and negative controls were used.

Breast cancer cells isolation

Fresh tissues were mechanically cut using the McIlwain tissue chopper and enzymatically digested with appropriate medium (Dulbecco's modified Eagle's medium [DMEM] F-12 (PAA), 0.3% Collagenase A (Roche Diagnostics, S.L.), 2.5U/mL Dispase (Gibco), 20 mM HEPES (Sigma-Aldrich), and antibiotics) 60 minutes at 37ºC with shacking. Samples were washed with Leibowitz-L15 medium (Gibco) supplemented with 10% fetal bovine serum (FBS) and penicillin/streptomycin between each step. Erythrocytes were eliminated by treating samples with hypotonic lysis buffer (ACK lysing buffer, Lonza Iberica) and incubated over night at 4ºC in the aforementioned Leibowitz-L15 medium. The following day, single epithelial cells were isolated by treating with trypsin (PAA Laboratories) 5 minutes at 37ºC and a mix of Dispase (Gibco life technologies, Invitrogen) DNAse (Invitrogen) for 10 minutes at 37ºC. Dead cells were first excluded by centrifugation with Lympholyte (Cedarlane laboratories) 800xg for 20 minutes and then 250xg for 10 minutes at room temperature. Cell aggregates were removed by filtering cell suspension with 40 µm filter and counted with Trypan Blue.

ALDH assay

Single cells were assessed for their ALDH activity using the ALDEFLUOR assay system (STEMCELL technologies). 4 x 10^5 cells were re-suspended in ALDEFLUOR buffer and activated ALDEFLUOR substrate was added. Immediately, half of them were separated in other tube with the inhibitor DEAB. The incubation was performed during 30 minutes at 37ºC. Mouse cells were excluded in flow cytometry using H2Kd-PECy7 (116622 from BioLegend). A population of 10,000 living cells was captured in all fluorescence-activated cell sorting (FACS) experiments. FACS analysis was performed using FACS Gallios cytometer (Beckman Coulter, Inc) and the FlowJo software package.

Therapeutic assays

For the generation of the docetaxel resistant-derived tumors, treatment started in mice bearing tumors of 6x6 mm size (L*l/100). Tumor growth and weight were evaluated twice a week. When tumor volume decreased below 3x3 mm, treatment was interrupted, and reinitiated when tumors re-grew over of 6x6 mm. Mice were ethically sacrificed after 10 to 12 doses of docetaxel when the tumor size surpassed a diameter of 1,5 cm or mouse weight decreased by 20%. Tumors were then excised, cut in pieces and re-implanted into new host mice (passage 2) in which docetaxel treatment was reinitiated following the same criteria, as shown in Supplementary Figure S3A.

For short term treatments, tumor bearing mice were treated with 2 to 5 doses of docetaxel, every 5-7 days for a period of 10 to 30 days, depending on tumors response. Mice were then sacrificed and tumors were analyzed for CD49f expression 3-5 days after the last dose of docetaxel, except for the most sensitive models that needed longer time to grow.

Lentiviral infection

Lentiviral infection was done following the manufacturer´s indications (Invitrogen). Briefly 293FT cells were used for the production of the virus. 293FT cells (5x106) were transfected with lentiviral pGIPZ empty or pGIPZ-shCD49f vectors (Dharmacon GE) and packaging (gag-pol, vsvg, rev) plasmids (Addgene) by calcium phosphate method. 25mM HEPES was added 16 h later. Virus supernatants were harvested 72h post transfection, centrigufated at 250G 5' and filtered with 0.22 µm filters. MDA-MB-436 cell lines were transduced in a ratio 1:3 with fresh growth medium and with 8µg/ml of polybrene. Plates were centrifuged 1 hour at 1.000 rpm at 37ºC to improve the infection. Selection started with puromycin antibiotic (Sigma-

Aldrich) at 1,5µg/ml. The resulting stable cell lines infected were maintained with 0,5 µg/ml . Medium was refreshed every three days. The following shCD49f sequences were tested: 77129: TATTCCATCTGCCTTGCTG; 77130:TAGTTACTGAATCTGAGAG; 77132: TTCTGAATATTAATCACAG; 77134: TTAGAAACAATACCTTTCC; 326014: ATTTCTAAAGCAATATCCT and 326015: TCAGTTGTACTTAAAACCA, and based on the results 77132 and 77134 were selected.

RNA extraction and RT-PCR

Total RNA from tissue was prepared with Tripure Isolation Reagent (Roche). Frozen tumor tissues were fractionated using the POLYTRON® system PT 1200 E (Kinematica). cDNA was produced by reverse transcription using 1 µg of RNA in a 35 µL reaction following manufacturer's instructions (Applied Biosystems). 20 ng/well of cDNA were used for the analysis performed in triplicate. Quantitative PCR was performed using the LightCycler® 480 SYBR green. Primer sequences are indicated below. Ct analysis was performed using LightCycler 480 software (Roche). All primers indicated below are in 5' → 3' direction.

hCD49f Forward	CTGGCCTCTTCATTTGGCTA
hCD49f Reverse	AAAATACTGTGGGGCTCCAAT
hEpCAM Forward	AATCGTCAATGCCAGTGTACTT
hEpCAM Reverse	TCTCATCGCAGTCAGGATCATAA
hPPiA Forward	ATGCTGGACCCAACACAAAT
hPPiA Reverse	TCTTTCACTTTGCCAAACACC
hCK18 Forward	TCAGCAGATTGAGGAGAGCA
hCK18 Reverse	GAGCTGCTCCATCTGTAGG
hCK5 Forward	ATCGCCACTTACCGCAAGCTGCTGGAG
hCK5 Reverse	AAACACTGCTTGTGACAACAGAG
hER Forward	ATCTCGGTTCCGCATGATGAATCTGC
hER Reverse	TGCTGGACAGAAATGTGTACACTCCAGA
hPR Forward	GGCATGGTCCTTGGAGGT
hPR Reverse	CACTGGCTGTGGGAGAGC
hHER2 Forward	TTCCTTCCTGCTTGAGTTCC
hHER2 Reverse	GRGCTGTTCCTCTTCCAACG
hITGB1 Forward	GCCGCGCGGAAAAGATG
hITGB1 Reverse	ACAATTTGGCCCTGCTTGTA
hITGB4 Forward	CCCCGAGGTAGGTCCAGG
hITGB4 Reverse	GTTTGCCAAGGTCCCAGAGA

Public clinical tools:

The web-based tools used include: Gene expression-based Outcome for Breast cancer Online (GOBO), comprising 1881 breast cancer-samples (Ringner et al., 2011) http://co.bmc.lu.se/gobo/gsa_information.pl, Kaplan Meier plotter (KM plotter), capable to assess the effect of 54,675 genes on survival using 5,143 breast cancer patients (Szasz et al., 2016) http://kmplot.com/analysis/, and PROGgeneV2, a tool that can be used to study prognostic implications of genes in various cancers, including breast (Goswami and Nakshatri, 2014) http://watson.compbio.iupui.edu/chirayu/proggene/database/index.php

Supplemental References

Bruna, A., Rueda, O. M., Greenwood, W., Batra, A. S., Callari, M., Batra, R. N., Pogrebniak, K., Sandoval, J., Cassidy, J. W., Tufegdzic-Vidakovic, A.*, et al.* (2016). A Biobank of Breast Cancer Explants with Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds. Cell *167*, 260-274 e222.

Carroll, A. M., and Bosma, M. J. (1991). T-lymphocyte development in scid mice is arrested shortly after the initiation of T-cell receptor delta gene recombination. Genes Dev *5*, 1357-1366.

Clarke, C., Madden, S. F., Doolan, P., Aherne, S. T., Joyce, H., O'Driscoll, L., Gallagher, W. M., Hennessy, B. T., Moriarty, M., Crown, J.*, et al.* (2013). Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. Carcinogenesis *34*, 2300-2308.

DeRose, Y. S., Wang, G., Lin, Y. C., Bernard, P. S., Buys, S. S., Ebbert, M. T., Factor, R., Matsen, C., Milash, B. A., Nelson, E.*, et al.* (2011). Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nat Med *17*, 1514-1520.

Desmedt, C., Di Leo, A., de Azambuja, E., Larsimont, D., Haibe-Kains, B., Selleslags, J., Delaloge, S., Duhem, C., Kains, J. P., Carly, B.*, et al.* (2011). Multifactorial approach to predicting resistance to anthracyclines. J Clin Oncol *29*, 1578-1586.

Garcia-Garcia, C., Ibrahim, Y. H., Serra, V., Calvo, M. T., Guzman, M., Grueso, J., Aura, C., Perez, J., Jessen, K., Liu, Y.*, et al.* (2012). Dual mTORC1/2 and HER2 blockade results in antitumor activity in preclinical models of breast cancer resistant to anti-HER2 therapy. Clin Cancer Res *18*, 2603-2612.

Goswami, C. P., and Nakshatri, H. (2014). PROGgeneV2: enhancements on the existing database. BMC Cancer *14*, 970.

Prat, A., Parker, J. S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J. I., He, X., and Perou, C. M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res *12*, R68.

Ringner, M., Fredlund, E., Hakkinen, J., Borg, A., and Staaf, J. (2011). GOBO: gene expression-based outcome for breast cancer online. PLoS One *6*, e17911.

Szasz, A. M., Lanczky, A., Nagy, A., Forster, S., Hark, K., Green, J. E., Boussioutas, A., Busuttil, R., Szabo, A., and Gyorffy, B. (2016). Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1,065 patients. Oncotarget *7*, 49322-49333.