Europe PMC Funders Group
Author Manuscript
Nat Cell Biol. Author manuscript; available in PMC 2021 April 27.

Published in final edited form as:
Nat Cell Biol. 2020 March 01; 22(3): 289–296. doi:10.1038/s41556-020-0474-3.

Cross-talk with lung epithelial cells regulates *Sfrp2*-mediated latency in breast cancer dissemination

Marco Montagner[1,2,*], **Rahul Bhome**[1], **Steven Hooper**[1], **Probir Chakravarty**[3], **Xiao Qin**[4], **Jahangir Sufi**[4], **Ajay Bhargava**[1], **Colin D. H. Ratcliffe**[1], **Yutaka Naito**[1], **Arianna Pocaterra**[2], **Chris Tape**[4], **Erik Sahai**[1,*]

[1]Tumour Cell Biology Laboratory, Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK

[2]Department. of Molecular Medicine, University of Padua, Viale G. Colombo 3, 35126 Padova, Italy

[3]Bioinformatics Platform, Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK

[4]Cell Communication Lab, Department of Oncology, University College London Cancer Institute, 72 Huntley Street, London, WC1E 6DD, UK

Abstract

The process of metastasis is complex[1]. In breast cancer, there are frequently long timespans between cells leaving the primary tumour and growth of overt metastases [2,3]. Reasons for disease indolence and subsequent transitioning back to aggressive growth include interplay with myeloid and fibroblastic cells in the tumour microenvironment and ongoing immune surveillance [4–6]. However, the signals causing actively growing cells to enter into an indolent state, and enabling them to survive for extended periods of time, are not well understood. We reveal how the behaviour of indolent breast cancer cells in the lung is determined by their interactions with alveolar epithelial cells, in particular AT1 cells. This promotes the formation of fibronectin (FN) fibrils by indolent cells that drive integrin-dependent pro-survival signals. Combined *in vivo* RNA sequencing and drop-out screening identified Secreted frizzled-related protein 2 (*Sfrp2*) as a key mediator of this interaction. *Sfrp2* is induced in breast cancer cells by signals emanating from lung epithelial cells and promotes FN fibril formation and survival, while blockade of *Sfrp2* expression reduces the burden of indolent disease.

To analyse indolent breast cancer, we utilised the D2.0R/D2.A1 model[7,8] (Extended Data Fig. 1a). D2.0R cells persisted for many weeks in the lungs (Fig. 1a and Extended Data Fig.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

*correspondence to marco.montagner@unipd.it and erik.sahai@crick.ac.uk.

Author contributions

M. M. and E. S. conceived, designed, and wrote the study. M. M. performed all the experiments with the exception of the CyTOF analysis, which was performed by R. B. with assistance from X. Q. and J. S. and supervision from C. T., some of the *in vitro* co-cultures, which were performed by S. H. and E. S., and the *in vivo* analysis of the proliferation in the lungs and sFRP2 over-expression, which were performed by S. H. with assistance from A. B., Y. N. and E. S.. C. D. H. R. and A. P. assisted with cell culture and analysis of gene expression. P. C. performed the bioinformatics analysis.

Competing interests

The authors declare no competing interests.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

1b), but did not form large colonies, whereas D2.A1 cells grew aggressively (Extended Data Fig. 1b). The indolent behaviour of the D2.0R cells parallels that of ESR1 +ve breast cancer; consistent with this, D2.0R cells express ESR1 *in vivo* and respond to estradiol [9].D2.0R cells were indolent in both Balb/C and Balb/C nude mice indicating that their phenotype is not due to the adaptive immune system (Extended Data Fig. 1c). Closer examination revealed that D2.0R cells had extravasated into the alveolar space and contacted the lung parenchyma after two days Fig. 1a. In this context, both D2.0R and D2.A1 cells formed long extensions reminiscent of filopodia-like protrusions (Fig. 1b and Extended Data Fig. 1d)[10, 11]. These protrusions increased the cell perimeter relative to the cell area, leading to a low circularity index of ~0.4 (1 = perfect circle). Immunostaining demonstrated that D2.0R had close contact with AQP5+ve and PDPN+ve alveolar type I (AT1) cells (Fig. 1c, left). Frequent contacts were also observed with SFTPC+ve alveolar type II (AT2) cells and MUC1+ve endothelial cells[12] (Extended Data Fig. 1e, f). EdU pulse labelling revealed that AT1 cells, which are normally quiescent, were proliferating proximal to D2.0R cells at 3 days and 14 days after arrival in the lungs. This suggests that the expansion of the lung parenchyma around indolent metastases at two weeks results mostly from proliferation of AT1 cells (Fig. 1a,c and Extended Data Fig. 1g, h). Similar contacts with PDPN+ve alveolar type I cells and an increase in EdU positivity were observed with human MCF7 cells (Figures 1d and Extended Data Fig. 1i).

To study how breast cancer cells might interact in the lung environment, we established a co-culture system that recapitulated the indolent behaviour of D2.0R cells. We co-cultured lung cells that express the markers of AT1 and AT2 cells and fibroblasts on a gas permeable substrate in Mitogen Low Glucose Low (MLNL medium)[13–15] (Fig. 1e). The addition of low numbers of either D2.0R or D2.A1 cells to these co-cultures recapitulated the indolent and aggressive growth of D2.0R and D2.A1 cells observed *in vivo* (Figures 1f, g and increased Ki67+ cells shown in Extended Data Fig. 2b). These differences could not be attributed to intrinsic differences in growth between D2.0R and D2.A1 cells in MHNH medium (Mitogen High Glucose High) or MLNL medium (Extended Data Fig. 2c). The indolent behaviour of D2.0R cells in the co-culture was reversible if cells were subsequently returned to conventional cell culture conditions (Extended Data Fig. 2d). We next explored the effect of individual cell types within the co-culture assay, something that is not possible in mice. Co-cultures with individual cell types in MLNL media indicated that AT1-like cells were able to boost D2.0R numbers, with AT2-like cells having a smaller positive effect (Fig. 1h). Similar results were obtained with 4T07 cells and MCF7 cells (Extended Data Fig. 3a). Time-lapse imaging revealed that AT1-like cells both suppress apoptosis and increase the mitotic rate of D2.0R cells (Extended Data Fig. 3b). The omission of individual epithelial cells types from the ‘full’ co-culture revealed a more nuanced picture of the interplay between breast cancer cells and AT1-like cells (Extended Data Fig. 3c). Increased growth was observed in the absence of AT1-like cells, suggesting that as well as generating pro-survival signals in the more restrictive MLNL conditions (Fig. 1h), they can also generate growth suppressive cues that counteract proliferative cues, most likely emanating from the AT2-like cells. To test directly whether AT1-like cells could suppress growth in the face of strong proliferative cues, we cultured D2.0R cells in MHNH media for 7 days. Extended Data Fig. 3d shows that AT1-like cells were able to reduce the growth of D2.0R cells in

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

favourable conditions. These data suggest that both pro-survival and growth restrictive signals from AT1 cells likely coexist *in vivo* and *in vitro.* We reasoned that the greatest clinical benefit would result from being able to target the supportive signals, therefore we concentrated on the interplay between lung epithelial cells and breast cancer cells in MLNL media.

In culture conditions where AT1-like cells support D2.0R cells (MLNL media), long protrusions were observed similar to those observed *in vivo* (Figures 1i, j). This change reduced the circularity of D2.0R cells from ~0.8 to 0.2-0.4 (Extended Data Fig. 3e); furthermore, the protrusions were positive for active p-Src, which is implicated in pro-metastatic signals[8,10,16](Fig. 1k). The protrusions formed by D2.0R cells in the presence of AT1-like cells were associated with fibronectin fibrils (Fig. 1i). Similar increases in cell protrusion were obtained with 4T07, MCF7, and T47D-DBM cells (Extended Data Figs. 3f-h), but not for the aggressive cell line D2.A1 that had a higher baseline of protrusions (Extended Data Fig. 3i). Blockade of integrins using cilengitide, which mimics the RGD integrin binding motif of fibronectin and other ECM molecules, reduced both numbers and protrusions of D2.0R and MCF7 cells (Figures 1l, m and Extended Data Fig. 3j). Most importantly, treatment of mice with cilengitide even after cells had already seeded the lungs reduced the number of metastases (Fig. 1n). These data suggest that persistence of indolent breast cancer cells and the induction of cellular protrusions by AT1 cells represent intertwined aspects of metastatic dissemination. Furthermore, they demonstrate that targeting this axis is a viable strategy for the elimination of indolent breast cancer cells.

To understand the signalling pathways involved in breast cancer-alveolar cell crosstalk, we undertook mass cytometry analysis of co-cultures (dExtended Data Fig. 4a). Consistent with data in Extended Data Fig. 3b, there was an increase in phosphorylation events associated with proliferation (S807/811-pRb and T37/46-p4E-BP1) in D2.0R and MCF7 upon co-culture (Fig. 2a). Changes in signalling were also observed, with prominent increases in ERK, MKK4, MKK3/6, PDPK1, β-catenin, and NFkB activity (Fig. 2a). Conversely, both D2.0R and MCF7 cells triggered proliferative responses in AT1-like cells (S–phase markers: S807/811-pRb and T37/46-p4E-BP1 and mitotic marker: pS28-H3), supporting the *in vivo* observation of EdU+ve nuclei around micrometastases (Extended Data Fig. 2b, c). Density Resampled Estimation of Mutual Information (DREMI) analysis revealed increased connectivity from PDPK1 to PKCα and AKT and from AKT to 4E-BP1 (Extended Data Fig. 2d – DREMI score in white). Several of these pathways have been linked to cancer dormancy[3,17–19], we therefore investigated how inhibition of these and other prominent signalling pathways affected D2.0R behaviour in presence of AT1-like cells. Blockade of EGFR, MEK, JNK and Src-family kinase (SFK) signalling, but not p38MAPK or β-catenin signalling, reduced the number of D2.0R cells when co-cultured with AT1-like cells (Fig. 2c). EGFR, MEK, and SFK inhibition both increased apoptosis and reduced mitotic events without greatly affecting D2.0R and AT1-like cells in monoculture (Fig. 2b and Extended Data Fig. 2e). Combining inhibitors with phospho-ERK analysis indicated that EGFRi, SFKi, and MEKi all reduced pERK levels, supporting a role for EGFR and SFK signalling upstream of ERK/MAP kinase (Extended Data Fig. 2f). The importance of ERK/MAP kinase activation was confirmed by a reduction in metastatic colony size *in vivo* (Extended Data Fig. 2g). The formation of cell protrusions and FN fibrils described in Fig. 1 depends

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

on EGFR and SFK signalling, but not on MEK signalling, indicating a bifurcation in the signalling cascade at a point downstream of SFK (Fig. 2c, Extended Data Figs 5a-e). The reduced FN staining was not correlated with reduced FN transcription (Extended Data Fig. 2j). These data reinforce the correlation between cell protrusions and signals that boost survival of indolent breast cancer cells.

In a parallel effort to understand indolent breast cancer cells, we investigated how the metastatic microenvironment alters cancer cell gene expression *in vivo*. We compared gene expression in indolent D2.0R cells isolated from the lungs with aggressively growing D2.A1 cells and both cell types grown in culture (Fig. 3a). As expected, the expression of cell cycle and DNA replication genes was reduced in indolent D2.0R cells and we observed up-regulation of BMP signalling[20,21] and the dormancy–associated factors *NR2F1* and *BHLHE41* (also known as *DEC2/SHARP1*)[18,22–24] (Fig. 3b and Extended Data Fig. 6a,b). We also noted an increase in extra-cellular matrix (ECM) genes, including several linked to aggressive metastatic behaviour (*POSTN, TNC*) and epithelial-to-mesenchymal transition (EMT) factors[25,26] (validated by QRT-PCR analysis - Fig. 3c). We next explored links between the transcriptome of indolent D2.0R cells and human breast cancer. A signature of genes highly expressed in indolent D2.0R cells *in vivo* compared to the other groups was clearly linked with improved distant metastasis free survival (DMFS) in human ER+ve breast cancer, which shows long latency periods before relapse[27] (Fig. 3d – genes listed in Supplementary Table 1 and compared with other 'dormancy signatures'[28,29] in Extended Data Fig. 7a,b). The signature remained associated with favourable outcome in ER+ve lymph node negative disease (p=0.00038), indicating that it is not simply a metric of lymph node status. Patients with high expression of the signature receiving tamoxifen therapy had increased DMFS at 10 years (confirmed using the GSE9515 dataset, Extended Data Fig. 7c). Conversely, patients with low expression of the 'indolence' signature had a significantly increased hazard ratio of 2.5 (Fig. 3e). Furthermore, re-plotting the analysis from two years onwards confirmed our signature's ability to indicate lower likelihood of distant relapses at prolonged time points (Extended Data Fig. 7c). Genes specifically up–regulated in D2.A1 cells or on plastic showed no link with outcomes (Extended Data Fig. 7d and Supplementary Table 1). This clear link to human outcomes further reinforces the relevance of our experimental analysis.

We next asked what in the lung environment might be responsible for triggering the transcriptomic changes in indolent D2.0R cells. We hypothesized that these AT1-like cells might trigger the up-regulation of genes in D2.0R cells in the lung.Fig. 3f and Extended Data Fig. 7e show that AT1-like cells could indeed induce the expression of genes that are highly expressed in indolent cells *in vivo*, including a wide range of ECM genes and EMT factors. Thus, interaction with the lung parenchyma can trigger the expression of indolence-associated genes *in vitro* and *in vivo*.

Within the genes up-regulated in indolent cells, some may play a role in maintaining the cells in a non-aggressive state, hence the overall correlation with good outcomes, and others might be involved in supporting their continued survival in the lung microenvironment. Genes in this latter class might be implicated in the survival signals emanating from AT1 cells. To identify these genes we performed an *in vivo* screen using an shRNA library

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

targeting 59 genes highly expressed in indolent D2.0R cells (Supplementary Table 2). Sub-pools of an shRNA library were transduced into D2.0R cells and injected into the tail vein of mice; reference genomic DNA was also prepared from the sub-pools before injection. After 3 weeks, D2.0R cells were isolated from the lungs, the relative representation of each shRNA compared with reference genomic DNA representing the initial composition of the library was determined by sequencing (Fig. 4a). We observed that depletion of some genes promoted the outgrowth of cells in the lungs, suggesting that they function to maintain dormancy, and others reduced the numbers the of cells recovered (Fig. 4b and Extended Data Fig. 8a). As the clinical imperative is to identify ways to eliminate indolent or latent disease, we concentrated on genes that, when depleted, yielded fewer cells in the lungs. Extended Data Fig. 4b confirms the effect of *SFRP2*, *HEYL*, *SHISA2* and *CDC42EP5* when individual shRNA-transduced cells were injected independently. CDC42EP5 was not pursued as there is already literature implicating cytoskeletal genes in extravasation, which is not the focus of this study. We focused our attention on SFRP2 as this family of proteins can modulate many signalling pathways, including WNT, BMP, and the assembly of pro-survival integrin/FN complexes, and is implicated in tumour-stroma crosstalk[30,3132]. Fig. 4c confirms that multiple independent shRNAs against SFRP2 all reduced metastatic burden. *Sfrp2* depletion did not affect the initial arrest and extravasation of D2.0R cells (Extended Data Fig. 8c). Loss of *Sfrp2* expression did not affect proliferation *in vitro* (Extended Data Fig. 8d).

SFRP2 is expressed at low levels in cell culture and primary tumours, but its levels dramatically increase when in the lung environment (Extended Data Fig. 8e). Co-culture experiments demonstrated that AT1-like lung epithelial cells could induce *SFRP2* in D2.0R cells in a Src-dependent manner (Figures 5a,b), thus providing a potential explanation for the effect of SFK inhibitor observed in Fig. 2. AT1-like cells also partially induced other SFRP family members in D2.0R cells and 4T07 cells (Extended Data Fig. 8f). SFRP2 has been widely reported as WNT-signaling regulator; however, we did not observe any consistent modulation of canonical WNT targets in cell depleted of *SFRP2* (Extended Data Fig. 9a). It has been reported that SFRP2 binds FN and is incorporated into an insoluble extracellular matrix fraction[30]. Furthermore, heparin binds the C-terminus of SFRP family proteins releasing them from the ECM leading to their inactivation[30]. We confirmed that heparin could increase the level of soluble inactive SFRP2 in the media (Extended Data Fig. 9b). Notably, this treatment also reduced D2.0R cells numbers in *in vitro* co-culture experiments (Extended Data Fig. 9c). These data support suggest that insoluble extracellular SFRP2 promotes cell numbers by increasing the deposition and organisation of FN (Fig. 5c). SFRP2 over-expression promoted the organisation of FN into fibrils (Fig. 5f, bottom) and increased numbers of cell protrusions (Fig. 5d and Extended Data Fig. 9d). Furthermore, CyTOF analysis revealed an overlap in the activation of intracellular pathways between D2.0R cells co-cultured with AT1-like-cells and D2.0R cells overexpressing SFRP2 (Fig. 5e and Fig. 2a), including pPDPK1, pMKK4, pMKK3/6, and pERK. These data, combined with the effect of SFKi on *SFPR2* induction, prompted us to perform epistasis experiments. SFRP2 over-expression reduced the ability of SFKi to block the formation of protrusions and FN fibrils (Figures 5d, f, g, compare with Extended Data Fig. 5a,b). Additionally, apoptosis in the presence of SFKi was reduced when SFRP2 was over-expressed (Fig. 5h,

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Extended Data Fig. 9e shows no effect on proliferation). These analyses argue that SFRP2 supports D2.0R persistence through pro-survival integrin/FN signalling leading to enhanced output across a range of oncogenic signalling pathways (Extended Data Fig. 10d).

Finally, we tested sFRP2 over-expression *in vivo*. SFRP2 over-expressing cells had more protrusions than control cells *in vivo* (assessed using the circularity metric - Fig. 5i). SFRP2 over-expression also increased the metastatic burden both human and mouse models, with a particularly pronounced increase in large metastases (area $>5x10^5 \mu m^2$) in the 4T07 model (Figures 5j,k and Extended Data Fig. 10a). Proliferation *in vitro* in absence of AT1-like cells was not affected (Extended Data Fig. 10b), nor was proximity to other stromal cells (Extended Data Fig. 10c).

Delayed recurrence of latent disseminated cells is an unmet clinical need. Our work argues that parenchymal epithelial cells constitute a critical and under-appreciated component of the microenvironment in metastases to epithelial organs. One possible reason for the lack of attention paid to epithelial cells in the tumour microenvironment is that they are out-competed by the malignant cells in growing tumours. However, in indolent micro-metastases, or during the first steps of colonisation, they are abundant relative to the cancer cells and can exert a greater influence on their behaviour. In the future, it will be interesting to study the signals from the lung epithelium that induce *Sfrp2* and determine why some highly aggressive cancers might be able to activate survival mechanisms upon arrival in the lung whilst not being subject to growth suppressive or limiting signals. To conclude, our data indicate that carcinoma cells originating in one tissue are highly responsive to signals coming from non-transformed epithelial cells at metastatic locations. We propose that this will prove to be a recurring theme in the metastatic spread of epithelial cancers to distant epithelial tissues and, crucially, we demonstrate that interference in this crosstalk reduces survival of disseminated indolent breast cancer cells.

Methods

Cell lines

Alveolar-Type1 like cells (TT1 cells) were a kind gift of Prof. J. Downward (The Francis Crick Institute, London) and were originally derived from Prof. Terry Tetley (Imperial College, London) as described in Ref.16. T47D-DBM cells were a gift of Prof. R. Gomis (IRB, Barcelona). Alveolar-Type2 cells (H441 cells) were purchased from ATCC (HTB-174). Human Normal Lung Fibroblasts (HNLF) were derived from primary lung fibroblasts (CRUK Cell Service AG02603) immortalized with pBABE-hygro-hTERT. D2.0R, D2.A1 and MCF7-GFP cells were a kind gift of D. Barkan (University of Haifa). 4T07 were gently provided by Prof. Stefano Piccolo (University of Padua). All the cells were cultivated under standard culture conditions in DMEM/10%FBS (Thermo Fisher Scientific, 41965-039) and routinely screened for mycoplasma at Cell Services facility at The Francis Crick Institute or with Universal Mycoplasma Detection kit (ATCC, 30-1012K).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Lung organotypic system and quantification

Lung cells and breast cancer cells were plated onto Lumox 24-multiwell plate (Sarstedt, 94.699.00.14) in Mitogen Low-Nutrient Low medium (MLNL, low glucose DMEM/1%FCS, Thermo Fisher Scientific 21885025) or Mitogen High-Nutrient High medium medium (MHNH, high glucose DMEM/10%FCS, Thermo Fisher Scientific, 41965-039) as indicated. In detail: AT1-like cells ($12{,}5x10^4$ cells/well) and AT2-like cells ($2{,}5x10^4$ cells/well) were plated at day 1, HNLFs at day 2 ($2{,}5x10^4$/well) and cancer cells at day 3 (100 cells/well). Medium was replaced every three days and GFP^+ cells were manually counted under an inverted fluorescent microscope after replacing medium with HBSS. For experiments where relative number of cells/ml is shown, cells from each well were trypsinized, filtered through a 70μm cell strainer and resuspended in 200μl of FACS buffer (PBS, 2mM EDTA, 3%BSA). Number of GFP^+ cells/ml was then measured with MACSQuant Analyzer (Miltenyi Biotec) with 96well plate module.

Drug treatments

Drugs and inhibitors were added in the medium together with cancer cells (unless stated otherwise) and replaced every other day together with fresh medium. Drugs and inhibitors included in the study are: MEK inhibitor (1μM PD184352, Sigma-Aldrich PZ0181), JNK inhibitor (10μM SP600125, Tocris 1496), p38 inhibitor (10μM SB203580, Tocris 1202), Tankyrase inhibitor (5μM XAV939, Sigma-Aldrich X3004), EGFR inhibitor (1μM Lapatinib, LCLabs.com L-4804), Src-family kinase inhibitor (250nM Dasatinib, LCLabs.com D-3307), PI3K inhibitor (1μM Pictilisib, GDC-0941, Selleckchem S1065), Cilengitide (10nM, MedChem Express, HY-16141).

Metastasis assays

The study is compliant with all relevant ethical regulations regarding animal research. All protocols were in accordance with UK Home Office regulations under project licence PPL80/2368 and subsequently PPL70/8380, which passed ethical review by the LRI Animal Welfare Ethical Review Board in 2014. Briefly, murine breast cancer cells were trypsinized, washing with PBS, and then resuspended at appropriate concentration before injecting into the tail vein of mice (100μl/mouse) using a 25G needle. Prior to analysis of the lung tissue, mice were culled by a schedule 1 method. Trametinib was administered by oral gavage three times a week (drug concentration 10mg/ml, 1mg/Kg) for up to three weeks. Cilengitide (Antibodies Online, ABIN4877733) was administered intraperitoneally four times (25mg/Kg) starting at the third day after injection of cells (days: 5, 7, 10, 12). Mice were then culled after 15 days.

Quantification of disseminated cells and metastasis

For quantification of disseminated indolent cells upon gene knockdown, $5x10^5$ D2.0R-mCherry-shControl cells (Sigma-Aldrich, SHC016) were injected into the tail vein of 6- to 8-weeks old female nude athymic BALB/c mice together with $5x10^5$ D2.0R-eGFP-shRNA targeting the indicated genes. After 3 weeks, lungs were collected, processed and stained for CD45 as below. Number of $CD45^-/GFP^+$ and $CD45^-/mCherry^+$ cells were quantified by FACS and the ratio eGFP/mCherry calculated to evaluate the survival of shRNA-bearing

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

cells (EGFP) relative to an internal control (mCherry, representative gating strategy in Supplementary Fig.1).

For quantification of disseminated cells and overt metastasis upon protein overexpression, $1x10^6$ D2.0R-eGFP-SFRP2 cells or $1x10^6$ MCF7-eGFP-SFRP2 cells or $1x10^6$ T47D-DBM-eGFP-SFRP2 cells or $3x10^5$ 4T07-eGFP-SFRP2 cells were injected into the tail vein of 6- to 8-weeks old female nude athymic BALB/c mice and compared to the same number of eGFP-Control cells. After the time indicated in relevant figure legend, lungs were harvested and metastatic burden and colony area were quantified by imaging GFP colonies or cells visible from the lung surface. The imaging set up of the LSM780 is capable of detecting GFP fluorescent up to ~30μm into the tissue.

For quantification of disseminated cells after extravasation, $5x10^5$ D2.0R-mCherry-shControl cells (Sigma-Aldrich, SHC016) were injected into the tail vein of 6- to 8-weeks old female nude athymic BALB/c mice together with $5x10^5$ D2.0R-eGFP-shRNA targeting the indicated genes. Three days post injection, lungs were collected and the area of the lung surface positive for either mCherry or GFP was measured and the ratio calculated. Similarly, for analysis of individual colony size in Fig. 5, Extended Data Fig. 4 and 10, the surface of the lung was imaged and analyzed using ImageJ software. Briefly, images were thresholded to exclude background autofluorescence and the ‘Analyze Particles’ command was used to acquire the metrics for every contiguous patch of signal (i.e. colony). For experiments using MCF7 and T47D cells mice were implanted with a beta-estradiol pellet one week before the injection of cancer cells (0.72mg/pellet, 90 day release).

Tissues dissociation

Lungs and primary tumors were harvested from mice, immersed in PBS, and promptly chopped up with scissors to small fragments. Minced lungs were then added to digestion solution (PBS buffer with 75μg/ml TM Liberase, Roche 05401151001, 75μg/ml TH Liberase, Roche 05401127001, 12.5μg/ml DNAse, Sigma-Aldrich DN25) for 1hr at 37°C on a rocker. Digested lung pieces were spun down for 5’ at 1300rpm, re-suspended in calcium- and magnesium-free PBS containing 1mM EDTA by vigorous pipetting until the solution was homogeneous and then filtered through a 70μm cell strainer to remove undigested fibrous tissue. In the case of stiffer tissues, such as primary tumors, tissue fragments were also mechanically disrupted by passing them through needles of decreasing thickness. Cells were then pelleted and red blood cells lysed with Red Blood Cells Lysis Solution (Miltenyi Biotec, 130-094-183) following manufacturer protocol. After washing, cells were re-suspended in FACS buffer (PBS, 2mM EDTA, 3%BSA) and labelled with CD45-APC antibody for 30min (eBiosciences, 30-F11, 1:400) to avoid contamination from leukocytes during sorting. Samples were then washed repeatedly, filtered through a 70μm cell strainer and kept on ice during fluorescence-activated cell sorting.

Gene expression studies

For gene expression studies of cancer cells co-cultured with lung stromal cells, $1,36x10^6$ AT1-like cells/dish were plated onto 6cm dishes on day 1 (in MLNL medium) followed by $6x10^4$ cancer cells the following day, in restrictive medium. On day 5, GFP^+ cells were

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

trypsinized, passed through a 40μm strainer, re-suspended in HBSS/2mM EDTA and sorted according to GFP positivity (Bio Rad S3e Cell Sorter) directly into lysis buffer (1,5-3x10^4 cells/sample). Total RNA extraction was performed using Total RNA Purification Plus Kit (Norgen Biotek, 48400) according to manufacturer protocol and the whole RNA eluate was retrotranscribed with SuperScript III (Thermo Fisher Scientific, 18080044) using oligo(dT) as primers. cDNA was further purified with QIAquick PCR Purification kit (Qiagen, 28106) before qPCR analysis was carried out with triplicate samplings of each sample cDNA on QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific) with a FastStart SYBR Green Master Mix (Roche 04673492001).

For RNA sequencing experiments of disseminated breast cancer cells *in vivo*, 1x10^6 D2.A1-eGFP cells or D2.0R-eGFP cells were injected into the tail veins of 6- to 8-weeks old female nude athymic BALB/c mice (Charles River). After 3 weeks lungs were removed, digested into a single cell suspension as described and labelled with CD45-APC as indicated above. $CD45^-$/$eGFP^+$ cells were sorted (Flow Cytometry Facility at CRUK-LRI and The Francis Crick Institute) directly into lysis buffer and total RNA was extracted with RNeasy Plus Micro kit (Qiagen) following manufacturer protocol. RNA samples were assessed for quantity and integrity using the NanoDrop 8000 spectrophotometer V2.0 (Thermo Fisher Scientific) and Agilent 2100 Bioanalyser (Agilent Technologies), respectively. Samples displayed low levels of degradation with RNA integrity numbers (RIN) between 6.4 and 7.8. Full-length cDNA molecules were generated from 4ng of total RNA per sample using the SMARTer kit for cDNA generation (Clontech). cDNA quantity was measured using the dsDNA High-sensitivity Qubit kit with the Qubit 2.0 Fluorometer (Thermo Fisher Scientific), and were checked for quality using a D1000 ScreenTape with the Agilent 2200 Tapestation (Agilent Technologies). Libraries were prepared using the Illumina Nextera XT Sample Preparation Kit (Illumina Inc.) with an input of 150pg of cDNA per sample. Resulting libraries were checked for average fragment size using the Agilent D1000 ScreenTape, and were quantified using the Qubit dsDNA High-sensitivity reagent kit. Equimolar quantities of each sample library were pooled together and 75bp paired-end reads were generated for each library using the Illumina NextSeq 500 High-output sequencing kit. For *in vitro* samples, breast cancer cells were grown in multiwell plates under standard culture conditions, trypsinised, sorted and processed in parallel with the *in vivo* samples.

For qPCR analysis of disseminated breast cancer cells *in vivo*, cells were isolated and total RNA purified as above. In order to obtain enough cDNA as template for qPCR analysis, total RNA was amplified with Arcturus RIboAmp HS PLUS kit before retrotranscription with with dT-primed M-MLV Reverse Trascriptase (Thermo Fisher Scientific, 28025013). qPCR analysis was carried out on QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific) with Fast SYBR Green Master Mix (Applied Biosystems 4385612).

For gene expression studies of orthotopic breast tumors, 1x10^6 D2.A1-eGFP cells or D2.0R-eGFP cells were injected into mammary fat pad of 6- to 8-weeks old female nude athymic BALB/c mice (Charles River). After 12 days tumor masses were harvested, processed and sorted as above. For *in vitro* samples, breast cancer cells were grown in multiwell plates, trypsinised, labelled and sorted in parallel with the *in vivo* samples.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

For gene expression studies of breast cancer cells treated with conditioned medium, $4x10^6$ AT1-like cells were plated in 10cm/dishes with MLNL medium. After 48hrs medium was collected, cleared from dead cells and debris by centrifugation (20min at maximum speed) and added to breast cancer cells. After 12hrs cells were collected and total RNA isolated using Total RNA Purification Plus Kit (Norgen Biotek, 48400) according to manufacturer protocol. Total RNA was retrotranscribed with dT-primed M-MLV Reverse Trascriptase. qPCR analyses were carried out with triplicate samplings of each sample cDNA on QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific) with a FastStart SYBR Green Master Mix.

All expression levels were calculated relative to *GAPDH*. Oligo sequences used in this study are listed in Supplementary Table 3.

Time lapse

$2x10^4$ AT1-like cells/well were plated onto Lumox 24-multiwell plate (Sarstedt, 94.699.00.14) in MLNL or MHNH medium as indicated. The following day 2000 D2.0R cells were plated in the same media. 3-4 hours after plating the indicated inhibitors were added and imaging for 48 hours using either a LSM510 or Nikon Eclipse Ti2 was commenced two hours later. The movies were analyzed manually to record the number of cells at the beginning, at the end, the number of mitoses, and the number of cell death events.

Library screening

A custom shRNA library was designed based on our *in vivo* gene expression data and synthesized by Sigma-Aldrich (custom MISSION shRNA library). All shRNAs are cloned inside pLKO.1-based plasmids (TRC version as indicated in Supplementary Table 2) and were individually amplified to avoid representation biases of the clones. We generated 12 shRNA pools, or sets, by randomly combining 14-15 shRNA plasmid clones per set and including a non-targeting control shRNA in each pool (Sigma-Aldrich, SHC016) as a quality control of the procedure (i.e. a shRNA not leading to enrich/depletion of cells) and not with normalization purposes. Plasmid DNA of each set was individually transfected in 293FT cells together with packaging plasmids (pMD2, psPAX2), harvested after 48hrs and added to D2.0R-eGFP cells at a low concentration to ensure a single shRNA integration per cell. Successfully transduced cells were selected with puromycin and injected into the tail veins of 6- to 8-weeks old female nude athymic BALB/c mice (3 mice/pool, $3x10^6$ cells/mouse). After 3 weeks, lungs have been collected and $CD45^-/EGFP^+$ D2.0R cells isolated as above. Genomic DNA was purified from sorted cells, as well as from pre-injection samples, with QIAmp DNA Micro Kit (Qiagen) and used as template for 2 rounds of PCR prior to Next Generation Sequencing. In the first round of PCR we used a forward primer with unique barcode sequence for each pool, while in the second reaction we used primers containing adaptor sequences for NGS. All primers and barcodes are listed in Supplementary Table 3. After PCR amplification, DNA fragments were purified and combined in order to obtain four sets, each one containing one sample/pool (one sample pre-injection, three samples after *in vivo* selection). Samples were sequenced on a Paired End 101 bp run (Illumina

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

HiSeq 2500) and the representation of each shRNA post-injection relative to the representation pre-injection was calculated as described in “Bioinformatic analysis” section.

Stable protein expression

Fluorescent proteins were stably expressed in cancer cells by transduction with retroviruses. pCX4-neo-GFP or pCX4-blasti-mCherry plasmids were transfected into 293T cells together with packaging plasmids (pGP, pVSVG). After two days, surnatants were collected, filtered through a 0.45μm filter and added to indicated cells for two days before selection with the appropriate drugs. SFRP2 protein was overexpressed in cancer cells by transduction with lentiviral particles. pLV-hygro-mSFRP2 (VectorBuilder, custom) plasmid was transfected into 293T cells together with packaging plasmids (pMD2, psPAX2). As control plasmid we used pCSII-IRES2-hygro (kind gift of Prof. S. Piccolo, University of Padua). After two days, surnatants were collected, filtered through a 0.45μm filter and added to indicated cells for two days before selection with hygromycin. Overexpression of SFRP2 mRNA was confirmed by qPCR using oligos amplifying a sequence within the coding sequence of the cDNA.

Proliferation assays

Breast cancer cells were plated on flat bottom 96 well plates (2000 cells/well) and confluency measured over time with Incucyte (Essen Bioscience) every 3-4hrs for 100hrs. Percentage of covered area was Log10-trasformed and plotted against time. The 95% confidence bands of the best-fit line were plotted and, for the purpose of plotting, line is forced to go through X=0.

Bioinformatic analysis

RNAseq. Sequencing was performed on biological replicates for each condition generating approximately 31.8 million 75 bp paired end reads. The RSEM package (version 1.2.11) and Bowtie2 were used to align reads to the mouse mm10 transcriptome, taken from refGene reference table available at UCSC downloaded on May 2014 [https://genome.ucsc.edu/]. For RSEM, all parameters were run as default. TMM (treated mean of M-values) normalisation and differential expression analysis using the negative binomial model was carried out with the R-Bioconductor package “Deseq2” (www.bioconductor.org R version 3. 1.0). Genes were considered to be differential expressed if the adjusted p value were less than 0.05. Geneset enrichment Analysis, GSEA, (version 2.2.3[33]) was carried out using ranked gene lists using the Wald statistic and genesets of C2 canonical pathways, C5 biological processes and additional published gene sets (Supplementary Table 4). All parameters were kept as default except for enrichment statistic (classic) and max size which was changed to 5000 respectively. Gene signatures with FDR q-value equal or less than 0.25 were considered statistically significant. For the heatmap in Fig. 3a, genes were clustered using a Euclidean distance matrix and average linkage clustering. Red indicates higher expression and blue indicates low expression relative to the mean expression of the gene across all samples. In Fig. 3b, GSEA results from [D2.0R vs other groups] were visualized using Cytoscape (version 3.6.0) and Enrichment Map plug-in[34]. The map has been manually annotated to reduce complexity and redundancy. Probe 223122_s_at on kmplot.com was used to stratify distant metastasis free survival of breast cancer patients according to *SFRP2* expression.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

shRNA library screening. Illumina sequence reads with “internal” barcodes (each barcode corresponds to a specific pool of shRNA) were demultiplexed into individual sample files, hairpin sequence was extracted from the backbone vector and common reads collapsed to “tags” providing one instance of each unique candidate hairpin sequence, along with a count of the total number of appearances of each in the original files using in house code. These ‘tag’ sequences were mapped against all annotated library sequences (Supplementary Table 2) using bwa-0.5.9[35] and counts of total sequences mapping to each target (counting the total original instances of each hairpin sequence) were generated. These counts were subsequently restricted to consider only targets appearing in the pool specific to that sample and these total raw counts were normalised to the maximum total number of reads across all samples to allow direct comparisons between samples. For each experimental set, a fold change of the representation of each shRNA post-selection relative to the control levels of the same shRNA pre-selection was calculated and these were log-2 transformed. To facilitate this, zero counts in the control were offset by 0.5 to allow the division and zero ratios were set to 1 to allow the extraction of logs and enable subsequent clustering. We then ranked genes according to a representation score, defined as the median of the log-2 fold change values. Candidate genes were selected based on two criteria: 1. knock-down of the gene led to loss of dormant cells carrying that shRNA, 2. consistent effect of at least two out of three shRNA sequences.

Single-Cell Signaling Analysis by Mass Cytometry

D2OR, D2OR-SFRP2 overexpressing, MCF7 and AT1-like cells alone or in coculture were treated with 25 μM ^{127}5-Iodo-2′-deoxyuridine (^{127}IdU - Fluidigm 201127) for 30 mins[36]. Thereafter, the media was removed and the cells were fixed with 4% PFA, and dissociated into single-cells using 2U/mL Dispase (Sigma D4693). Cells from each experimental condition were barcoded using the Cell-ID™ 20-Plex Pd Barcoding Kit (Fluidigm 201060)[37], pooled into a single-tube, blocked with Cell Staining Buffer (CSB, Fluidigm 201068), and stained with extracellular rare-earth metal conjugated antibodies (listed below). Cells were then washed in CSB, permeabilised with 0.1 % Triton X-100 in PBS and then with ice-cold 50% methanol, and stained with intracellular rare-earth metal conjugated antibodies (listed below). Cells were then washed in CSB, fixed in 1.6% FA (Pierce 28906) for 10 mins and then incubated in DNA Intercalator (^{191}Ir & ^{193}Ir - Fluidigm 201192) overnight at 4°C. Cells were then washed in water, diluted to 0.5x10^6 cells/mL and EQ Four Element Calibration Beads (Fluidigm 201078) were added at a 1:5 ratio[38]. Cells were analysed using a Helios Mass-Cytometer (Fluidigm) at 100-300 events/sec. Files were normalised against EQ beads and de-barcoded into each experimental condition using Fluidigm’s CyTOF Software (version 6.7.1014) and uploaded to the Cytobank platform (www.cytobank.com). Events were gated for Gaussian parameters (Event length, Centre, Residual, and Width values) and DNA^{high} (^{191}Ir and ^{193}Ir) to identify cells. Earth Mover’s Distance (EMD)[36] was calculated with the Python package *scprep* (https://github.com/KrishnaswamyLab/scprep) using default parameters[39], DREVI (conditional-Density Rescaled Visualization) plots and DREMI (conditional- Density Resampled estimate of Mutual Information) scores were generated using the MATLAB program *simpledremi* (https://github.com/dpeerlab/DREMI). Signalling network models were compiled in OmniGraffle 7. All the scripts used for data analysis are available at the repository of the

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Cell Communication Lab at UCL's Cancer Institute (https://github.com/TAPE-Lab/CyTOF_DataAnalysis). List of antibodies used is provided in Supplementary Table 5.

Immunohistochemistry

FFPE material was cut into 5μm sections and subject to antigen retrieval using heated citrate buffer (pH6). Incubation with both primary and secondary antibody was performed at room temperature for 45-60 minutes. GFP was detected using Goat anti-GFP (1:300, Abcam AB6673) followed by Donkey anti-Goat 555 (Invitrogen A-21432).

For frozen sections, lungs were perfused with 4% PFA in PBS immediately post mortem before transitioning through 30% sucrose for 24 hours into OCT and rapid freezing. 10μm sections were cut before staining. Slides were fixed in 4% PFA for 15 minutes at room temperature. After washes, cells were permeabilized with PBS/0.2%-TritonX for 5 minutes at room temperature and blocked with IF buffer (PBS/0.05%-Tween20/3%BSA for Ki67 or PBS/3%BSA for other staining) for 1hr. Primary antibodies were incubated in IF buffer overnight at 4°C in a wet chamber. The day after, cells were washed several times with IF buffer and incubated with secondary antibodies for at least 1hr at room temperature together with DAPI (1mg/ml stock, 1:500, Sigma-Aldrich D9542) and Phalloidin (Phalloidin-Atto633, 20μM stock, 1:1000, Sigma-Aldrich 68825) when indicated. Images were acquired with a Zeiss LSM 780 using ZEN software. Antibodies used in this study are: PDPN (1:100, Acris DM3501), AQP5 (1:100, Abcam ab78486), SP-C (1:100, Abcam Ab90716), CD68 (1:100, Biolegend 137004), Vimentin (1:100, Abcam ab92547), αSMA (1:200, Sigma C6198). EdU incorporation was visualized with Click-iT Plus Edu Alexa Fluor 647 (Invitrogen C10640) in accordance with the manufacturer's instructions. For *in situ* staining, the same steps were performed (excluding the freezing in OCT and sectioning) with the modification that all blocking and antibody steps were performed for at least 24 hours at 4°C.

Immunofluorescence

Cells were fixed in 4% PFA for 15 minutes at room temperature. After washes, cells were permeabilized with PBS/0.2%-TritonX for 5 minutes at room temperature and blocked with IF buffer (PBS/0.05%-Tween20/3%BSA for Ki67 or PBS/3%BSA for other staining) for 1hr. Primary antibodies were incubated in IF buffer overnight at 4°C in a wet chamber. The day after, cells were washed several times with IF buffer and incubated with secondary antibodies for at least 1hr at room temperature together with DAPI (1mg/ml stock, 1:500, Sigma-Aldrich D9542) and Phalloidin (Phalloidin-Atto633, 20μM stock, 1:1000, Sigma-Aldrich 68825) when indicated. Images were acquired with a Zeiss LSM 780 using ZEN software. Antibodies used in this study are: Ki-67 (1:1000, Abcam ab15580), Fibronectin (1:500, Sigma F3648), phospho-SRC Y418 (1:100, Invitrogen, 44-660G).

Western blotting of conditioned medium

To visualize soluble SFRP2 protein, confluent D2.0R cells were cultivated in DMEM without serum. After 5 days, conditioned medium was pooled from three 15cm dishes/condition, spun 20' at maximum speed to remove debris and then concentrated by spinning the samples for 30' at 4°C at 3000rcf (Amicon Ultra-15 Centrifugal Filter Devices 30,000

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

MWCO, Millipore). As loading control, remaining cells were harvested and processed as in[40]. Western blotting was performed as in[40]. Antibodies: SFRP2 (1:1000, Abcam, ab137560), GAPDH (1:25000, Millipore, MAB374). Antibody for SFRP2 has been validated with recombinant mouse SFRP2 (R&D, 1169-FR).

Cell morphology assessment

To calculate circularity we used the Image J plug-in described in the following link: https://imagej.nih.gov/ij/plugins/circularity.html This calculates *circularity = 4pi(area/perimeter^2)* When <50 cells were being measured, manual tracing of cell outline was used to ensure that single cells were being analysed; when n>50 then automatic thresholding was used. This latter method precludes a definitive determination of whether a GFP+ve patch contains a single cell or a small cluster of cells. Hence, we utilize the term cell/colony circularity to reflect that the measurement includes both isolated cell and micro-cluster values. Cell extensions >15 microns in length were classified as protrusions in manual scoring.

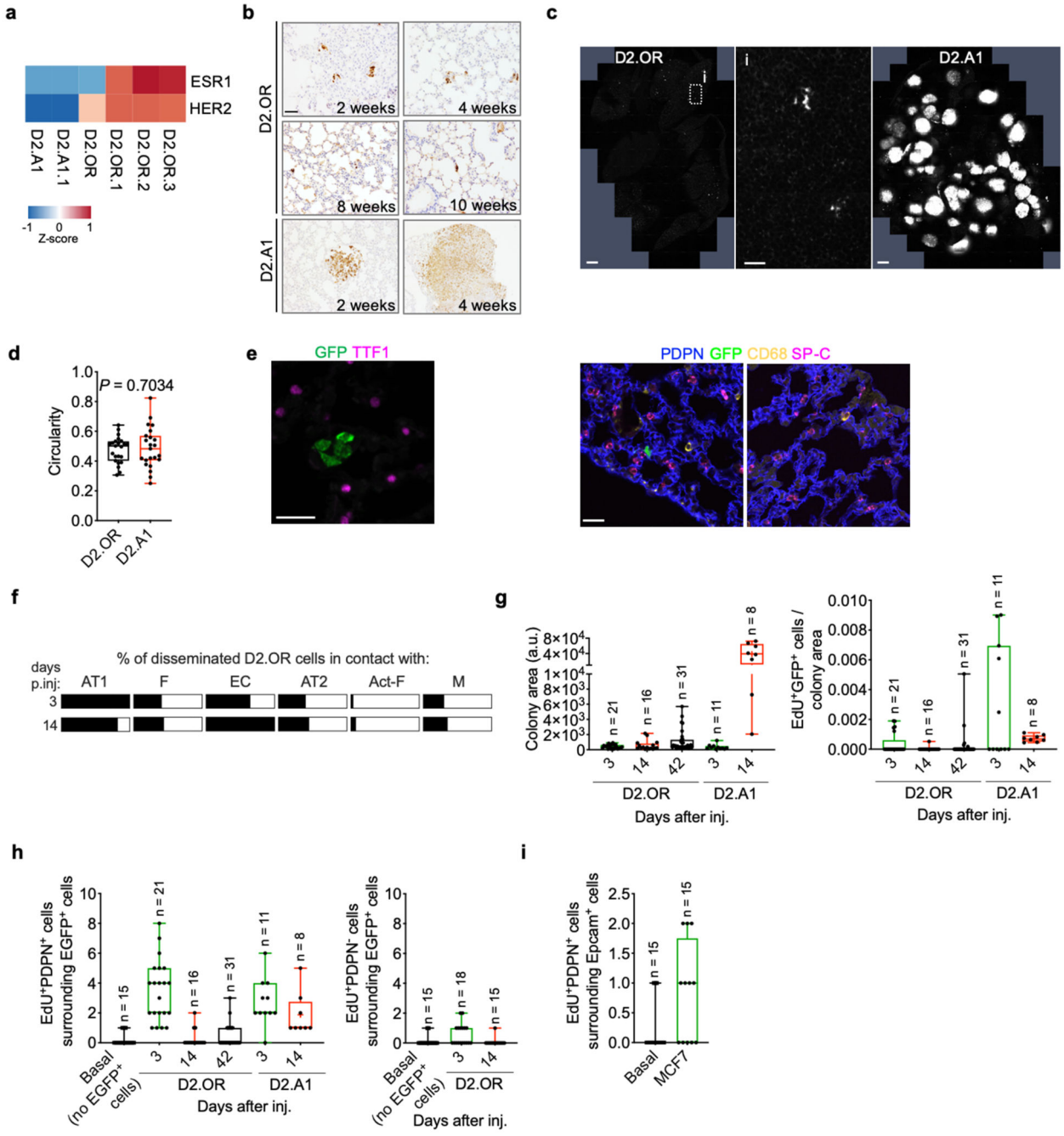
Statistics and reproducibility

Statistical analyses used GraphPad Prism software. For experiments with samples-sizes greater than 10, normality of data was tested with Shapiro-Wilk test. For normally distributed samples, we performed Student's two-tailed t-test for single comparisons (paired or unpaired) and one-way ANOVA analysis for multiple comparisons. In case of different variances within samples to be compared we applied Welch's correction. For non-normal data, we performed two-tailed Mann-Whitney test for analysis of unpaired data and two-tailed Wilcoxon matched pairs rank test for paired data. For multiple comparisons of non-normal data, we applied Dunn's test. For samples below 10 in size, it is not easy to assess the underlying distribution of the data and non-parametric tests were preferred, unless the sample-size was below 5, where we preferred parametric tests due to the minimum possible p-value becoming large in the non-parametric case. Data are plotted as the mean of all independent experiments. In some experiments the mean-normalised values from all independent experiments are plotted to provide information about assay variability. For animal experiments, each mouse was considered as a biologically independent sample. Linear regression p-values are calculated from the observed t-statistic ratio of the parameter estimates to their standard errors. For survival plots (Kaplan-Meier analysis), data were analysed with GraphPad Prism software, GOBO (http://co.bmc.lu.se/gobo/gsa.pl) or KM Plotter (https://kmplot.com/analysis/) online tools which all calculate log-rank p-value (Mantel-Cox method). For analysis with GraphPad Prism, p-value calculated with Gehan-Breslow-Wilcoxon methods is provided. GSEA is generated from GSEA online tool (http://software.broadinstitute.org/gsea/index.jsp), which also calculates the two primary statistics of the analysis: NES and FDR. Normalised Enrichment Score (NES) is calculated by normalising Enrichment Score to gene sets size, False Discovery Rate (FDR) represents an estimated likelihood that a gene set with a given NES represents a false positive. The threshold for significance was set at 0.05 for all experiments except for GSEA where we considered a significant FDR as below 0.25. Data in histograms are presented as mean -/+ SD (if present) unless stated otherwise.

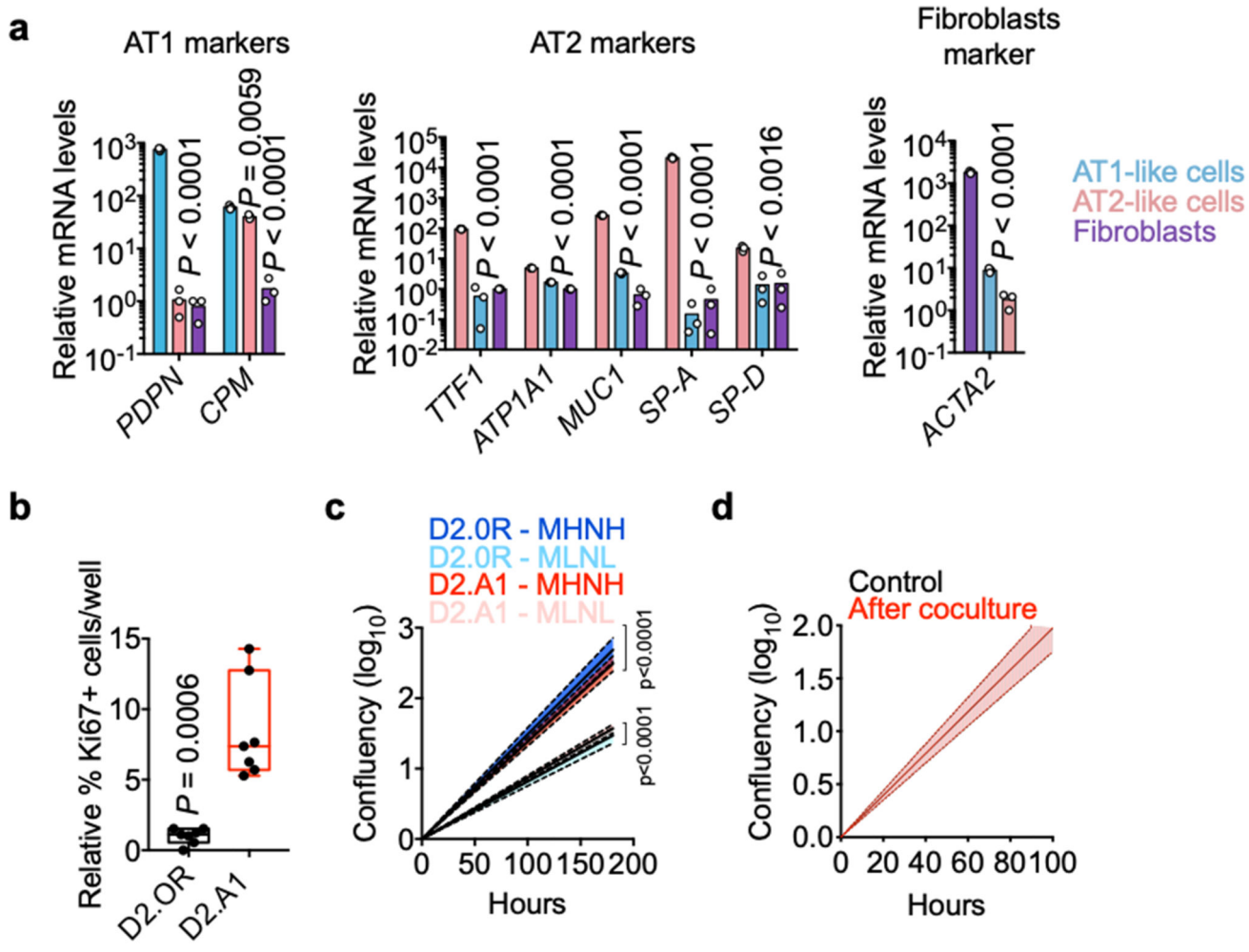
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Extended Data

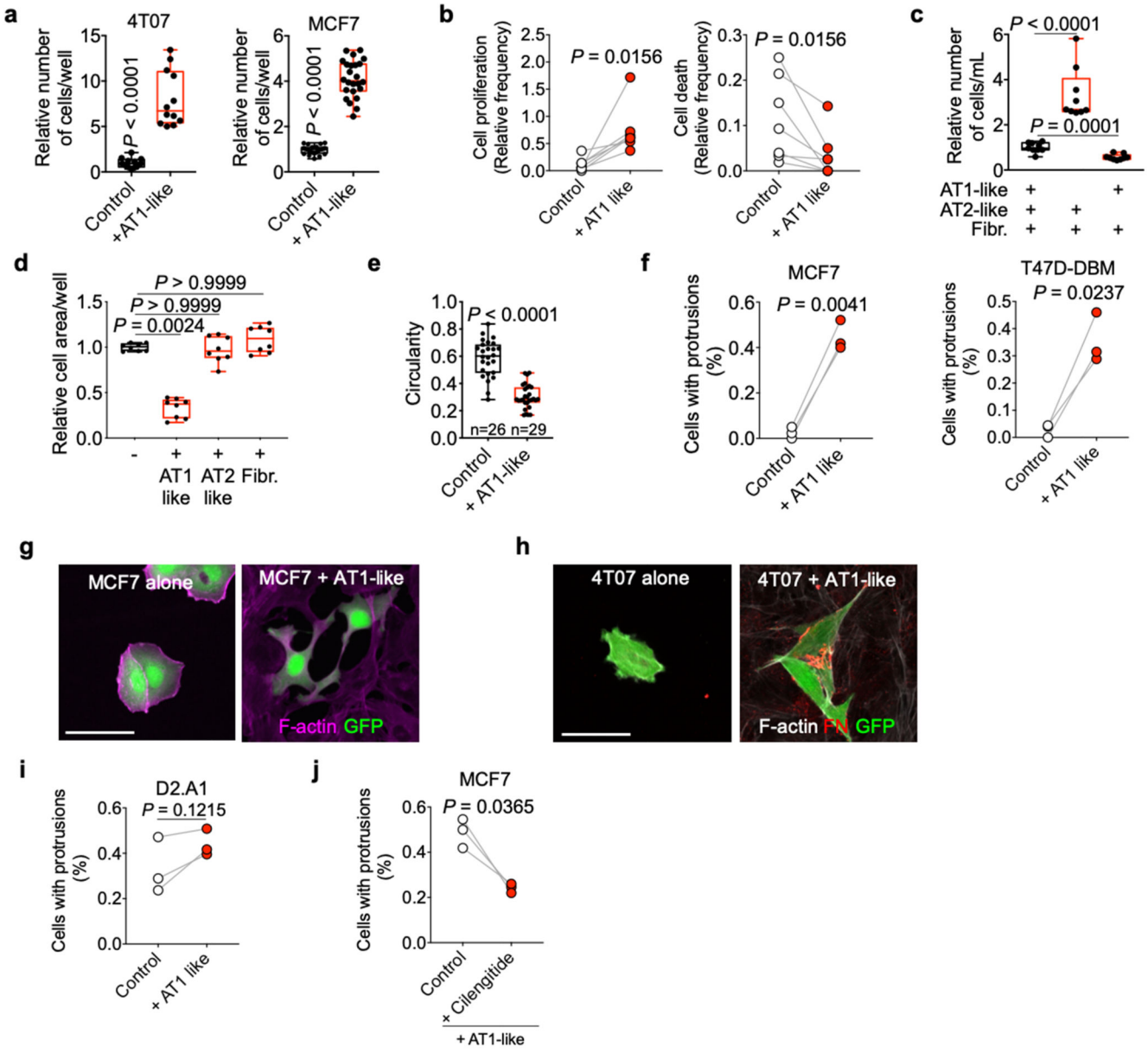


Extended Data Figure 1.



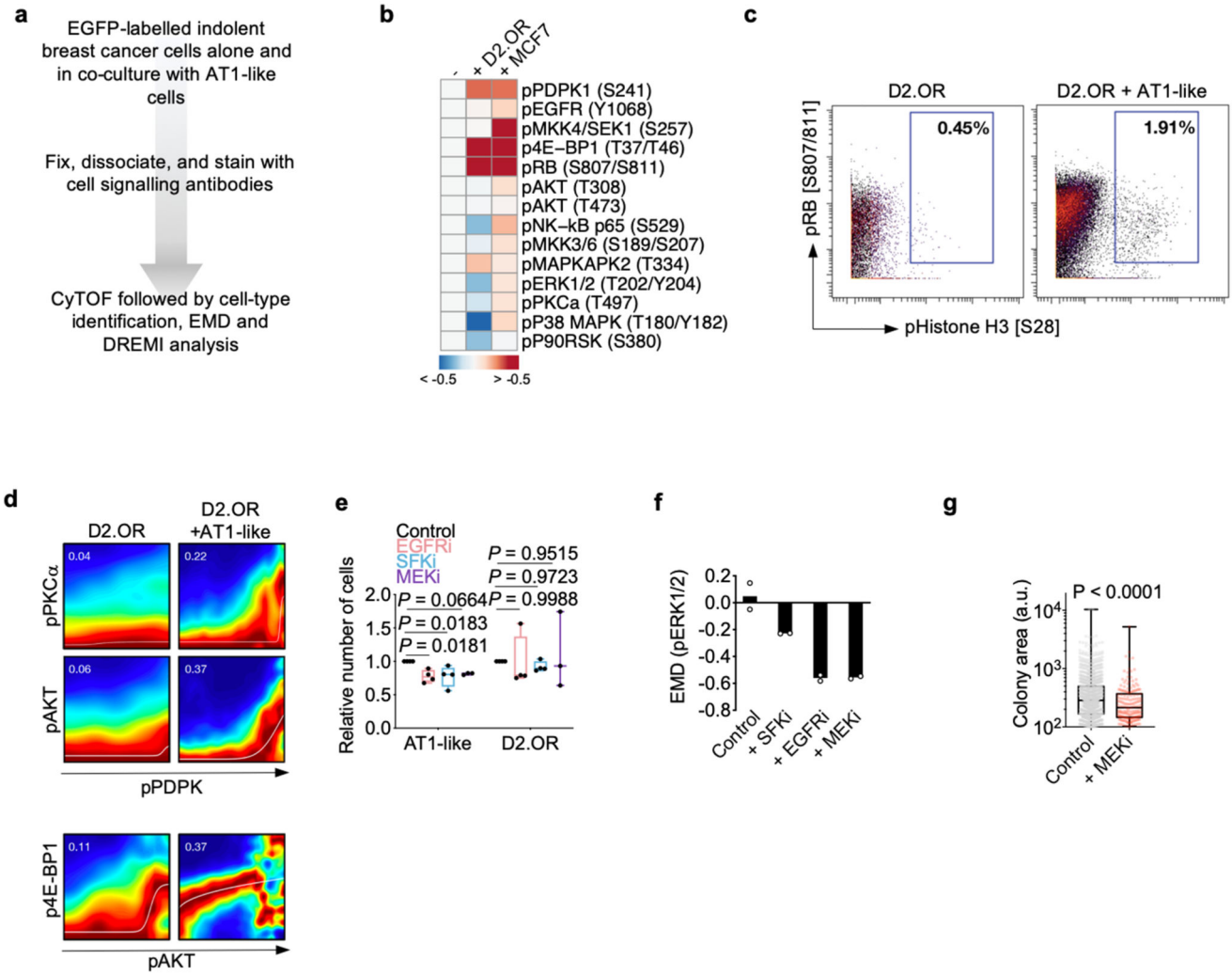
Extended Data Figure 2.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts



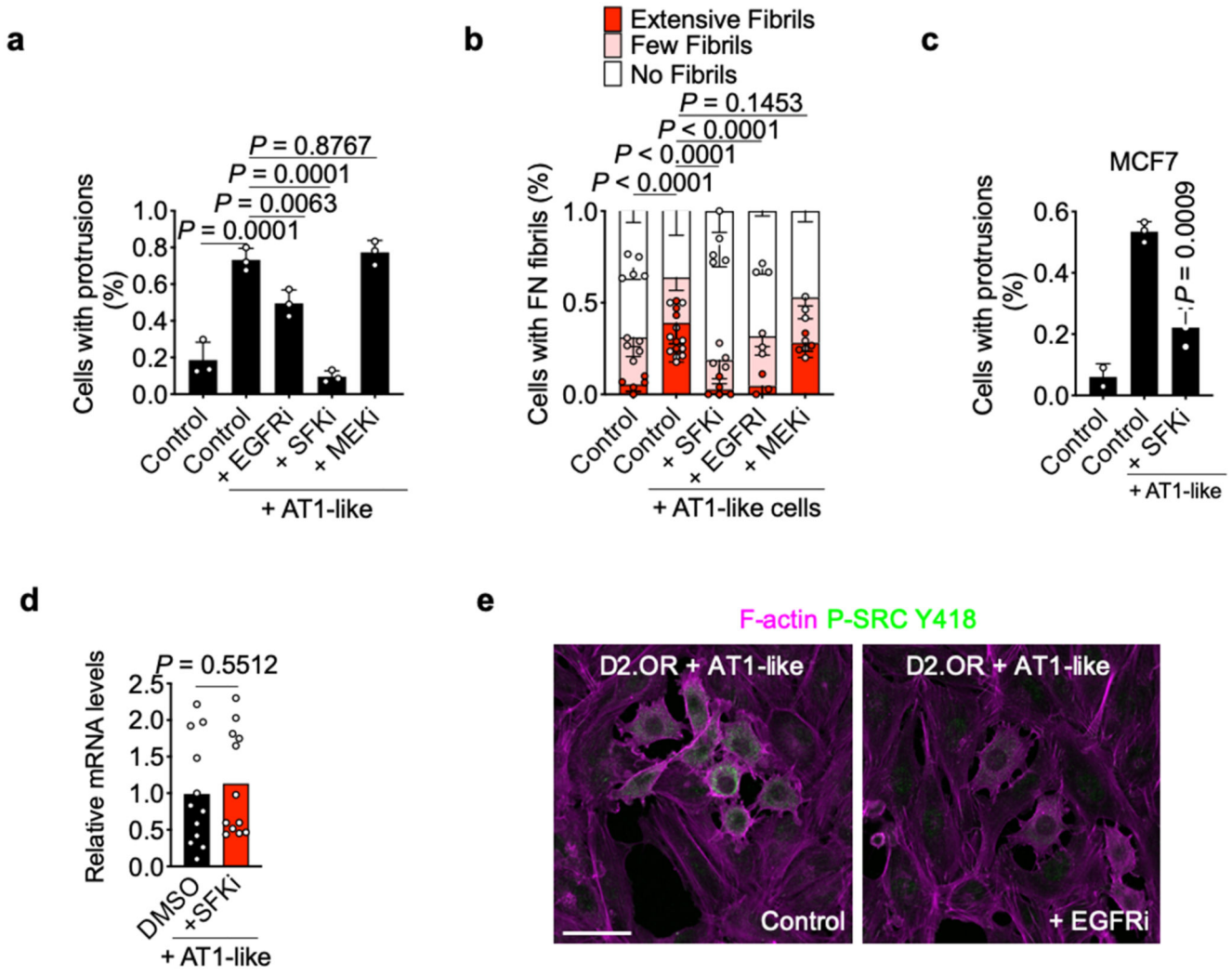
Extended Data Figure 3.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts



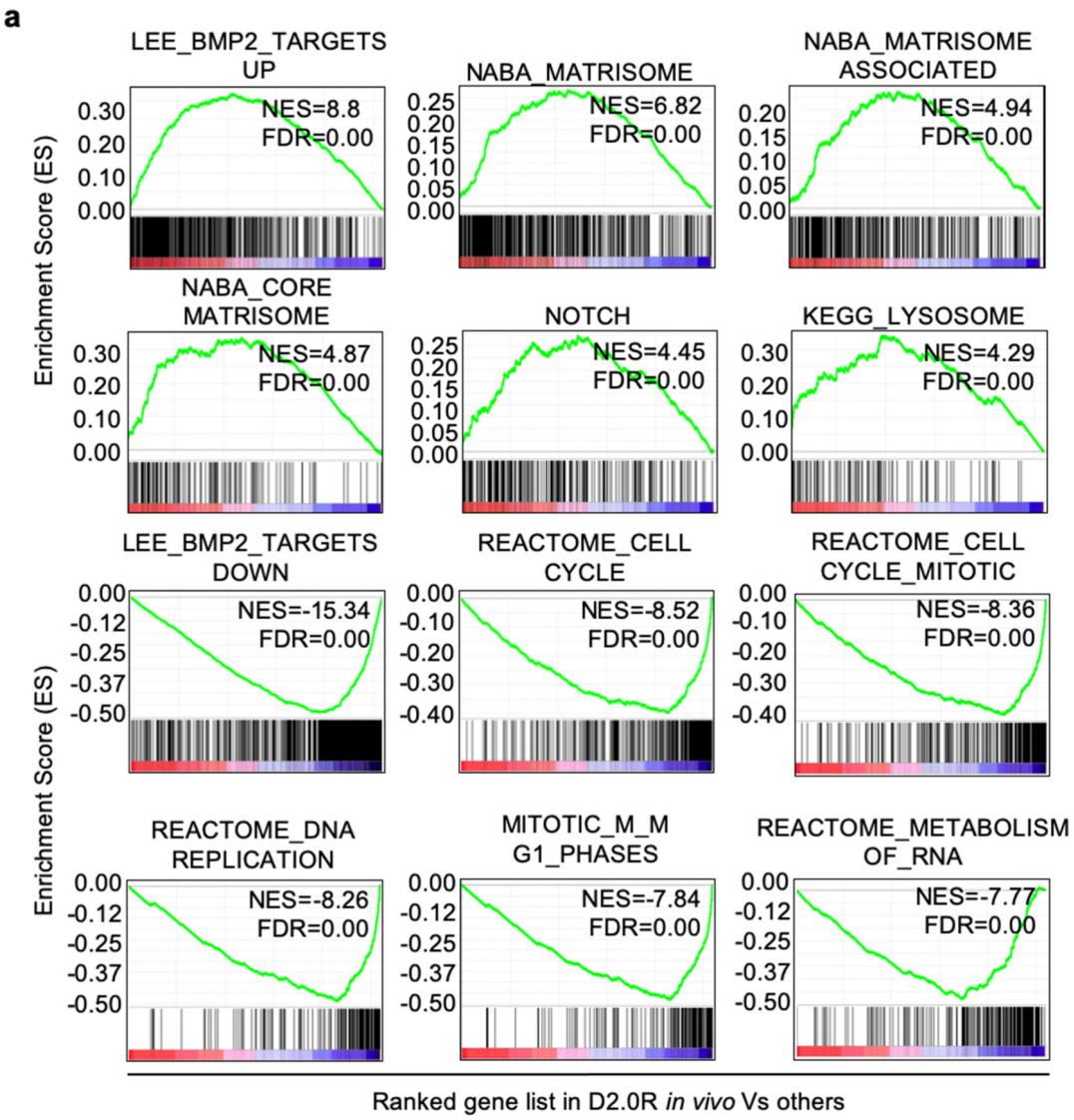
Extended Data Figure 4.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts



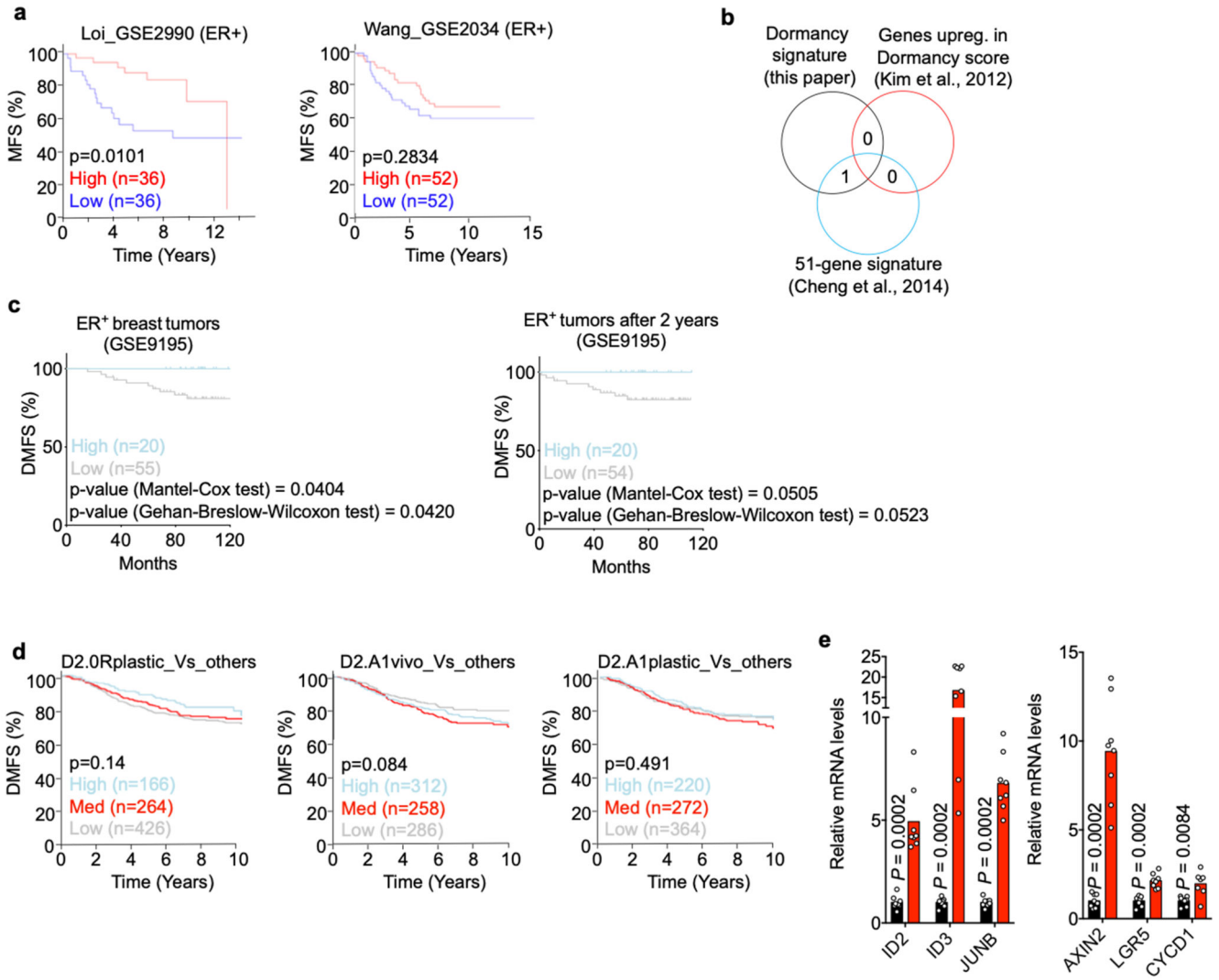
Extended Data Figure 5.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts



Extended Data Figure 6.

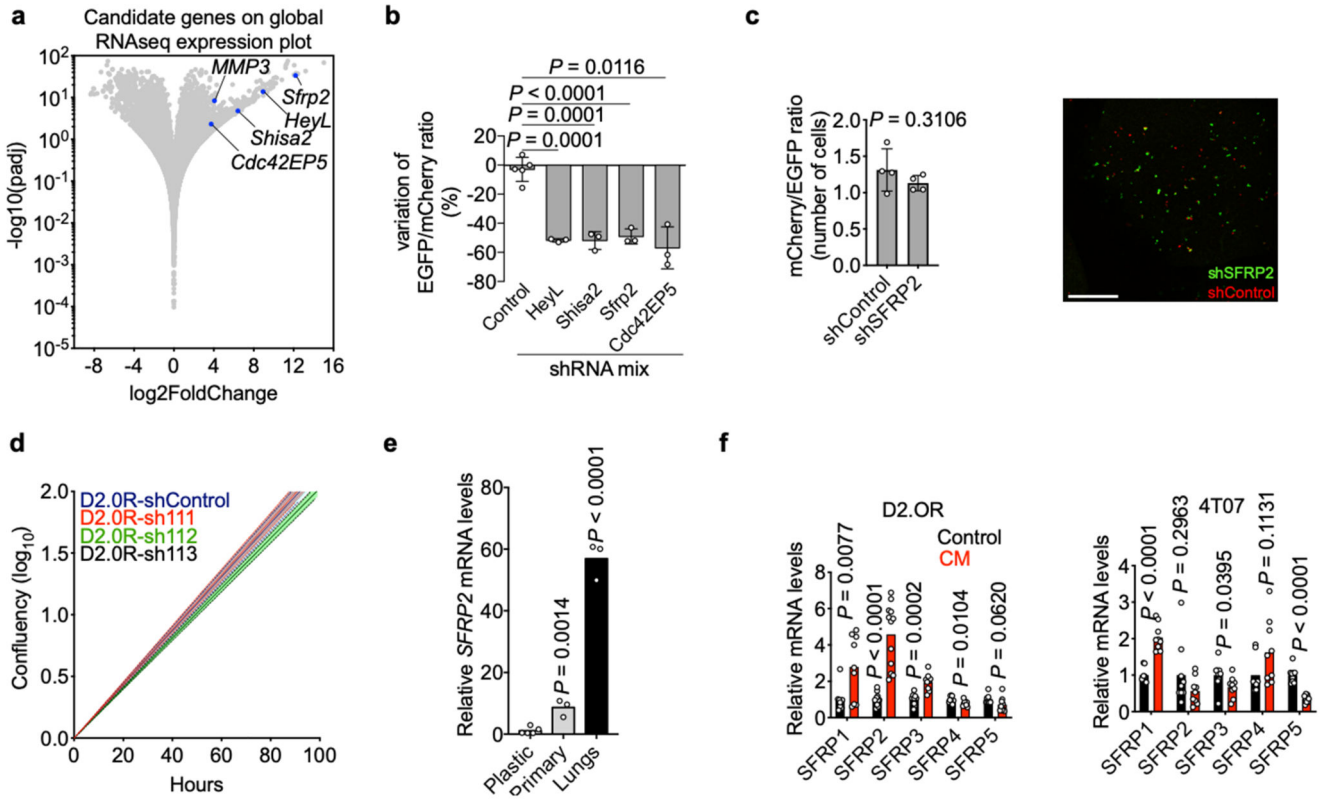
Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts



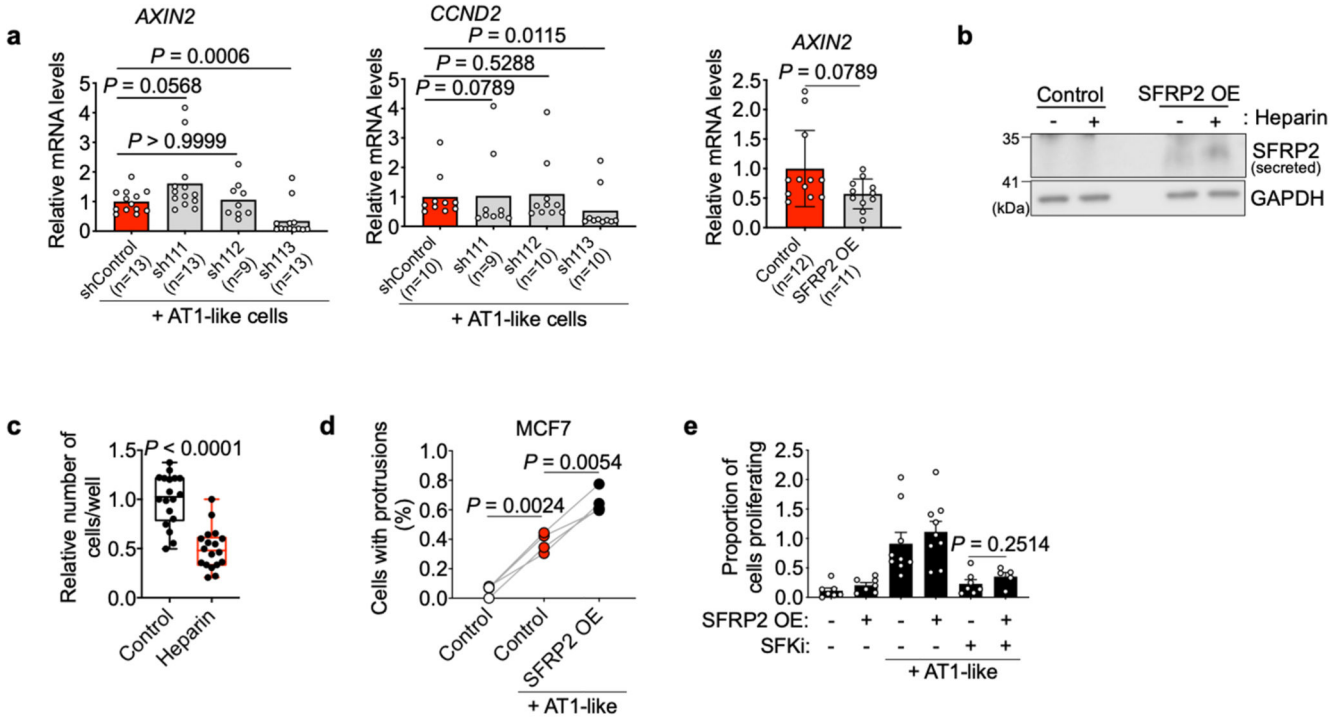
Extended Data Figure 7.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

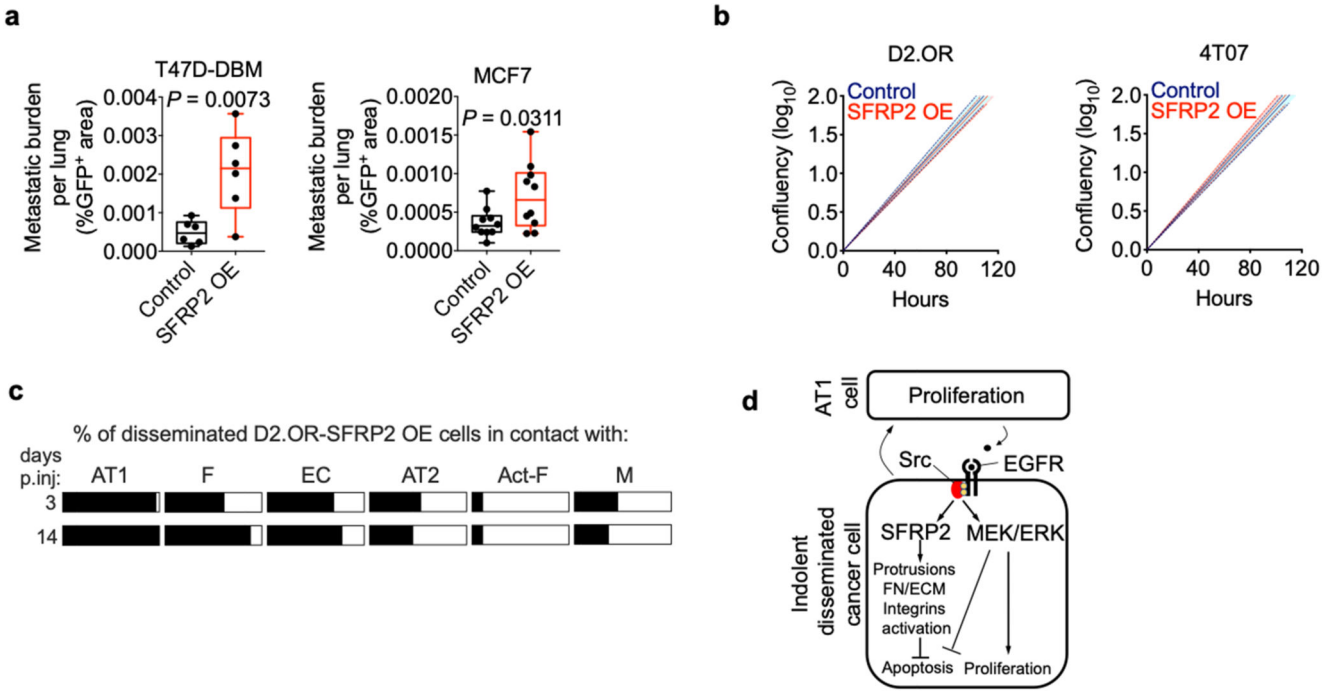
Europe PMC Funders Author Manuscripts



Extended Data Figure 8.



Extended Data Figure 9.



Extended Data Figure 10.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Julian Downward (Crick Institute), Dalit Barkan (University of Haifa) and Prof. R. Gomis (IRB, Barcelona) for gifts of cell lines. We are indebted to Ilaria Malanchi (Crick Institute), Stefano Piccolo, Sirio Dupont, and Graziano Martello (University of Padua) for thoughtful discussion and reagents. We are indebted to Flow cytometry, Experimental Histopathology, Bioinformatics and Biostatistics (in particular Stuart Horswell), Biological research, Cell services and Advanced sequencing facilities at the Crick Institute for exceptional scientific and technical support throughout the project. We thank Charles Mein (Bart's and the London School of Medicine and Dentistry) for support and advice with RNA sequencing. E.S. and M.M were funded by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001144), the UK Medical Research Council (FC001144) and the Wellcome Trust (FC001144). M.M. also received funding from Marie Curie Actions—Intra-European Fellowships #625496 and BIRD Seed grant from Department of Molecular Medicine (University of Padua). C.T. and J.S. are supported by a Cancer Research UK Career Development Fellowship awarded to C.T.

Data availability

RNAseq data have been deposited at GEO Database (GSE120628). Other data that support the findings are available upon reasonable request from the corresponding authors.

References

1. Lambert AW, Pattabiraman DR, Weinberg RA. Emerging Biological Principles of Metastasis. Cell. 2017; 168:670–691. [PubMed: 28187288]

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

2. Hedley BD, Chambers AF. Tumor dormancy and metastasis. Adv Cancer Res. 2009; 102:67–101. [PubMed: 19595307]
3. Sosa MS, Bragado P, Aguirre-Ghiso JA. Mechanisms of disseminated cancer cell dormancy: an awakening field. Nat Rev Cancer. 2014; 14:611–622. [PubMed: 25118602]
4. Obenauf AC, Massague J. Surviving at a Distance: Organ-Specific Metastasis. Trends Cancer. 2015; 1:76–91. [PubMed: 28741564]
5. Malladi S, et al. Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. Cell. 2016; 165:45–60. [PubMed: 27015306]
6. Avgustinova A, et al. Tumour cell-derived Wnt7a recruits and activates fibroblasts to promote tumour aggressiveness. Nat Commun. 2016; 7 10305 [PubMed: 26777421]
7. Naumov GN, et al. Persistence of solitary mammary carcinoma cells in a secondary site: a possible contributor to dormancy. Cancer Res. 2002; 62:2162–2168. [PubMed: 11929839]
8. Barkan D, et al. Inhibition of metastatic outgrowth from single dormant tumor cells by targeting the cytoskeleton. Cancer Res. 2008; 68:6241–6250. [PubMed: 18676848]
9. Sowder ME, Johnson RW. Enrichment and detection of bone disseminated tumor cells in models of low tumor burden. Sci Rep. 2018; 8 14299 [PubMed: 30250146]
10. Shibue T, Brooks MW, Weinberg RA. An integrin-linked machinery of cytoskeletal regulation that enables experimental tumor initiation and metastatic colonization. Cancer cell. 2013; 24:481–498. [PubMed: 24035453]
11. Shibue T, Brooks MW, Inan MF, Reinhardt F, Weinberg RA. The outgrowth of micrometastases is enabled by the formation of filopodium-like protrusions. Cancer Discov. 2012; 2:706–721. [PubMed: 22609699]
12. Ghajar CM, et al. The perivascular niche regulates breast tumour dormancy. Nature cell biology. 2013; 15:807–817. [PubMed: 23728425]
13. van den Bogaard EH, Dailey LA, Thorley AJ, Tetley TD, Forbes B. Inflammatory response and barrier properties of a new alveolar type 1-like cell line (TT1). Pharm Res. 2009; 26:1172–1180. [PubMed: 19199008]
14. Bohinski RJ, Huffman JA, Whitsett JA, Lattier DL. Cis-active elements controlling lung cell-specific expression of human pulmonary surfactant protein B gene. J Biol Chem. 1993; 268:11160–11166. [PubMed: 8496176]
15. Kemp SJ, et al. Immortalization of human alveolar epithelial cells to investigate nanoparticle uptake. Am J Respir Cell Mol Biol. 2008; 39:591–597. [PubMed: 18539954]
16. Seguin L, Desgrosellier JS, Weis SM, Cheresh DA. Integrins and cancer: regulators of cancer stemness, metastasis, and drug resistance. Trends Cell Biol. 2015; 25:234–240. [PubMed: 25572304]
17. El Touny LH, et al. Combined SFK/MEK inhibition prevents metastatic outgrowth of dormant tumor cells. J Clin Invest. 2014; 124:156–168. [PubMed: 24316974]
18. Bragado P, et al. TGF-beta2 dictates disseminated tumour cell fate in target organs through TGF-beta-RIII and p38alpha/beta signalling. Nature cell biology. 2013; 15:1351–1361. [PubMed: 24161934]
19. Gawrzak S, et al. MSK1 regulates luminal cell differentiation and metastatic dormancy in ER(+) breast cancer. Nature cell biology. 2018; 20:211–221. [PubMed: 29358704]
20. Gao H, et al. The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. Cell. 2012; 150:764–779. [PubMed: 22901808]
21. Kobayashi A, et al. Bone morphogenetic protein 7 in dormancy and metastasis of prostate cancer stem-like cells in bone. J Exp Med. 2011; 208:2641–2655. [PubMed: 22124112]
22. Cackowski FC, et al. Mer Tyrosine Kinase Regulates Disseminated Prostate Cancer Cellular Dormancy. J Cell Biochem. 2017; 118:891–902. [PubMed: 27753136]
23. Fluegen G, et al. Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic microenvironments. Nature cell biology. 2017; 19:120–132. [PubMed: 28114271]
24. Sosa MS, et al. NR2F1 controls tumour cell dormancy via SOX9- and RARbeta-driven quiescence programmes. Nat Commun. 2015; 6:6170. [PubMed: 25636082]

Europe PMC Funders Author Manuscripts

25. Malanchi I, et al. Interactions between cancer stem cells and their niche govern metastatic colonization. Nature. 2011; 481:85–89. [PubMed: 22158103]
26. Oskarsson T, et al. Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. Nat Med. 2011; 17:867–874. [PubMed: 21706029]
27. Zhang XH, Giuliano M, Trivedi MV, Schiff R, Osborne CK. Metastasis dormancy in estrogen receptor-positive breast cancer. Clin Cancer Res. 2013; 19:6389–6397. [PubMed: 24298069]
28. Kim RS, et al. Dormancy signatures and metastasis in estrogen receptor positive and negative breast cancer. PLoS One. 2012; 7:e35569. [PubMed: 22530051]
29. Cheng Q, et al. A signature of epithelial-mesenchymal plasticity and stromal activation in primary tumor modulates late recurrence in breast cancer independent of disease subtype. Breast Cancer Res. 2014; 16:407. [PubMed: 25060555]
30. Lee JL, Lin CT, Chueh LL, Chang CJ. Autocrine/paracrine secreted Frizzled-related protein 2 induces cellular resistance to apoptosis: a possible mechanism of mammary tumorigenesis. J Biol Chem. 2004; 279:14602–14609. [PubMed: 14709558]
31. Bovolenta P, Esteve P, Ruiz JM, Cisneros E, Lopez-Rios J. Beyond Wnt inhibition: new functions of secreted Frizzled-related proteins in development and disease. J Cell Sci. 2008; 121:737–746. [PubMed: 18322270]
32. Kaur A, et al. sFRP2 in the aged microenvironment drives melanoma metastasis and therapy resistance. Nature. 2016; 532:250–254. [PubMed: 27042933]
33. Liberzon A, et al. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011; 27:1739–1740. [PubMed: 21546393]
34. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One. 2010; 5:e13984. [PubMed: 21085593]
35. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010; 26:589–595. [PubMed: 20080505]
36. Behbehani GK, Bendall SC, Clutter MR, Fantl WJ, Nolan GP. Single-cell mass cytometry adapted to measurements of the cell cycle. Cytometry A. 2012; 81:552–566. [PubMed: 22693166]
37. Zunder ER, et al. Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. Nat Protoc. 2015; 10:316–333. [PubMed: 25612231]
38. Finck R, et al. Normalization of mass cytometry data with bead standards. Cytometry A. 2013; 83:483–494. [PubMed: 23512433]
39. van Dijk D, et al. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. Cell. 2018; 174:716–729. e727 [PubMed: 29961576]
40. Pocaterra A, et al. F-actin dynamics regulates mammalian organ growth and cell fate maintenance. J Hepatol. 2019; 71:130–142. [PubMed: 30878582]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

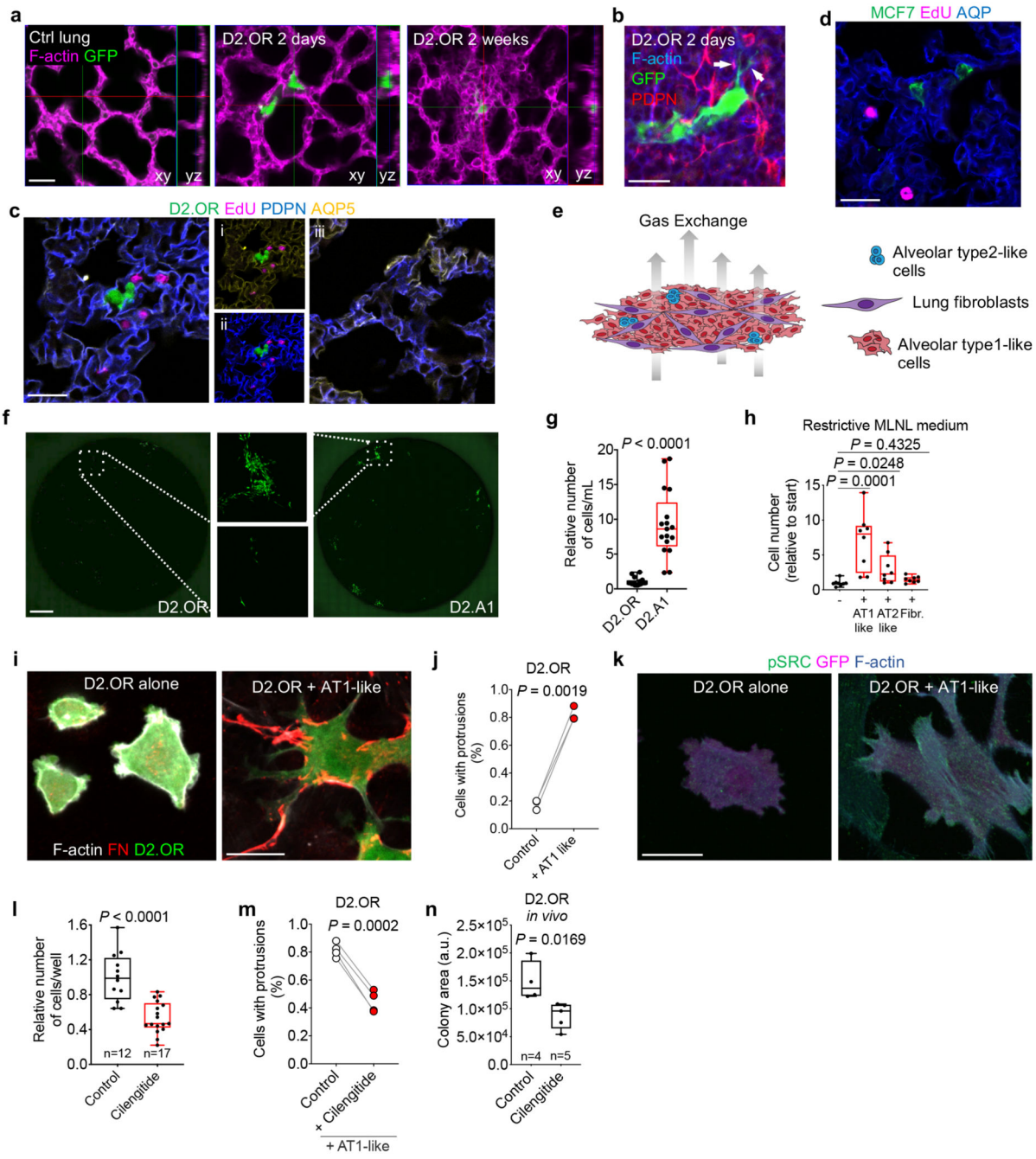


Figure 1. Alveolar type1 cells (AT1) regulate behavior of disseminated indolent breast cancer cells.

a, Fluorescent *in situ* images of the lung alveolar space in control and D2.0R-EGFP-injected mice. n=3. Scale bar, 20μm. **b**, Fluorescent IHC of AT1 cells (PDPN+ve) and lung-disseminated D2.0R cells (GFP) upon tail vein injection. Arrows indicate protrusions formed by D2.0R cells. n=5. Scale bar, 20μm. **c**, Fluorescent IHC of D2.0R cells in the lungs two weeks after intravenous injection showing surrounding proliferating (EdU+) mature AT1 cells (PDPN+/AQP5+). i and ii, separate staining for PDPN and AQP5. iii, control

uninjected lung. n=3. Scale bar, 20μm. **d**, Disseminated MCF7 cells in the lung showing similar pattern of surrounding proliferating mature AT1 cells. n=3. Scale bar, 20μm. **e**, Schematic of the lung organotypic system. Validation of marker expression in Extended Data Fig. 2a **f**, Representative (n=6 independent experiments) immunofluorescence of GFP+ breast cancer cells co-cultured with lung stromal cells. Dashed squares highlight indolent, scattered D2.0R cells and active proliferating colonies of D2.A1 cells. Scale bar, 2mm. **g**, Quantification of breast cancer cells in the co-culture after 5 days. Mean normalized pooled samples (n=18) from independent experiments (n=6). Mann-Whitney test. **h**, Quantification of D2.0R cells co-cultured with individual lung stromal cells after 5 days in MLNL medium. Pooled samples (n=8) from independent experiments (n=2). Dunn's multiple comparisons test. **i**, Immunofluorescence of D2.0R cells cultured alone (left) or co-cultured with AT1-like cells (right). FN, Fibronectin. n=5. Scale bar, 20μm. **j**, Percentage of D2.0R cells with protrusions alone or in coculture with AT1-like cells. Means from n=3 independent experiments. Paired t-test. **k**, Immunofluorescence of D2.0R cells cultured with or without AT1-like cells. pSRC, Phospho-SRC. n=2. Scale bar, 20μm. **l**, Relative number of D2.0R cells after 5 days of treatment with cilengitide. Mean normalized pooled samples from independent experiments (n=3). Mann-Whitney. **m**, Percentage of D2.0R cells with protrusions upon coculture with AT1-like cells and treatment with cilengitide. Means from n=4 independent experiments. Paired t-test. **n**, Quantification of disseminated D2.0R-EGFP cells after tail vein injection in BALB/C nude mice and treatment with cilengitide. Unpaired t-test. **g**, **h**, **l**, **n** plots show data as whisker plots: midline, median; box, 25–75th percentile; whisker, minimum to maximum.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

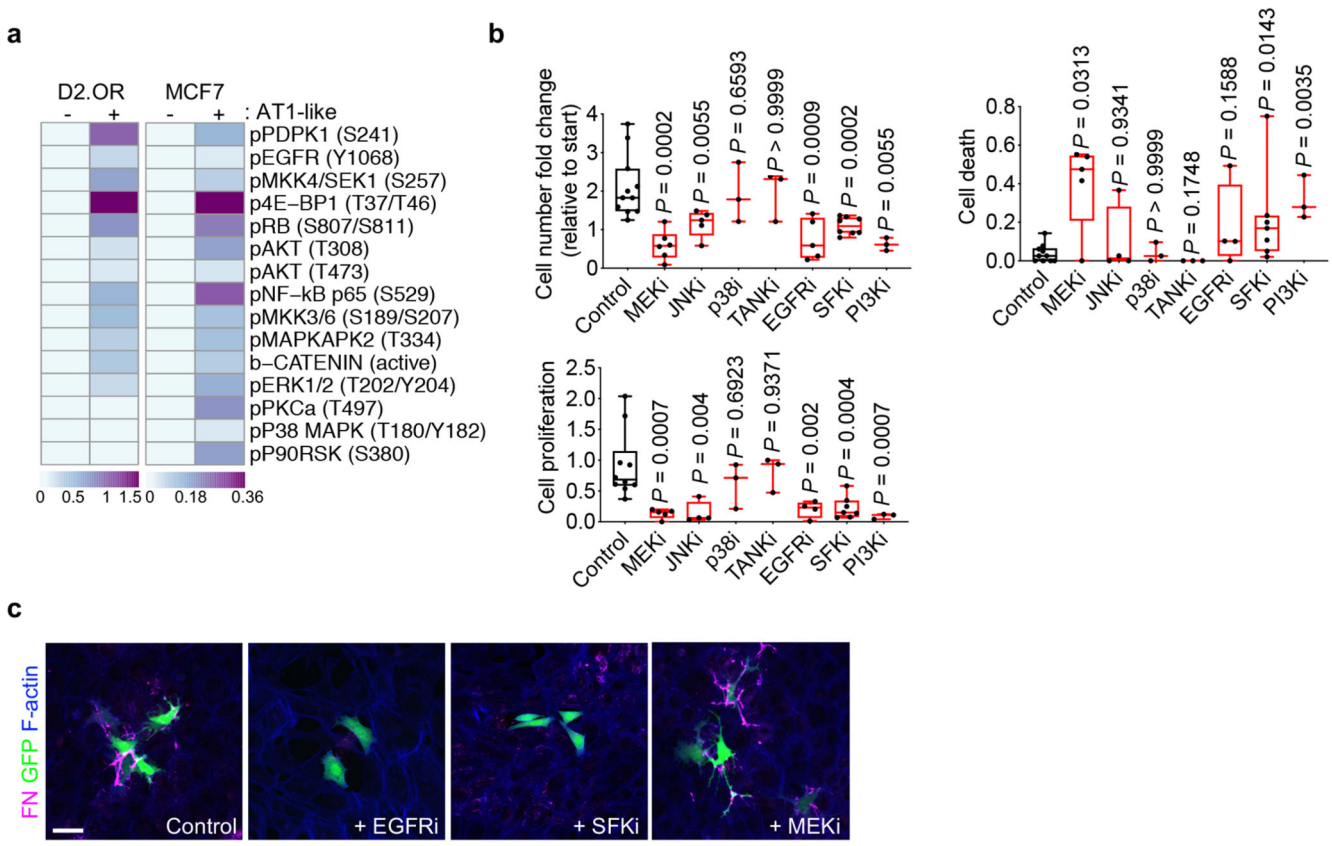


Figure 2. Mass cytometry analysis reveals signaling pathways involved in the crosstalk between AT1 and indolent breast cancer cells.

a, Heatmaps of EMD values (Earth Mover's Distance) estimating the activation of the indicated molecules in D2.0R or MCF7 alone or in coculture with AT1-like cells. Representative of three independent experiments. **b**, Plots show cell number fold change and relative frequency (number of events/starting number of D2.0R cells) of apoptotic and mitotic events in D2.0R cells determined from movies of D2.0R cells cultured with AT1 cells in MLNL media in the presence of inhibitors of the indicated targets. Each data point represents mean of an independent experiment. Cell number: n=11 (Control), 8 (for SFKi), 6 (for MEKi), 5 (for JNKi and EGFRi), 3 (for p38i, TANKi and PI3Ki). Cell death and cell proliferation: n=10 (Control), 7 (for SFKi), 5 (for MEKi), 4 (for JNKi and EGFRi), 3 (for p38i, TANKi and PI3Ki). Mann-Whiney test. Data are presented as whisker plots: midline, median; box, 25–75th percentile; whisker, minimum to maximum. **d**, Images show F-actin and fibronectin (FN) staining of D2.0R-EGFP cells co-cultured +/- AT1-like cells in MLNL medium with Dasatinib (SFKi), Lapatinib (EGFRi) or PD184352 (MERKi) for 48hrs. Similar results were obtained with an additional SFKi (AZD0530). n=3. Scale bar is 20μm.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

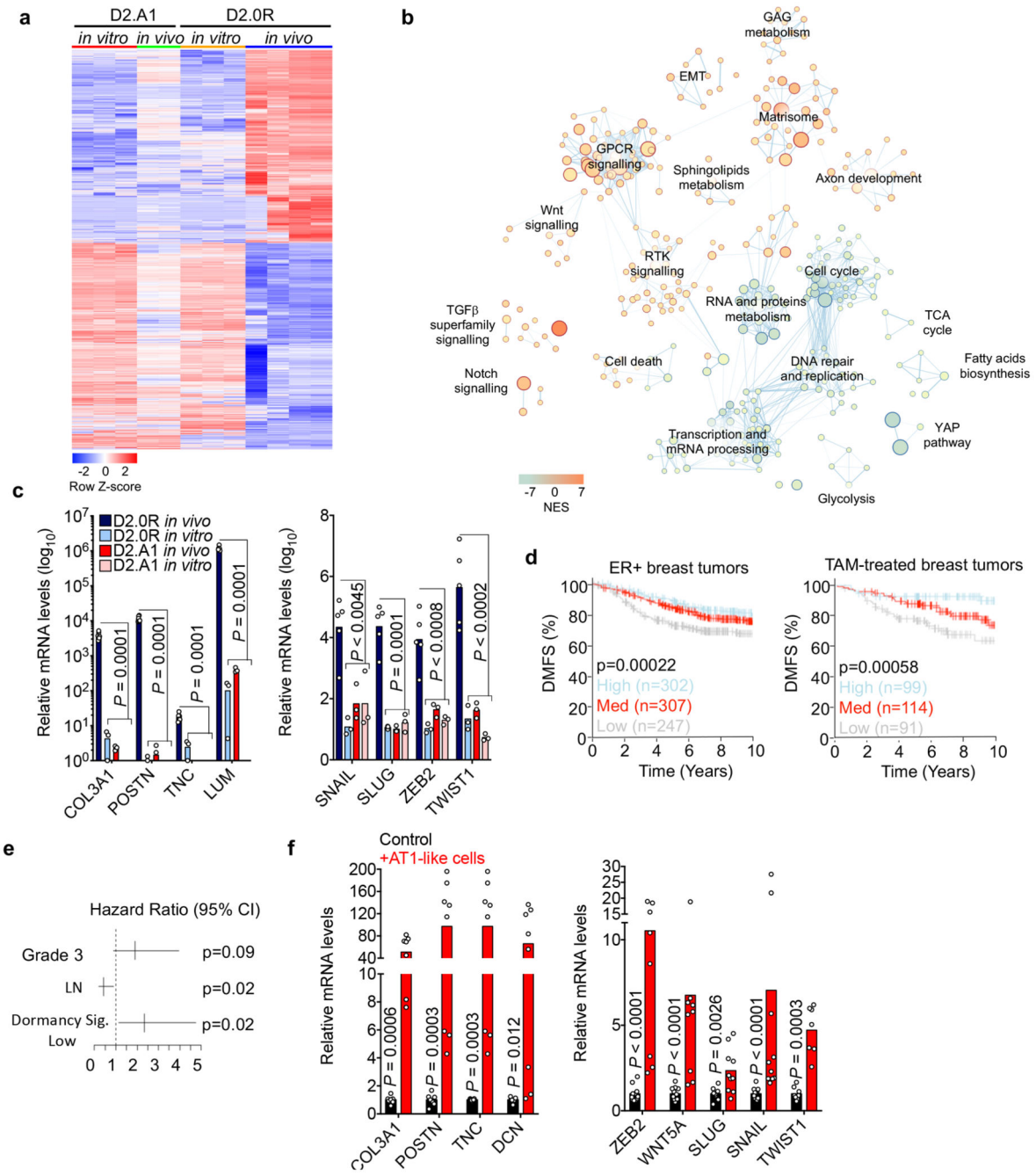


Figure 3. Gene expression analysis of lung-disseminated indolent breast cancer cells *in vivo*.

a, D2.0R-EGFP cells or D2.A1-EGFP cells were injected intravenously in mice and recovered from lungs after 3 weeks. Cells were then processed for RNA sequencing. Heatmap shows normalized expression data for genes that were differentially regulated in the D2.0R *in vivo* compared to D2.0R *in vitro*, D2.A1 *in vivo* and *in vitro*. **b**, Enrichment map for disseminated indolent breast cancer cells *in vivo*. The map shows gene-set enrichment results of D2.0R cells *in vivo* compared to the other groups. Node size, genes in pathway; node color, enrichment score (orange indicates enrichment in D2.0R *in vivo*, blue

indicates enrichment in the other groups combined); edge width, overlap size between connected nodes. **c**, qPCR analysis of selected genes from independent *in vitro* and *in vivo* samples (n=3 mice or wells for D2.A1 *in vivo* and *in vitro* and D2.0R *in vitro*; n=6 mice for D2.0R *in vivo*). Selected genes belong to two processes (extracellular matrix proteins, ECM, and epithelial to mesenchymal transition, EMT) identified in the gene-set enrichment analysis (GSEA). One-way ANOVA test. **d**, Kaplan-Meier curves showing distant metastasis free survival (DMSF) of patients derived from http://co.bmc.lu.se/gobo/gsa.pl database stratified according to the dormancy signature. Left plot displays ER+ breast cancer patients, right plot displays patients that have undergone treatment with tamoxifen. **e**, Plot shows Cox multivariate analysis of stage, lymph node status and dormancy signature (split in high and low groups) in tamoxifen-treated breast cancer patients (n=142). Cox proportional hazard regression model. **f**, qPCR analysis for ECM and EMT genes of D2.0R-EGFP cells cultured alone or together with AT1-like cells for 4 days in MLNL medium. Mean normalized pooled samples (n=7 or 8 samples from 3 independent experiments for left plot; n=11 samples from 4 independent experiments for right plot, except *TWIST1* data that was generated from n=8 samples from 3 independent experiments). Mann-Whitney test.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

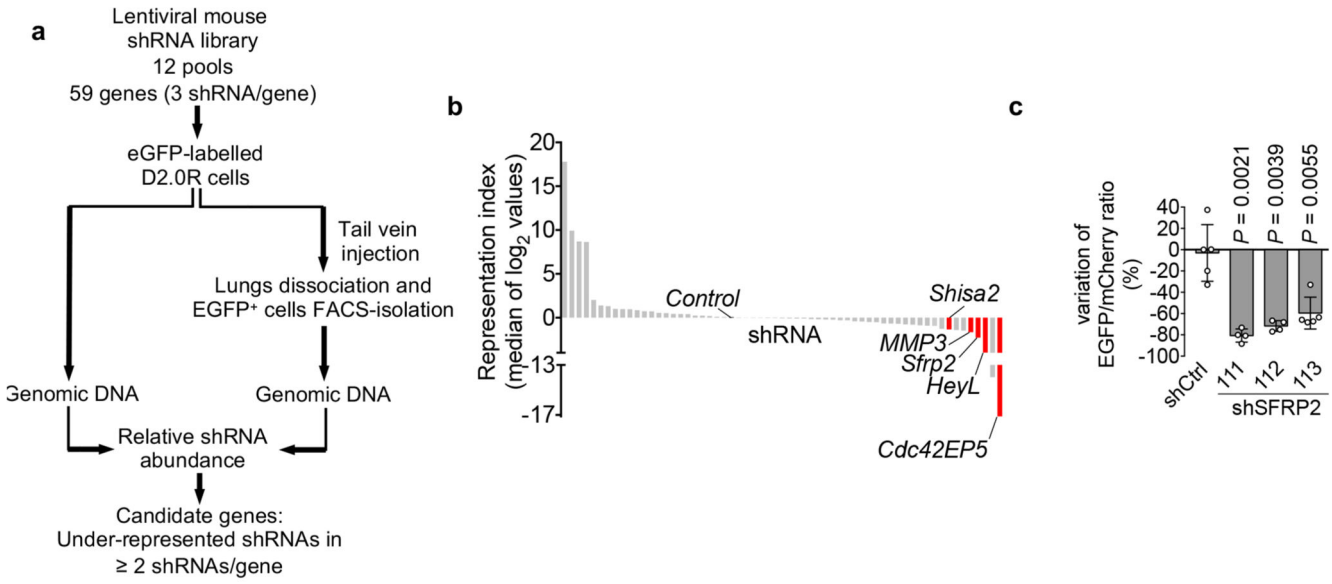


Figure 4. A loss-of-function screen *in vivo* identifies SFRP2 as survival regulator in lung disseminated indolent breast cancer cells.

a, Schematic showing the screening strategy *in vivo*. **b**, Histogram of representation scores for each gene calculated from the fold-change of representation of each shRNA relative to pre-injection abundance. On the left side of the plot there are genes whose knock-down led to increased proliferation; on the right side of the plot there are genes that, once downregulated, led to reduced representation of the clones. Red bars highlight genes that were selected for further validation. **c**, D2.0R-EGFP-shSfrp2 or -shControl cells (3 independent shRNA sequences) were injected with an equal amount of D2.0R-mCherry-shControl cells intravenously (ratio=1). After 3 weeks, breast cancer cells were isolated and the ratio EGFP+-cells/mCherry+-cells calculated (n=5 mice for Control group, n=4 mice for the other groups). Unpaired t-test with Welch's correction.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

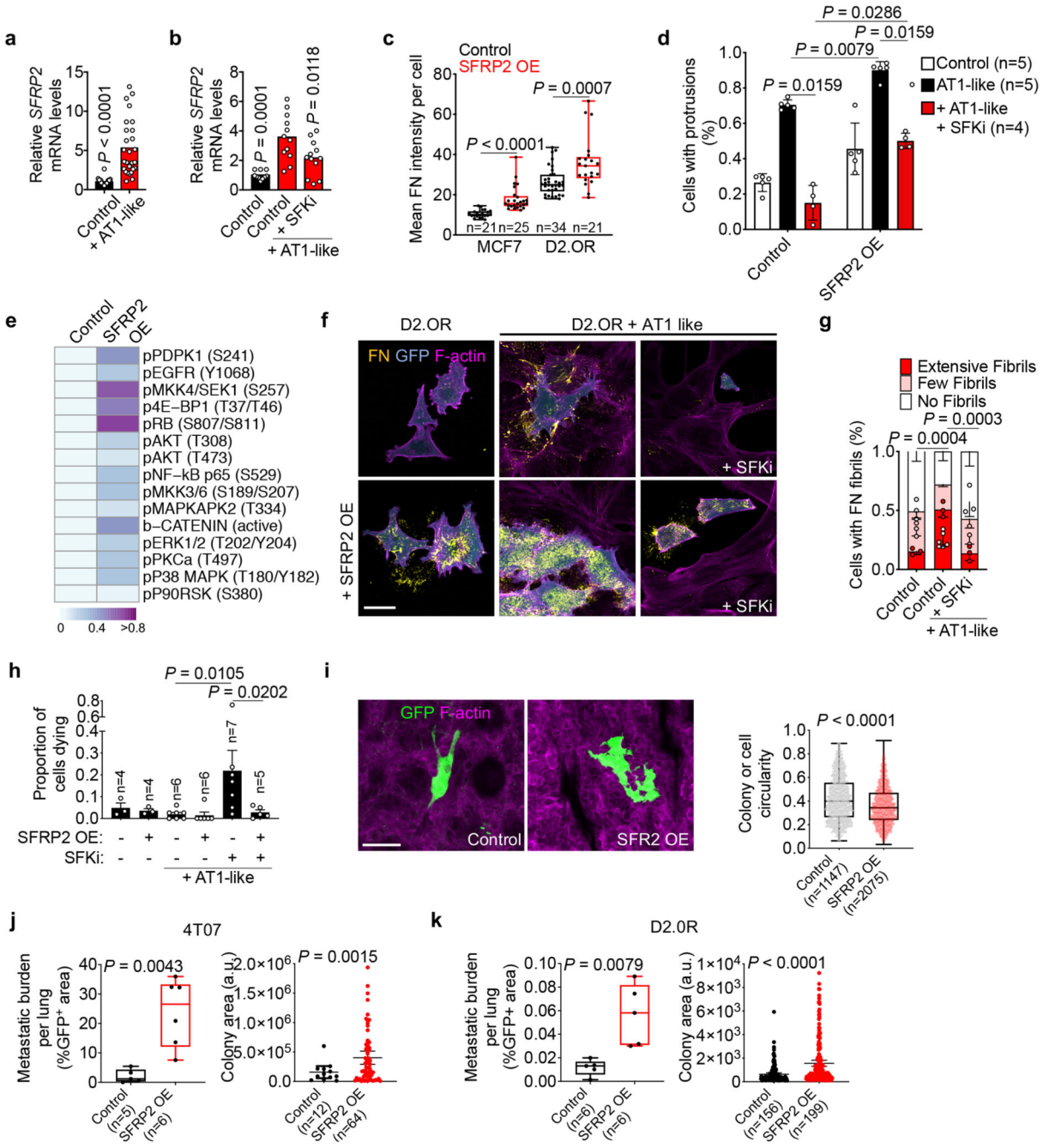


Figure 5. SFRP2 regulates Src-mediated fibrillogenesis, protrusions and survival.

a, qPCR for *SFRP2* in D2.0R-EGFP cells cultivated alone or with AT1-like cells in MLNL medium for 4 days. Mean normalized pooled samples (n=24 for Control, n=27 for AT1-like group) from 7 independent experiments. Mann-Whitney test. **b**, as in **a**, in addition cells were treated with the SFK-inhibitor (SFKi, 50nM) or DMSO. Mean normalized pooled samples (n=10 for Control, n=12 for the other groups) from 3 independent experiments. One-way ANOVA test. **c**, Mean Fibronectin (FN) intensity per cell in control and SFRP2 overexpressing (OE) breast cancer cells. Mann-Whitney test. Pooled data from n=2

independent experiments. **d**, Protrusions in Control or SFRP2-OE D2.0R-EGFP cells alone or with AT1-like cells in presence or not of SFK-inhibitor. Mean values of independent experiments. Mann-Whitney test. **e**, Heatmaps of EMD values in Control or SFRP2 OE D2.0R cells. n=1 experiment. **f**, Immunofluorescence for F-actin and FN in Control or SFRP2-OE- D2.0R-EGFP cells alone or with AT1-like cells in MLNL medium with SFKi for 48hrs. n=3. Scale bar is 20μm. **g**, SFRP2-OE-D2.0R-EGFP cells alone or in coculture with AT1-like cells have been treated with SFK inhibitor for 48hrs. Quantification as in Extended Data Figure 5b. n=3 experiments. **h**, Quantification of cell death in Control or SFRP2-OE D2.0R cells, -/+ SFK-inhibition (48h). Quantification as in Figure 2b. Mean and S.E.M. are shown (n=3-7 independent experiments). Mann-Whitney test. **i**, Left, fluorescent *in situ* images of Control or SFRP2-OE- D2.0R-EGFP cells in the lung alveolar space after tail vein injection. Scale bar, 20μm. Right, circularity of lung-disseminated Control or SFRP2-OE- D2.0R-EGFP cells (across 4 mice/group). Mann-Whitney test. **j** and **k**, Quantification of metastatic burden and metastatic colony area 2 weeks after intravenous injection of Control or SFRP2-OE 4T07 or D2.0R cells, respectively, into wt or nude mice, respectively. Mann-Whitney test for metastatic burden. Unpaired t-test with Welch's correction for colony area experiments. **c**, **i** (right), **j** and **k** (left) plots show data as whisker plots: midline, median; box, 25–75th percentile; whisker, minimum to maximum. **j** (right), **k** (right) plots are mean with 95% confidence interval.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts