Cross-talk with lung epithelial cells regulates *Sfrp2* expression enabling disseminated breast cancer cell latency

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

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

Abstract

The process of metastasis is highly complex[1]. In the case of breast cancer, there are frequently long timespans between cells leaving the primary tumour and growth of overt metastases[2, 3]. Possible 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. In this work, 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 crosstalk 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.

Main

To analyse indolent breast cancer, we utilised the D2.OR/D2.A1 model[7-9] (Supplementary Figure 1a). As expected, D2.OR cells persisted for many weeks in the lungs (Figure 1a and Supplementary Figure 1b), but did not form large colonies, whereas D2.A1 cells grew aggressively (Supplementary Figure 1b). The indolent behaviour of the D2.OR cells parallels that observed clinically in ESR1+ve breast cancer; consistent with this, D2.OR cells express ESR1 *in vivo* and respond to

estradiol[10]. D2.OR cells were similarly indolent in both Balb/C and Balb/C nude mice indicating that their phenotype is not due to the adaptive immune system (Supplementary Figure 1c). Closer examination revealed that D2.OR cells had extravasated into the alveolar space and were in close contact with the lung parenchyma after two days, and that they remained in this location at least for two weeks (Figure 1a). In this context, both D2.OR and D2.A1 cells formed long extensions reminiscent of filopodia-like protrusions observed by other researchers (Figure 1b and Supplementary Figure 1d)[11, 12]. The formation of protrusions increased the cell perimeter relative to the cell area and this is reflected in a low circularity index of ~0.4 (1 = perfect circle). Immunostaining demonstrated that D2.OR invariably had close contact with AQP5+ve and PDPN+ve alveolar type I (AT-1) cells (Figure 1c, left). Frequent contacts were also observed with SFTPC+ve alveolar type II (AT2) cells and MUC1+ve endothelial cells, which is consistent with previous reports[13] (Supplementary Figures 1e, f). EdU pulse labelling revealed that AT1 cells, which are normally quiescent, were proliferating proximal to D2.OR cells at both 3 days and 14 days after arrival in the lungs, with the greatest proliferation at the earlier time-point. This suggests that the expansion of the lung parenchyma around indolent metastases visible in Figure 1a at two weeks results mostly from proliferation of AT1 cells (Figure 1c and Supplementary Figures 1g, h). Similar contacts with PDPN+ve alveolar type I cells and an increase in EdU positivity were observed with the few human MCF7 cells that persisted 3 days following tail vein injection (Figures 1d and Supplementary Figure 1i). These data uncover a proliferative response in AT1 cells to the arrival of indolent breast cancer cells, and suggest that interplay between indolent breast cancer cells and the lung parenchyma may influence metastatic behaviour.

To date, research into the metastatic microenvironment has focused on leukocytes, fibroblasts, and endothelial cells with little attention given to epithelial cells, which are a major component of lung tissue[1, 13]. To study how breast cancer cells might interact in the lung environment, we established a co-culture system that replicated key features of the lung and could recapitulate the indolent behaviour of D2.OR cells. To this end, we co-cultured lung epithelial cells that express the key markers of AT1 and AT2 cells and fibroblasts on a gas permeable substrate in Mitogen Low Glucose Low (MLNL medium)[14-16] (schematic illustration in Figure 1e - with validation of cell type specific marker expression in Supplementary Figure 1j). Strikingly, the addition of low numbers of either D2.OR or D2.A1 cells to these co-cultures recapitulated the indolent and aggressive growth of D2.OR and D2.A1 cells observed *in vivo*, respectively (Figures 1f, g and increased Ki67+ cells shown in Supplementary Figure 1k). These differences could not be attributed to intrinsic differences in growth rates between D2.OR and D2.A1 cells in either MHNH medium or MLNL medium (Supplementary Figure 1l). Crucially, the indolent behaviour of D2.OR cells in the co-culture was reversible if cells were subsequently returned to conventional cell culture conditions, further reinforcing the similarities with long latency metastatic recurrence *in vivo* (Supplementary Figure 1m). We next explored the effect of individual cell types within the co-culture assay on D2.OR cells, something that is not possible in mice as elimination of key cell lineages in the lung is not compatible with life. Somewhat unexpectedly, co-cultures with individual cell types in MLNL media indicated that AT1-like cells were able to boost D2.OR numbers, with AT2-like cells having a smaller positive effect (Figure 1h). Similar results were obtained with 4T07 cells, an additional murine model for indolent metastases, and human MCF7 cells (Supplementary Figure 1n). Time-lapse imaging revealed that AT1-like cells both

suppress apoptosis and increase the mitotic rate of D2.OR cells (Supplementary Figure 1o). 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 (Supplementary Figure 1p). Notably, 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 (Figure 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.OR cells in MHNH media for 7 days. Supplementary Figure 1q shows that AT1-like cells were able to reduce the growth of D2.OR cells in favourable conditions. Together, these data suggest a complex relationship between lung epithelial cells and breast cancer micro-metastases. In particular, both pro-survival and growth restrictive signals from AT1 cells likely coexist *in vivo* and *in vitro,* which can be highlighted by modulating the experimental conditions. We reasoned that the greatest eventual 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 provided supportive signals to D2.OR cells (MLNL media), we noted that co-culture with AT1-like cells induced the formation of long protrusions similar to those observed *in vivo* (Figures 1i, j). This change in morphology reduced the circularity of D2.OR cells from ~0.8 to 0.2-0.4 (Supplementary Figure 1r), further the protrusions were positive for active p-Src, which has been extensively implicated in pro-metastatic signals[9, 11, 17](Figure 1k). The protrusions formed by D2.OR cells in the presence of AT1-like cells were associated

with prominent fibronectin fibrils (Figure 1i). Similar increases in cell protrusion were obtained with 4T07, MCF7, and T47D-DBM cells (Supplementary Figures 1s-u), but not for the aggressive cell line D2.A1 that had a higher baseline of protrusions (Supplementary Figure 1v). Blockade of integrins using cilengitide, which mimics the RGD integrin binding motif of fibronectin and other ECM molecules, reduced both protrusions and the numbers of D2.OR and MCF7 cells (Figures 1l, m and Supplementary Figure 1w). Most importantly, treatment of mice with cilengitide even after cells had already seeded the lungs reduced the number of metastases (Figure 1n). These data suggest that persistence of indolent breast cancer cells at the metastatic site and the induction of cellular protrusions by AT1 cells might represent intertwined aspects of metastatic dissemination. Further, they demonstrate that targeting this axis is a viable strategy for the elimination of indolent breast cancer cells.

To understand better the signalling pathways involved in breast cancer-alveolar cell crosstalk, we undertook mass cytometry analysis of co-cultures using a panel of metal labelled antibodies. D2.OR or MCF7 and AT1-like cells were either cultured alone or co-cultured in MLNL, fixed, dissociated into single cells, and stained in suspension with a panel of antibodies covering for a broad array of proteins involved in signalling and proliferation (Figure 2a). The identity of breast cancer cells and AT1-like clusters in the co-culture could be inferred from the mono-cultures and was confirmed by the GFP signal (for D2.OR and MCF7 cells). Consistent with data in Supplementary Figure 1o, there was an increase in phosphorylation events associated with proliferation (S807/811-pRb and T37/46-p4E-BP1) in D2.OR and MCF7 upon co-culture (Figure 2b). Further, mouse and human indolent breast cancer cells activate a similar intracellular response upon contact with AT1-like cells (Figure 2b), with prominent increases in ERK, MKK4, MKK3/6, PDPK1, β-catenin, and NFkB signalling.

Conversely, both D2.OR and MCF7 breast cancer cells tested triggered the same proliferative response in AT1-like cells (S807/811-pRb and T37/46-p4E-BP1), together with the increase in S28 phosphorylation of Histone H3, a marker of cellular proliferation, supporting the *in vivo* observation of EdU+ve nuclei around micrometastases (Supplementary Figures 2a, b). More in depth analysis using conditional Density Resampled Estimation of Mutual Information (DREMI) analysis, which generates a score that reflects the linkage of signalling between the two variables analysed, revealed increased connectivity from PDPK1 to PKCα and AKT and from AKT to 4E-BP1 (Supplementary Figure 2c – DREMI score in white). Several of these pathways have been linked to the dormant phenotype[3, 18-20], we therefore investigated how pharmacological inhibition of these and other prominent signalling pathways affected D2.OR 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.OR cells when co-cultured with AT1-like cells (Figure 2c). Notably, EGFR, MEK, and SFK inhibition both increased apoptosis and reduced mitotic events without greatly affecting D2.OR and AT1-like cells in monoculture (Figure 2c and Supplementary Figure 2d). Combining inhibitors with phosho-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 (Supplementary Figure 2e). The importance of ERK/MAP kinase activation was confirmed by a reduction in metastatic colony size *in vivo* (Supplementary Figure 2f), further supporting the concept of targeting signalling between AT-1 cells and breast cancer cells to eliminate indolent micro-metastases.

We next investigated if there was a relationship between the signalling pathways required for boosting proliferation and cell survival and the protrusions observed in indolent cells interacting with alveolar epithelial cells. Interestingly, the formation of cell protrusions and FN fibrils described in Figure 1 depends on EGFR and SFK signalling, but not on MEK signalling, potentially indicating a bifurcation in the signalling cascade at a point downstream of SFK (Figure 2d, Supplementary Figures 2g-l and Supplementary Figure 4o). The reduced FN staining was not correlated with reduced FN transcription (Supplementary Figure 2j). Of note, EGFR upstream signalling was required for Src activation (Supplementary Figure 2k). These data reinforce the correlation between cell protrusions and signals that boost survival of indolent breast cancer cells when cultured with lung epithelial cells.

In a parallel effort to better understand the biology of indolent breast cancer cells, we investigated how the metastatic microenvironment alters cancer cell gene expression *in vivo*. We isolated D2.OR cells from lungs 3 weeks after injection in mice and compared their transcriptional profile with D2.A1 cells isolated from the lungs and both cell types grown in culture. Non-hierarchical clustering analysis revealed that the D2.OR cells isolated from the lungs had very distinctive transcriptomes from lung isolated D2.A1 cells and both cell types grown *in vitro* (Figure 3a). The expression of cell cycle and DNA replication genes was dramatically reduced and, consistent with previous reports, we observed up-regulation of Bmp signalling[21, 22] and the dormancy-associated factors *Nr2f1* and *Sharp1* (also known as *Dec2*)[19, 23-25] (Figure 3b and Supplementary Figures 3a, b). Of more interest, we 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[26, 27]. QRT-PCR

analysis of selected genes from the RNAseq analysis provided independent corroboration of the sequencing data (Figure 3c). We next explored links between the transcriptome of indolent D2.OR cells and human breast cancer. Strikingly, a signature of genes highly expressed in indolent D2.OR cells *in vivo* compared to the other groups was clearly linked with improved distant metastasis free survival (DMFS) in human ER positive breast cancer, which is known to have long latency periods before relapse[28] (Figure 3d – genes listed in Supplementary Table 1). Patients receiving tamoxifen therapy with the D2.OR-derived (indolent) signature responded incredibly well to treatment. Of note, in these analyses the outcome between patients with high and low expression of our dormancy signature showed similar metastatic recurrence rates over the first two years. This was confirmed in another dataset of endocrine therapy treated patients (GSE9515) and re-plotting the analysis from two years onwards confirmed the signatures ability to indicate lower likelihood of distant relapses at prolonged time points (Supplementary Figure 1c). Conversely, those patients with low expression of the 'indolence' signature had a significantly increased hazard ratio of 2.5 (Figure 3e). Genes specifically up-regulated in D2.A1 cells or on plastic showed no link with outcomes (Supplementary Figure 3e and Supplementary Table 1). Of note, our signature does not contain genes overlapping with the dormancy score genes identified by Kim et al., and Cheng et al., (Supplementary Table 1[29, 30]) and performed as expected in publicly available databases used in the same publication (Supplementary Figure 3d). This clear link to human outcomes further reinforced 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.OR cells. Following our analysis in Figures 1 and 2, we hypothesized that these AT1-like cells might trigger the up-regulation of

genes in D2.OR cells in the lung. To explore this idea, we asked whether AT1-like cells could promote the expression of the genes identified as being associated with indolence *in vivo*. Figure 3f and Supplementary Figure 3f 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 as well as BMP and Wnt target genes. 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, we hypothesized that some would 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. Further, genes in this latter class might be implicated in the survival signals emanating from AT1 cells. To identify these genes we performed a functional screen *in vivo* using shRNA targeting genes up-regulated in D2.OR cells isolated from the lungs. The small number of cells that could be isolated from the lungs of mice 3 weeks post-injection (3000-10000) placed constraints on the complexity of the size of the library that could be screened. We therefore selected a subset of genes involved in cell-cell communication, cell signalling, the extra-cellular environment, and control of cell state for screening (Supplementary Table 2). We transduced D2.OR-EGFP cells with a MOI optimised for a single shRNA per cell. Three shRNA per gene were used against 59 genes. Sub-pools of the shRNA library were prepared and injected into the tail vein of mice in triplicate. Concomitant with this, reference DNA was prepared from the sub-pools before injection. After 3 weeks, D2.OR cells were isolated from the lungs, their DNA sequenced and the relative representation of each shRNA compared with reference DNA representing the initial composition of the library (workflow illustrated in Figure

4a). Figure 4b shows the relative representation of shRNAs of each single gene. Consistent with our original hypothesis 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 (Figure 4b). 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. A second screen was carried out on the best hits in this category showing a consistent effect with at least 2 out of 3 interfering sequences: *Cdc42ep5*, *Sfrp2*, *Heyl*, *Mmp3* and *Shisa2* (Figure 4b and Supplementary Figures 4a, b). Cells containing shRNA against the putative hits were labelled with GFP, control cells were labelled with mCherry and co-injected into the same mice. Supplementary Figure 4b shows that the effect of *Sfrp2*, *Heyl*, *Shisa2* and *Cdc42ep5* was confirmed when shRNA-transduced cells were injected independently. *Cdc42ep5*[31] was not pursued as there is already extensive literature implicating cytoskeletal genes in the process of extravasation, which is not the focus of this study. We instead 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[32, 33]. Further, it has been previously linked with survival and crosstalk between cancer cells and stroma[34]. Figure 4c confirms that multiple independent shRNAs against SFRP2 all reduced metastatic burden. *Sfrp2* depletion did not affect the initial arrest and extravasation of D2.OR cells as equal numbers of control and depleted cells were observed in the lungs 72hrs after intravenous injection (Supplementary Figure 4c). Loss of *Sfrp2* expression did not affect proliferation *in vitro* (Supplementary Figure 4d).

Having established the importance of *SFRP2 in vivo*, we considered whether *Sfrp2* might be regulated by crosstalk between breast cancer cells and the lung

epithelium. While *Sfrp2* was expressed at low levels in cell culture and primary tumours, its levels dramatically increased when in the lung environment (Supplementary Figure 4e). Co-culture experiments demonstrated that AT1-like lung epithelial cells could induce *Sfrp2* in D2.OR cells in a Src-dependent manner (Figures 4d, e), thus providing a potential explanation for the effect of SFK inhibitor observed in Figure 2. A broader analysis revealed that AT1-like cells also partially induced other SFRP family members in D2.OR cells and 4T07 cells (Supplementary Figure 4f). 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, indicating that Wnt signalling is likely not involved in the observed phenotype (Supplementary Figure 4g and effects of Tankyrase inhibitor in Figure 2c). It has been previously reported that SFRP2 binds FN and is incorporated into an insoluble extracellular matrix fraction [32]. Further, heparin binds the C-terminus of SFRP family proteins releasing them from the ECM leading to their inactivation [32]. We confirmed that heparin could increase the level of soluble inactive SFRP2 in the media (Supplementary Figure 4h). Notably, this treatment was also associated with reduced D2.OR cells numbers when co-cultured with AT1-like cells (Supplementary Figure 4i - it should be noted that heparin's anti-coagulation function is not relevant in this *in vitro* assay). These data support a model in which insoluble extracellular SFRP2 promotes cell numbers by increasing the deposition and organisation of FN (Figure 4f). In D2.OR cells over-expressing SFRP2, the FN was organised into fibrils (Figure 4i, bottom) and was correlated with increased numbers of cell protrusions (Figure 4g). This increase in protrusions was further enhanced by co-culture with AT1-like cells (Figure 4g and Supplementary Figure 4j). To obtain a more comprehensive molecular overview of how SFRP2 might boost D2.OR cells, we returned to CyTOF analysis of cell signalling.

We observed a striking overlap in the action of intracellular pathways between D2.OR cells co-cultured with AT1-like-cells and D2.OR cells overexpressing SFRP2 (Figure 4h and Figure 2b), including pPDPK1, pMKK4, pMKK3/6, and pERK. These data, combined with the effect of SFKi on SFPR2 induction, prompted us to perform epistasis experiments. This revealed two things: first, SFRP2 over-expression reduced the ability of SFKi to block the formation of protrusions and FN fibrils (Figures 4g, i, j and Supplementary Figures 2g, h). Second, apoptosis in the presence of SFKi was reduced when SFRP2 was over-expressed (Figure 4k, Supplementary Figure 4k shows no effect on proliferation). Together with data in Figure 1, these analyses argue that SFRP2 supports D2.OR persistence through pro-survival integrin/FN signalling leading to enhanced output across a range of oncogenic signalling pathways.

To test further the importance of sFRP2 *in vivo*, we examined the effect of its over-expression in mouse and human indolent breast cancer cells. Consistent with our *in vitro* data, SFRP2 over-expressing cells had more protrusions than control cells *in vivo*, as assessed using the circularity metric to evaluate the cell perimeter relative to cell area (Figure 4l). SFRP2 over-expression increased the size of colonies observed both human and mouse models with a particularly pronounced increase in large metastases (area $>5x10^5 \mu m^2$) in the more aggressive 4T07 model (Figures 4m, n and Supplementary Figure 4l) without affecting proliferation *in vitro* in absence of AT1-like cells (Supplementary Figure 4m) nor proximity to other stromal cells (Supplementary Figure 4n). The data establish that, upon arrival in the lungs, indolent breast cancer cells engage in complex reciprocal signalling with lung epithelial cells resulting the induction of sFRP2 and enhancing cell survival (Supplementary Figure 4o). Targeting this signalling whilst retaining the growth suppressive signals within the lung environment represents an appealing approach to eliminating dormant cancer cells.

Delayed recurrence of latent disseminated cells is a relevant unmet clinical need. Our current knowledge of the dormant phenotype is mainly limited to the signals that drive metastatic outgrowth. Albeit important, this doesn't explain how disseminated cells survive for such a long time in a foreign environment and how cancers of epithelial origin integrate in a different epithelial tissue. This work argues that parenchymal epithelial cells constitute a critical and previously un-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 rapidly out-competed by the malignant cells in growing tumours. However, in the context of indolent micro-metastases, or during the first steps of colonisation, they remain abundant relative to the cancer cells and therefore can exert a greater influence on their behaviour. Crosstalk between heterogeneous epithelial cells commonly regulates cell competition and tissue homeostasis. However, mechanisms underpinning cell competition can be hijacked by transformed epithelial cells in the early stages of primary tumour formation[35]. Here we report that the crosstalk between lung parenchymal cells and breast cancer cells is a key determinant of their indolent behaviour. Interaction between indolent cancer cells and AT1 cells contributes to the induction of the dormant transcriptional program and provides microenvironmental signals that support the persistence of latent cells within the lung parenchyma. We describe the transcriptional profile of indolent disseminated breast cancer cells *in vivo* highlighting a complex landscape including metabolic rewiring, synthesis of ECM niche and activation of specific signalling pathways. Combined *in vivo* loss-of-function screening and a novel *in vitro* organotypic system identify *Sfrp2* as central mechanism boosting the formation of cell protrusions and enabling the long-term survival breast cancer cells in the lung microenvironment. We identify with time

lapse analysis and single cell mass cytometry that EGFR, MEK, PI3K and SFK pathways impact on latent cell proliferation, quiescence and death. It is interesting to note that EGFR signalling is crucial in both alveolar responses to damage[36,37] and our experimental model, possibly indicating that the arrival of metastatic cells in the lungs triggers a tissue damage response. Upon co-culture with lung epithelial cells, activated Src is enriched in protrusions of cancer cells and contributes to the transcriptional induction of *Sfrp2*. SFRP2 in turn, coordinates pericellular fibronectin fibrillogenesis that leads to activation of integrin and survival cues. Of note, while depletion of *Sfrp2* inhibits long term survival of disseminated breast cancer cells, increasing *Sfrp2* expression leads to more aggressive metastatic lesions, suggesting that survival mechanisms involved in the metastatic outgrowth might also be essential for persistence of indolent cells. 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. With our work we identify key mechanisms that foster persistence of indolent cells in a secondary organ, providing new possible targets for adjuvant therapies that aim at killing disseminated cells before their awakening.

References

1. Lambert, A.W., Pattabiraman, D.R. & Weinberg, R.A. Emerging Biological Principles of Metastasis. *Cell* **168**, 670-691 (2017).
2. Hedley, B.D. & Chambers, A.F. Tumor dormancy and metastasis. *Adv Cancer Res* **102**, 67-101 (2009).
3. Sosa, M.S., Bragado, P. & Aguirre-Ghiso, J.A. Mechanisms of disseminated cancer cell dormancy: an awakening field. *Nat Rev Cancer* **14**, 611-622 (2014).
4. Obenauf, A.C. & Massague, J. Surviving at a Distance: Organ-Specific Metastasis. *Trends Cancer* **1**, 76-91 (2015).
5. Malladi, S. *et al.* Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. *Cell* **165**, 45-60 (2016).
6. Avgustinova, A. *et al.* Tumour cell-derived Wnt7a recruits and activates fibroblasts to promote tumour aggressiveness. *Nat Commun* **7**, 10305 (2016).
7. Naumov, G.N. *et al.* Ineffectiveness of doxorubicin treatment on solitary dormant mammary carcinoma cells or late-developing metastases. *Breast Cancer Res Treat* **82**, 199-206 (2003).
8. Naumov, G.N. *et al.* Persistence of solitary mammary carcinoma cells in a secondary site: a possible contributor to dormancy. *Cancer Res* **62**, 2162-2168 (2002).
9. Barkan, D. *et al.* Inhibition of metastatic outgrowth from single dormant tumor cells by targeting the cytoskeleton. *Cancer Res* **68**, 6241-6250 (2008).
10. Sowder, M.E. & Johnson, R.W. Enrichment and detection of bone disseminated tumor cells in models of low tumor burden. *Sci Rep* **8**, 14299 (2018).
11. Shibue, T., Brooks, M.W. & Weinberg, R.A. An integrin-linked machinery of cytoskeletal regulation that enables experimental tumor initiation and metastatic colonization. *Cancer cell* **24**, 481-498 (2013).
12. Shibue, T., Brooks, M.W., Inan, M.F., Reinhardt, F. & Weinberg, R.A. The outgrowth of micrometastases is enabled by the formation of filopodium-like protrusions. *Cancer Discov* **2**, 706-721 (2012).
13. Ghajar, C.M. *et al.* The perivascular niche regulates breast tumour dormancy. *Nature cell biology* **15**, 807-817 (2013).
14. van den Bogaard, E.H., Dailey, L.A., Thorley, A.J., Tetley, T.D. & Forbes, B. Inflammatory response and barrier properties of a new alveolar type 1-like cell line (TT1). *Pharm Res* **26**, 1172-1180 (2009).
15. Bohinski, R.J., Huffman, J.A., Whitsett, J.A. & Lattier, D.L. Cis-active elements controlling lung cell-specific expression of human pulmonary surfactant protein B gene. *J Biol Chem* **268**, 11160-11166 (1993).
16. Kemp, S.J. *et al.* Immortalization of human alveolar epithelial cells to investigate nanoparticle uptake. *Am J Respir Cell Mol Biol* **39**, 591-597 (2008).
17. Seguin, L., Desgrosellier, J.S., Weis, S.M. & Cheresh, D.A. Integrins and cancer: regulators of cancer stemness, metastasis, and drug resistance. *Trends Cell Biol* **25**, 234-240 (2015).
18. El Touny, L.H. *et al.* Combined SFK/MEK inhibition prevents metastatic outgrowth of dormant tumor cells. *J Clin Invest* **124**, 156-168 (2014).
19. 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* **15**, 1351-1361 (2013).
20. Gawrzak, S. *et al.* MSK1 regulates luminal cell differentiation and metastatic dormancy in ER(+) breast cancer. *Nature cell biology* **20**, 211-221 (2018).
21. Gao, H. *et al.* The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. *Cell* **150**, 764-779 (2012).
22. Kobayashi, A. *et al.* Bone morphogenetic protein 7 in dormancy and metastasis of prostate cancer stem-like cells in bone. *J Exp Med* **208**, 2641-2655 (2011).

23. Cackowski, F.C. *et al.* Mer Tyrosine Kinase Regulates Disseminated Prostate Cancer Cellular Dormancy. *J Cell Biochem* **118**, 891-902 (2017).
24. Fluegen, G. *et al.* Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic microenvironments. *Nature cell biology* **19**, 120-132 (2017).
25. Sosa, M.S. *et al.* NR2F1 controls tumour cell dormancy via SOX9- and RARbeta-driven quiescence programmes. *Nat Commun* **6**, 6170 (2015).
26. Malanchi, I. *et al.* Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature* **481**, 85-89 (2011).
27. Oskarsson, T. *et al.* Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nat Med* **17**, 867-874 (2011).
28. Zhang, X.H., Giuliano, M., Trivedi, M.V., Schiff, R. & Osborne, C.K. Metastasis dormancy in estrogen receptor-positive breast cancer. *Clin Cancer Res* **19**, 6389-6397 (2013).
29. Kim, R.S. *et al.* Dormancy signatures and metastasis in estrogen receptor positive and negative breast cancer. *PLoS One* **7**, e35569 (2012).
30. 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* **16**, 407 (2014).
31. Calvo, F. *et al.* Cdc42EP3/BORG2 and Septin Network Enables Mechano-transduction and the Emergence of Cancer-Associated Fibroblasts. *Cell Rep* **13**, 2699-2714 (2015).
32. Lee, J.L., Lin, C.T., Chueh, L.L. & Chang, C.J. Autocrine/paracrine secreted Frizzled-related protein 2 induces cellular resistance to apoptosis: a possible mechanism of mammary tumorigenesis. *J Biol Chem* **279**, 14602-14609 (2004).
33. Bovolenta, P., Esteve, P., Ruiz, J.M., Cisneros, E. & Lopez-Rios, J. Beyond Wnt inhibition: new functions of secreted Frizzled-related proteins in development and disease. *J Cell Sci* **121**, 737-746 (2008).
34. Kaur, A. *et al.* sFRP2 in the aged microenvironment drives melanoma metastasis and therapy resistance. *Nature* **532**, 250-254 (2016).
35. Gil, J. & Rodriguez, T. Cancer: The Transforming Power of Cell Competition. *Curr Biol* **26**, R164-166 (2016).
36. Desai, T.J., Brownfield, D.G. & Krasnow, M.A. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **507**, 190-194 (2014).
37. Herriges, M. & Morrisey, E.E. Lung development: orchestrating the generation and regeneration of a complex organ. *Development* **141**, 502-513 (2014).

Materials and 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.OR, 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).

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/Antibody treatments Drugs, inhibitors and blocking antibodies were added in the medium together with cancer cells (unless stated otherwise) and replaced every other day together with fresh medium. Drugs, inhibitors, antibodies 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 All animal experiments were kept in accordance with UK regulations under project licence PPL80/2368 and subsequently PPL70/8380. 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.OR-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.OR-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 cells (EGFP) relative to an internal control (mCherry).

For quantification of disseminated cells and overt metastasis upon protein overexpression, $1x10^6$ D2.OR-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 amount 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.OR-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.OR-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 Supplementary Figure 2 and Figure 5, 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 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 ($1{,}5\text{-}3x10^4$ 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.OR-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.OR-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.

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 SupplementaryTable 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.OR 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.OR-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.OR 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 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.

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[38]. 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)[39], 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 ºC. 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 [40]. 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)[38] was calculated with the Python package *scprep* (https://github.com/KrishnaswamyLab/scprep) using default parameters[41], 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)[42]. Signalling network models were compiled in OmniGraffle 7.

Metal	Antibody Name	Clone	Supplier
089-Y	pHistone H3	HTA28	Biolegend UK
In-113	CEACAM1 (CD66a)	CC1	Thermofisher Scientific
In-115	Pan-CK	AE1/AE3	Biolegend UK
La-139	cPARP [D214]	F21-852	BD Biosciences
Pr-141	pPDK1 [S241]	J66-653.44.22	BD Biosciences
Nd-142	cCaspase 3 [D175]	D3E9	Fluidigm
Nd-143	C-MYC	Y69	Abcam
Nd-146	pEGFR [Y1068]	D7A5	Fluidigm
Sm-147	pMKK4/SEK1 [S257]	C36C11	CST
Nd-148	pSRC [Y418]	SC1T2M3	Thermofisher Scientific
Sm-149	p4E-BP1 [T37/46]	236B4	Fluidigm
Nd-150	pRB [S807/811]	J112-906	Fluidigm
Eu-151	pPKCα [T497]	K14-984	BD Biosciences
Sm-152	pAKT [T308]	J1-223.371	BD Biosciences
Eu-153	pCREB [S133]	87G3	CST
Sm-154	pSMAD1 [S463/465] /pSMAD5 [S463/465] /pSMAD9 [S465/467]	D5B10	CST
Gd-155	pAKT [S473]	D9E	BD Biosciences
Gd-156	pNF-κB p65 [S529]	K10-895.12.50	BD Biosciences
Gd-157	pMKK3/MKK6 [S189/207]	D8E9	CST
Gd-158	pP38 [T180/Y182]	D3F9	CST
Tb-159	pMAPKAPK2 [T334]	27B7	Fluidigm
Gd-160	pAMPKα [T172]	40H9	CST
Dy-161	pBAD [S112]	40A9	CST
Dy-162	pMTOR [S2448]	D9C2	CST
Dy-163	pP90RSK [T359]	D1E9	CST

Dy-164	p120-Catenin [T310]	22/p120 (pT310)	BD Biosciences
Ho-165	Beta-Catenin [Active]	D13A1	CST
Er-166	pGSK-3β [S9]	D85E12	CST
Er-167	pERK1/2 [T202/Y204]	20A	BD Biosciences
Er-168	pSMAD2 [S465/467] /pSMAD3 [S423/425]	D27F4	CST
Tm-169	GFP	5F12.4	Fluidigm
Er-170	pMEK1/2 [S221]	166F8	CST
Yb-172	pS6 [S235/236]	D57.2.2E	CST
Lu-175	CD44	IM7	Biolegend
Yb-176	Cyclin B1	GNS-11	BD Biosciences
Ir-191	Cell-ID Intercalator	-	Fluidigm
Ir-193	Cell-ID Intercalator	-	Fluidigm
Bi-209	acetyl Histone H3 [K27]	D5E4	CST

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[43]) and Bowtie2[44] 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"[45] (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[46, 47]) 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[48-54]). 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 Figure 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 Figure 3b, GSEA results from [D2.OR vs other groups] were visualized using Cytoscape (version 3.6.0) and Enrichment Map plug-in[55]. The map has been manually

687 annotated to reduce complexity and redundancy. Probe 223122_s_at on kmplot.com was
688 used to stratify distant metastasis free survival of breast cancer patients according to *SFRP2*
689 expression.

690 shRNA library screening. Illumina sequence reads with "internal" barcodes (each
691 barcode corresponds to a specific pool of shRNA) were demultiplexed into individual sample
692 files, hairpin sequence was extracted from the backbone vector and common reads collapsed
693 to "tags" providing one instance of each unique candidate hairpin sequence, along with a count
694 of the total number of appearances of each in the original files using in house code. These
695 'tag' sequences were mapped against all annotated library sequences (Supplementary Table
696 2) using bwa-0.5.9[56] and counts of total sequences mapping to each target (counting the total
697 original instances of each hairpin sequence) were generated. These counts were
698 subsequently restricted to consider only targets appearing in the pool specific to that sample
699 and these total raw counts were normalised to the maximum total number of reads across all
700 samples to allow direct comparisons between samples. For each experimental set, a fold
701 change of the representation of each shRNA post-selection relative to the control levels of the
702 same shRNA pre-selection was calculated and these were log-2 transformed. To facilitate this,
703 zero counts in the control were offset by 0.5 to allow the division and zero ratios were set to 1
704 to allow the extraction of logs and enable subsequent clustering. We then ranked genes
705 according to a representation score, defined as the median of the log-2 fold change values.
706 Candidate genes were selected based on two criteria: 1. knock-down of the gene led to loss
707 of dormant cells carrying that shRNA, 2. consistent effect of at least two out of three shRNA
708 sequences.

709

710 **Immunohistochemistry** FFPE material was cut into 5µm sections and subject to antigen
711 retrieval using heated citrate buffer (pH6). Incubation with both primary and secondary
712 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.OR 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 MWCO, Millipore). As loading control, remaining cells were harvested and processed as in [57]. Western blotting was performed as in [57]. 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 Mann-Whitney test for analysis of unpaired data and 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 unless stated otherwise.

Data availability

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

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.

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.

Additional References

38. Behbehani, G.K., Bendall, S.C., Clutter, M.R., Fantl, W.J. & Nolan, G.P. Single-cell mass cytometry adapted to measurements of the cell cycle. *Cytometry A* **81**, 552-566 (2012).
39. Zunder, E.R. *et al.* Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. *Nat Protoc* **10**, 316-333 (2015).
40. Finck, R. *et al.* Normalization of mass cytometry data with bead standards. *Cytometry A* **83**, 483-494 (2013).
41. van Dijk, D. *et al.* Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**, 716-729 e727 (2018).
42. Krishnaswamy, S. *et al.* Systems biology. Conditional density-based analysis of T cell signaling in single-cell data. *Science* **346**, 1250689 (2014).
43. Li, B. & Dewey, C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
44. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).
45. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
46. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739-1740 (2011).
47. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545-15550 (2005).
48. Enzo, E. *et al.* Aerobic glycolysis tunes YAP/TAZ transcriptional activity. *The EMBO journal* **34**, 1349-1370 (2015).
49. Gujral, T.S. *et al.* A noncanonical Frizzled2 pathway regulates epithelial-mesenchymal transition and metastasis. *Cell* **159**, 844-856 (2014).
50. Harper, K.L. *et al.* Mechanism of early dissemination and metastasis in Her2(+) mammary cancer. *Nature* (2016).
51. Horton, J.D. *et al.* Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 12027-12032 (2003).
52. Porstmann, T. *et al.* PKB/Akt induces transcription of enzymes involved in cholesterol and fatty acid biosynthesis via activation of SREBP. *Oncogene* **24**, 6465-6481 (2005).
53. Yu, M. *et al.* RNA sequencing of pancreatic circulating tumour cells implicates WNT signalling in metastasis. *Nature* **487**, 510-513 (2012).
54. Romani, P. *et al.* Extracellular matrix mechanical cues regulate lipid metabolism through Lipin-1 and SREBP. *Nature cell biology* **21**, 338-347 (2019).
55. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G.D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, e13984 (2010).
56. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).
57. Pocaterra, A. *et al.* F-actin dynamics regulates mammalian organ growth and cell fate maintenance. *J Hepatol* **71**, 130-142 (2019).

Figure Legends

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.OR-injected mice ($5x10^5$ D2.OR-eGFP cells/mouse) at the indicated time points along the xy and yz axis. Images highlight thickening of the alveolar wall around disseminated cancer cells over time. F-actin is shown in magenta and GFP (D2.OR cells) in green. Scale bar, 20μm. **b**, Fluorescent IHC for filamentous actin (F-actin), GFP (D2.OR) and Podoplanin (PDPN, AT1 cells) shows that breast cancer cells are intimately connected to AT1 cells *in vivo* and form long protrusions (arrows). Scale bar, 20μm. **c**, Fluorescent IHC of D2.OR cells in the lungs two weeks after intravenous injection ($5x10^5$ D2.OR-eGFP cells/mouse) showing surrounding proliferating (EdU+) mature AT1 cells (PDPN+/AQP5+). i and ii, separate staining for PDPN and AQP5. iii, control uninjected lung. Scale bar, 20μm. **d**, Disseminated MCF7 cells ($1x10^6$ MCF7-GFP cells/mouse) in the lung showing similar pattern of proliferating mature AT1 cells. **e**, Schematic of the lung organotypic system. **f**, Representative immunofluorescence of GFP+ breast cancer cells co-cultured with lung stromal cells. Dashed squares highlight indolent, scattered D2.OR 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. Data points indicate the relative number of cells/mL of each co-culture. Mean normalized pooled samples (n=18) from independent experiments (n=6). Mann-Whitney test. **h**, Quantification of D2.OR cells co-cultured with individual lung stromal cells after 5 days in Mitogen Low-Nutrients Low medium (MLNL). Pooled samples (n=8) from independent experiments (n=2). Dunn's multiple comparisons test. **i**, Immunofluorescence of D2.OR cells cultured alone (left) or co-cultured with AT1-like cells (right). Cells have been stained for fibronectin (FN) and F-actin. 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 two-tailed t-test. **k**, Immunofluorescence of D2.OR cells cultured with or without AT1-like cells. Cells have been stained for phospho-SRC and F-actin. Scale bar, 20μm. **l**, Relative number of D2.OR cells after 5 days of treatment with

cyclic RGD pentapeptide cilengitide. Mean normalised pooled samples (n=12-18) from independent experiments (n=3). Mann-Whitney. **m**, Cilengitide inhibits the formation of protrusions in D2.0R cells cocultured with AT1-like cells. Means from n=4 independent experiments. Paired two-tailed t-test. **n**, 10^6 D2.0R-EGFP cells were injected i.v. in BALBC nude mice. Indicated cohorts were treated four times with 500µg of cilengitide at two days intervals. After 15 days, lungs have been collected and colony area quantified. n=4-5 mice/group. Unpaired two-tailed Student's t-test. **g**, **h**, **l**, **n** plots show data as whisker plots: midline, median; box, 25–75th percentile; whisker, minimum to maximum.

Supplementary Figure 1. Alveolar type1 cells (AT1) regulate behavior of disseminated indolent breast cancer cells. a, Heatmap of Estrogen Receptor (ESR1) and HER2 (Erbb2) expression in D2 cells *in vivo* based on RNAseq presented in Figure 3. Heatmap has been generated with ClustVis tool (https://biit.cs.ut.ee/clustvis/#pathways). **b**, Representative images of lung-disseminated GFP+ breast cancer cells at the indicated time points after tail vein injection. D2.OR and D2.A1 cells are syngeneic cell lines with latent and aggressive behavior respectively. Images show immunohistochemistry (IHC) staining for GFP. Scale bar is 100µm. **c**, Representative images of lungs from wild-type BALB/c mice injected either with D2.OR-EGFP or with D2.A1-EGFP. Lungs were collected and imaged on the GFP channel at the lung surface. Dashed box, lung area magnified in the middle image. Scale bars, 1mm (low magnification) or 100µm (high magnification). **d**, Circularity of D2.OR and D2.A1 cells within the lung parenchyma at 4 days after injection (n=23 cells). n.s., not significant by unpaired two-tailed t-test. **e**, Fluorescent IHC of D2.OR cells in the lungs two weeks after intravenous injection Left: Green, D2.OR cells (EGFP+); Magenta, AT2 cells (TTF1+). Right: Green, D2.OR cells (EGFP+); Magenta, AT2 cells (SP-C+); Yellow: myeloid cells (CD68+); Blue, AT1 cells (PDPN+). Scale bar, 20µm. **f**, Proximity of disseminated D2.OR cells to indicated lung stromal cells at 3 or 14 days post-injection. Lung slices from 3 mice injected with D2.OR-EGFP cells have been stained with multiple markers for different stromal subpopulations. Graphs indicate the percentage of EGFP+ cells in contact with each stromal cells subtype (black: in

916 contact; white: not in contact). AT1: Alveolar Type 1 cells (PDPN+); F: Fibroblasts (VIM+); EC:
917 Endothelial cells (MUC+); AT2: Alvelolar Type 2 cells (SFPC+); Act-F: Activated Fibroblasts
918 (aSMA+); M: Macrophages (CD68+). **g**, 10^6 D2.A1-EGFP or D2.0R-EGFP cells were injected
919 i.v. in BALB/C nude mice. After the indicated time, lungs have been collected, colony area and
920 number of proliferating EGFP+ cells per metastatic lesion were quantified. n = 3 mice/group.
921 **h**, Number of proliferating PDPN-ve and PDPN+ve cells surrounding metastatic lesions and
922 disseminated cells in Supplementary Figure 1g. n = 3 mice/group. **i**, 10^6 MCF7-EGFP cells
923 were injected i.v. in BALBC nude mice. After 3 days, lungs have been collected, number of
924 proliferating PDPN+ve cells surrounding the metastatic lesion was quantified. n = 3
925 mice/group. **j**, Relative mRNA levels of stromal cell-type specific markers of the different
926 cellular populations included in the lung coculture system. AT1, alveolar type1 cells; AT2,
927 alveolar type2 cells. Dots are means from independent experiments (n=3). Unpaired two-tailed
928 t-test. **k**, Percentage of Ki67+-D2.OR-EGFP or -D2.A1-EGFP cells cultivated together with
929 lung stromal cell lines in Mitogen Low-Nutrients Low medium for 4 days. Mean normalized
930 pooled samples (n=7) from independent experiments (n=2). Unpaired two-tailed t-test. **l**,
931 Growth curves of D2.OR-EGFP and D2.A1-EGFP cells *in vitro* with permissive (MHNH) or
932 restrictive (MLNL) medium. Confluency values at indicated time points were log10-
933 transformed and linear regression was calculated. Line was forced to go through the origin.
934 Solid line, mean of best-fit line; dashed lines, 95% confidence bands. **m**, D2.OR-EGFP cells
935 were cocultured with lung stromal cells for 5 days (or on air-permeable surface only as control),
936 isolated by fluorescence-activated cell sorting (FACS), and their growth kinetic *in vitro* in MLNL
937 on standard plastic plates measured over time (lines are overlapped). n=2 independent
938 experiments. **n**, Relative number of 4T07-EGFP or MCF7-EGFP cells cultivated alone or
939 together with AT1-like cells in MLNL medium for 5 days. Mean normalized pooled samples
940 (n=12-24) from independent experiments (n=3-4). Mann-Whitney test for 4T07, unpaired two-
941 tailed t-test for MCF7 data. **o**, Plots show the relative frequency (number of events/starting
942 number of D2.OR cells) of mitotic (left) and apoptotic (right) events in D2.OR cells cultured in
943 MLNL media in the absence or presence of AT1-like cells. Linked points indicate mean data

from individual experiments (n=7). Wilcoxon test. **p**, Relative number of D2.OR-EGFP cells cocultured with different combinations of lung stromal cells (as indicated) in MLNL medium for 5 days. Mean normalized pooled samples (n=9) from independent experiments (n=3). Mann-Whitney test. **q**, Quantification of D2.OR cells co-cultured with individual lung stromal cells in Mitogen High-Nutrients High medium (MHNH) for 7 days. Relative number of cells was calculated by measuring the GFP+ area per well. Pooled samples (n=8) from independent experiments (n=3). Dunn's multiple comparisons test. **r**, Circularity of D2.OR cells alone or cocultured with AT1-like cells (n=26-29 cells). Unpaired two-tailed t-test. **s**, Percentage of indicated human breast cancer cells with protrusions alone or in coculture with AT1-like cells. n=3 independent experiments. Paired t-test. **t**, Representative images of protrusions induced by coculturing of MCF7 cells with AT1-like cells. Scale bar, 20μm. **u**, Immunofluorescence of 4T07 cells cultured alone (left) or co-cultured with AT1-like cells (right). Cells have been stained for fibronectin (FN) and F-actin. Scale bar, 20μm. **v**, Percentage of D2.A1 cells with protrusions alone or in coculture with AT1-like cells. n=3 independent experiments. Paired t-test. **w**, Cilengitide inhibits the presence formation of protrusions in MCF7 cells cocultured with AT1-like cells. n=3 independent experiments. Paired t-test. **d**, **g**, **h**, **i**, **l**, **n**, **p**, **q**, **r** plots show data as whisker plots: midline, median; box, 25–75th percentile; whisker, minimum to maximum.

Figure 2. Mass cytometry analysis reveals signaling pathways involved in the crosstalk between AT1 and indolent breast cancer cells. **a**, Schematics representation of the experimental outline of mass cytometry assay. **b**, Heatmaps of EMD values (Earth Mover's Distance) estimating the activation of the indicated molecules in D2.OR or MCF7 alone or in coculture with AT1-like cells. Representative of three independent experiments. **c**, Plots show cell number fold change and relative frequency (number of events/starting number of D2.OR cells) of apoptotic and mitotic events in D2.OR cells determined from movies of D2.OR 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 (n=3-11). 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.OR-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). Scale bar is 20µm.

Supplementary Figure 2. Mass cytometry analysis reveals signaling pathways involved in the crosstalk between AT1 and indolent breast cancer cells. **a**, Heatmaps of EMD values showing the activation of relevant markers in AT1-like cells cocultured with D2.OR or MCF7. Representative results from n=3 independent repetitions. **b**, Plot showing increase phospho-HistoneH3 (S28) signal in AT1-like cells co-cultured with D2.OR cells. **c**, DREVI plots showing the relationship between the indicated phospho-antibody signals in D2.OR monocultures or cocultures with AT1-like cells (DREMI score in upper left corner). **d**, Number of cells after the indicated treatment (for two days) relative to untreated cells. Mean of n = 3-4 independent experiments. One-way ANOVA test. **e**, Histogram of EMD values showing the inhibition of P-ERK abundance in D2.OR cells cocultured with AT1-like cells. Bars show the average of two technical replicates. Representative results from n=3 independent repetitions. **f**, Plot shows the area of D2.OR colonies ten days after intravenous delivery into either control Balb/C nude mice or Trametinib treated mice – 5 control and 4 trametinib treated mice were analyzed. Mann-Whitney test. **g**, Percentage of D2.OR cells with protrusions after treatment with indicated inhibitors for two days. Data are means of n=3 independent experiments. One-way ANOVA test. **h**, Control D2.OR-EGFP cells have been treated for two days with indicated drugs. Fibronectin fibrils were quantified after immunostaining. n=3-5 experiments. t-test with Welch correction: comparisons between "extensive fibrils" category. One-way ANOVA test. **i**, Percentage of MCF7 cells with protrusions after treatment with SFKi in monoculture or coculture with AT1-like cells. One-way ANOVA test. **j**, Relative expression of fibronectin mRNA in D2.OR-EGFP cells cultivated with AT1-like cells in MLNL medium -/+ SFKi for 4 days. Mean normalised pooled samples (n=12) from independent experiments (n=3). Mann-

Whitney test. **k**, Images showing F-actin and activated P-Src (Y418) in D2.OR cells cocultured with AT1-like cells (left). The signal is lost upon treatment of cocultures with EGFRi. Scale bar is 20µm. Plots in **e**, **g, h** and **i** are as mean and SD.

Figure 3. Gene expression analysis of lung-disseminated indolent breast cancer cells *in vivo*. **a**, D2.OR-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.OR *in vivo* compared to D2.OR *in vitro*, D2.A1 *in vivo* and *in vitro*. Red indicates higher expression and blue indicates low expression relative to the mean expression of the gene across all samples. **b**, Enrichment map for disseminated indolent breast cancer cells *in vivo*. The map shows gene-set enrichment results of D2.OR cells *in vivo* compared to the other groups. Node size, genes in pathway; node color, enrichment score (orange indicates enrichment in D2.OR *in vivo*, blue indicates enrichment in the other groups); edge width, overlap size between connected nodes. **c**, qPCR analysis of selected genes from independent *in vitro* and *in vivo* samples (n=3-6 mice or wells). 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 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 multivariate analysis of stage, lymph node status and dormancy signature in tamoxifen-treated breast cancer patients. x-axis represents the hazard ratio. **f**, AT1-like cells trigger expression of ECM and EMT genes in D2.OR cells *in vitro*. qPCR analysis of D2.OR-EGFP cells cultured alone or together with AT1-like cells for 4 days in MLNL medium. Mean normalized pooled samples (n=8-9) from independent experiments (n=3-4). Mann-Whitney test.

Supplementary Figure 3. Gene expression analysis of lung-disseminated indolent breast cancer cells *in vivo*. **a**, Representative GSEA analysis from the top up- and down-regulated gene sets in D2.OR cells *in vivo* compared to the other groups. NES, normalized enrichment score. FDR, false discovery rate. **b**, Heatmap shows normalized expression values for two dormancy markers (Sharp1 and Nr2f1). **c**, Kaplain-Meier curves showing DMFS of ER+ breast with high and low expression of dormancy signature. Data have been plotted starting from month 0 (left) or month 24 (right). **d**, Kaplain-Meier curves of ER+ breast cancer patients from publicly available datasets used in Ref. 29, stratified according to the dormancy signature (left). Right, Overlap between our dormancy signature and genes included in the dormancy score (Supplementary Table 1). **e**, Kaplan-Meier curves showing distant metastasis free survival (DMSF) of patients derived from http://co.bmc.lu.se/gobo/gsa.pl stratified according additional signatures generated from the other groups analysed with RNAseq in Figure 3 (Supplementary Table 1). **f**, BMP and Wnt target genes expression as in Figure 3f. Mann-Whitney test.

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*. We first selected 59 candidates among the top upregulated genes in D2.OR cells *in vivo* and designed a shRNA library including 3 shRNA sequences for each gene. shRNA were combined in pools of 14-15 shRNAs/pool. We then transduced D2.OR-EGFP cells with each pool of shRNA-containing lenviruses at a MOI optimized to ensure a single integration per genome. After puromycin selection, cells were injected intravenously in triplicated mice and collected after 3 weeks. Genomic DNA from *in vivo* selected cells as well as from cell populations before injection as reference, and relative abundance of each shRNA sequence (relative to pre-injection abundance) was estimated after Next Generation Sequencing. **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.OR-EGFP-shSfrp2 or -shControl cells (3 independent shRNA sequences) were injected with an equal amount of D2.OR-mCherry-shControl cells intravenously (ratio=1). After 3 weeks, breast cancer cells were isolated and the ratio EGFP+-cells/mCherry+-cells calculated (n=4-5 mice). Unpaired two-tailed t-test with Welch's correction. **d**, qPCR for Sfrp2 of D2.OR-EGFP cells cultivated alone or cocultured with AT1-like cells in MLNL medium for 4 days. Mean normalized pooled samples (n=24-27) from independent experiments (n=7). Mann-Whitney test. **e**, as in **d**, in addition cells were treated with the SFK-inhibitor (Dasatinib, 50nM) or DMSO, as control. Mean normalized pooled samples (n=10-12) from independent experiments (n=3). One-way ANOVA test. **f**, Mean Fibronectin (FN1) intensity per cell in control and SFRP2 OE indolent breast cancer cells. Mann-Whitney test. Representative results from n=2 independent repetitions. **g**, SFRP2 overexpression rescues loss-of-protrusion following SFKi inhibition. Control or SFRP2-overexpressing D2.OR-EGFP cells have been cultured alone or with AT1-like cells in presence or not of SFK-inhibitor (Dasatinib). n=4-5 independent experiments. Mann-Whitney test. **h**, Heatmaps of EMD values estimating the activation of the indicated molecules in control and SFRP2 OE D2.OR cells. Representative of three independent experiments. **i**, Images show F-actin and FN staining of D2.OR-EGFP cells +/- SFRP2 over-expression co-cultured +/- AT1-like cells in MLNL medium with Dasatinib (SFKi) for 48hrs. GFP labelling of D2.OR cells in shown in cyan. Scale bar is 20µm. **j**, SFRP2 overexpressing D2.OR-EGFP cells have been treated for two days with indicated drugs. Fibronectin fibrils were quantified after immunostaining. n=3 experiments. One-way ANOVA comparisons between "extensive fibrils" category. **k**, Quantification of D2.OR cell death in the indicated conditions: +/- AT1-like cells, +/- SFRP2 over-expression and +/- Dasatinib treatment (SFKi). Cells have been treated for two days and quantified as in Figure 2c. Mean and S.E.M. are shown (n=3-7 independent experiments). Mann-Whitney test. **l**, Left, fluorescent in situ images of D2.OR and SFRP2 over-expressing D2.OR cells in the lung alveolar space. F-actin is shown in magenta and GFP (D2.OR cells) in green. Scale bar, 20µm. Right, circularity of lung

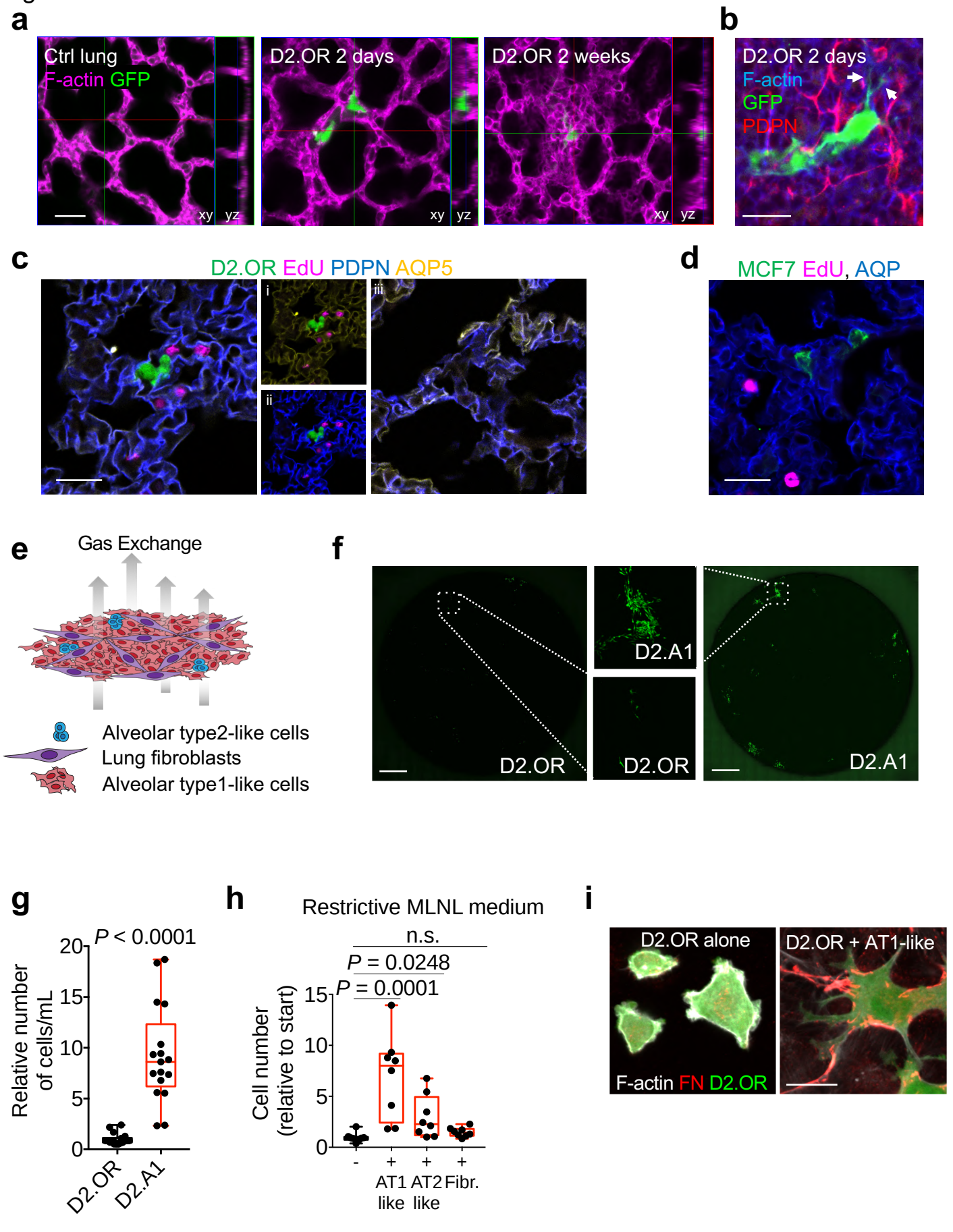
disseminated wt or SFRP2-overexpressing D2.OR-EGFP cells (n=4 mice). Mann-Whitney test. **m**, Quantification of the metastatic burden and metastatic colony area 2 weeks after intravenous injection of 4T07-EGFP cells (+/- SFRP2 over-expression, n=5 mice for control, n=6 mice for SFRP2) into Balb/C mice. Mann-Whitney test for metastatic burden. Unpaired two-tailed t-test with Welch's correction for colony area experiments. **n**, Quantification of the metastatic burden and metastatic colony area 2 weeks after in the intravenous injection of D2.OR-EGFP cells (+/- SFRP2 over-expression n=5 mice) into Balb/C nude mice. Mann-Whitney test for metastatic burden. Unpaired two-tailed t-test with Welch's correction for colony area experiments. **f**, **l**, **m**, **n** plots show data as whisker plots: midline, median; box, 25–75th percentile; whisker, minimum to maximum. Plots in **g** and **j** are as mean and SD.

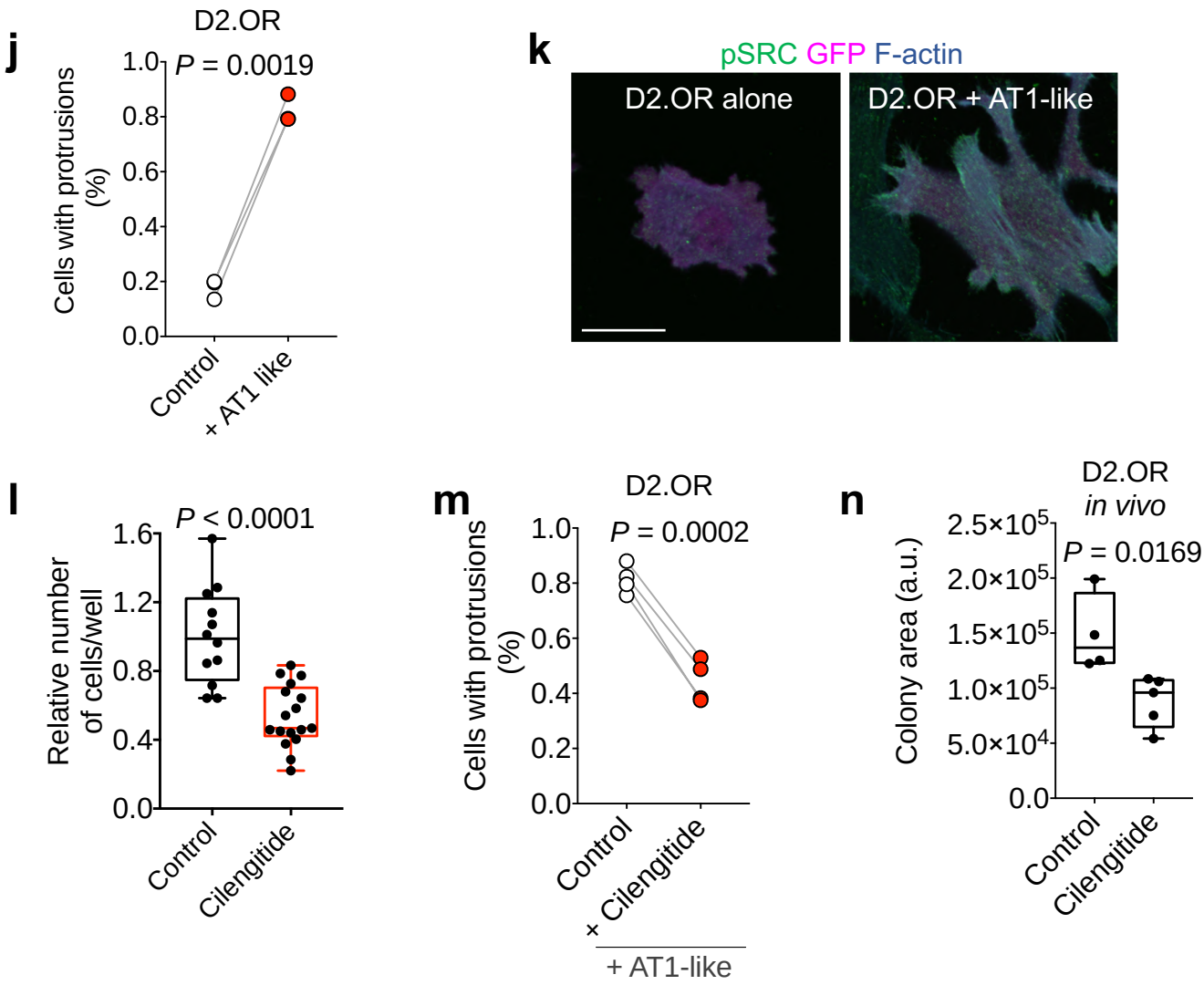
Supplementary Figure 4. A loss-of-function screen *in vivo* identifies SFRP2 as survival regulator in lung disseminated indolent breast cancer cells. **a**, Volcano plot of RNAseq expression data of D2.OR cells *in vivo* compared to the other groups. In blue, candidate genes selected for step 2 validation. **b**, Step 2 validation of candidate genes. Subpopulations of D2.OR-EGFP cells bearing a single shRNA for the indicated gene were individually generated (3 shRNA sequences/gene). Cells with shRNA for the same gene were mixed together in equal amount, injected in tail vein of BALB/c nude mice (n=3-6 mice) and processed as in Figure 4c. Unpaired two-tailed t-test with Welch's correction. **c**, Subpopulations of D2.OR-EGFP-shSfrp2 cells were mixed and injected in the tail vein with an equal amount of D2.OR-mCherry-shControl. After 3 days to allow seeding and extravasation in the lung parenchyma, lungs were collected and GFP+ and mCherry+ simultaneously quantified to rule out pre-dissemination role of SFRP2 (n=4 mice). Scale bar, 500μm. Unpaired two-tailed t-test with Welch's correction. **d**, *In vitro* growth curves of D2.OR-EGFP cells bearing the indicated shRNAs for Sfrp2. Confluency values at indicated time points were log10-transformed and linear regression was calculated. Line was forced to go through the origin. Solid line, mean of best-fit line; dashed lines, 95% confidence bands. **e**, Relative expression levels of Sfrp2 in

D2.OR.EGFP cells on plastic, isolated from mammary fat pad or lung-disseminated (n=3-5 mice or wells). Unpaired two-tailed t-test. **f**, Histogram showing the induction of SFRP family members by AT1-like conditioned media in both D2.OR and 4TO7 cells. Mean normalized pooled samples (n=9) from independent experiments (n=3-4). Mann-Whitney test. **g**, Left, qPCR for canonical Wnt target genes of D2.OR-EGFP carrying interfering sequences for SFRP2 cultivated with AT1-like cells in MLNL medium for 4 days. Right, qPCR for the Wnt target Axin2 in control and SFRP2-overexpressing cells. Mean normalized pooled samples (n=9-13) from independent experiments (n=3-4). Mann-Whitney test. **h**, Conditioned media from confluent D2.0R-EGFP-Control or SFRP2 OE cells plated in MLNL were concentrated and analyzed by Western Blotting. Cells have been treated or not with 50ug/mL of Heparin to allow SFRP2 solubilization in the medium. **i**, Plot shows the effect of heparin, which binds and inhibits SFRP family proteins, on D2.OR cell number when co-cultured with AT1-like cells. Mean normalized pooled samples (n=18) from independent experiments (n=3). Unpaired two-tailed t-test. **j**, Control or SFRP2 overexpressing MCF7 cells were plated alone or in presence of AT1-like cells. Plot shows the percentage of cells with protrusions in each experiment. n=3 independent experiments. Paired two-tailed t-test. **k**, Quantification of cell D2.OR cell proliferation (as judged by mitoses) in the indicated conditions: +/- AT1-like cells, +/- SFRP2 over-expression, and +/- SFKi treatment. Mean and S.E.M. are shown (n=5 independent experiments). Unpaired two-tailed t-test. **l**, Quantification of the metastatic burden and metastatic colony area two weeks after intravenous injection of human indolent breast cancer cell lines (T47D-DBM and MCF7) (+/- SFRP2 over-expression, n=3 mice for control, n=3 mice for SFRP2) into Balb/C nude mice. Unpaired two-tailed t-test. **m**, *In vitro* growth curves of control and SFRP2 over-expressing D2.OR and 4T07 cells. Confluency values at indicated time points were log10-transformed and linear regression was calculated. Line was forced to go through the origin. Solid line, mean of best-fit line; dashed lines, 95% confidence bands. **n**, Proximity of disseminated SFRP2-overexpressing D2.OR cells to indicated lung stromal cells at 3 or 14 days post-injection. Lung slices from 3 mice injected with D2.OR-EGFP cells have been stained with multiple markers for different stromal subpopulations. Graphs indicate the

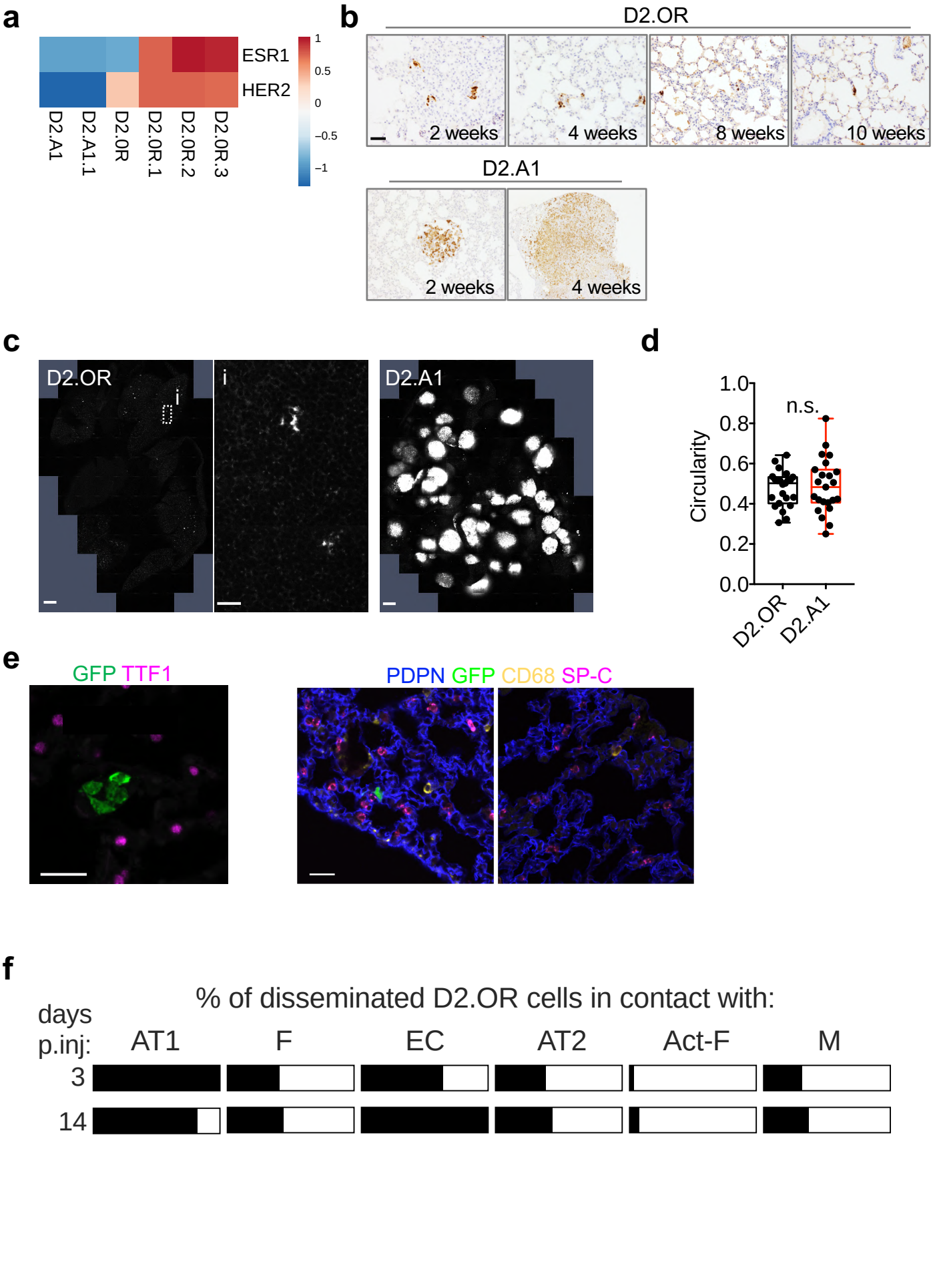
percentage of EGFP+ cells in contact with each stromal cells subtype (black: in contact; white: not in contact). Staining as in Suppl. Figure 1f. **o**, Schematic illustration of the signalling between AT1 cells and breast cancer cells that supports metastatic persistence. **i**, **j**, **l** plots show data as whisker plots: midline, median; box, 25–75th percentile; whisker, minimum to maximum.

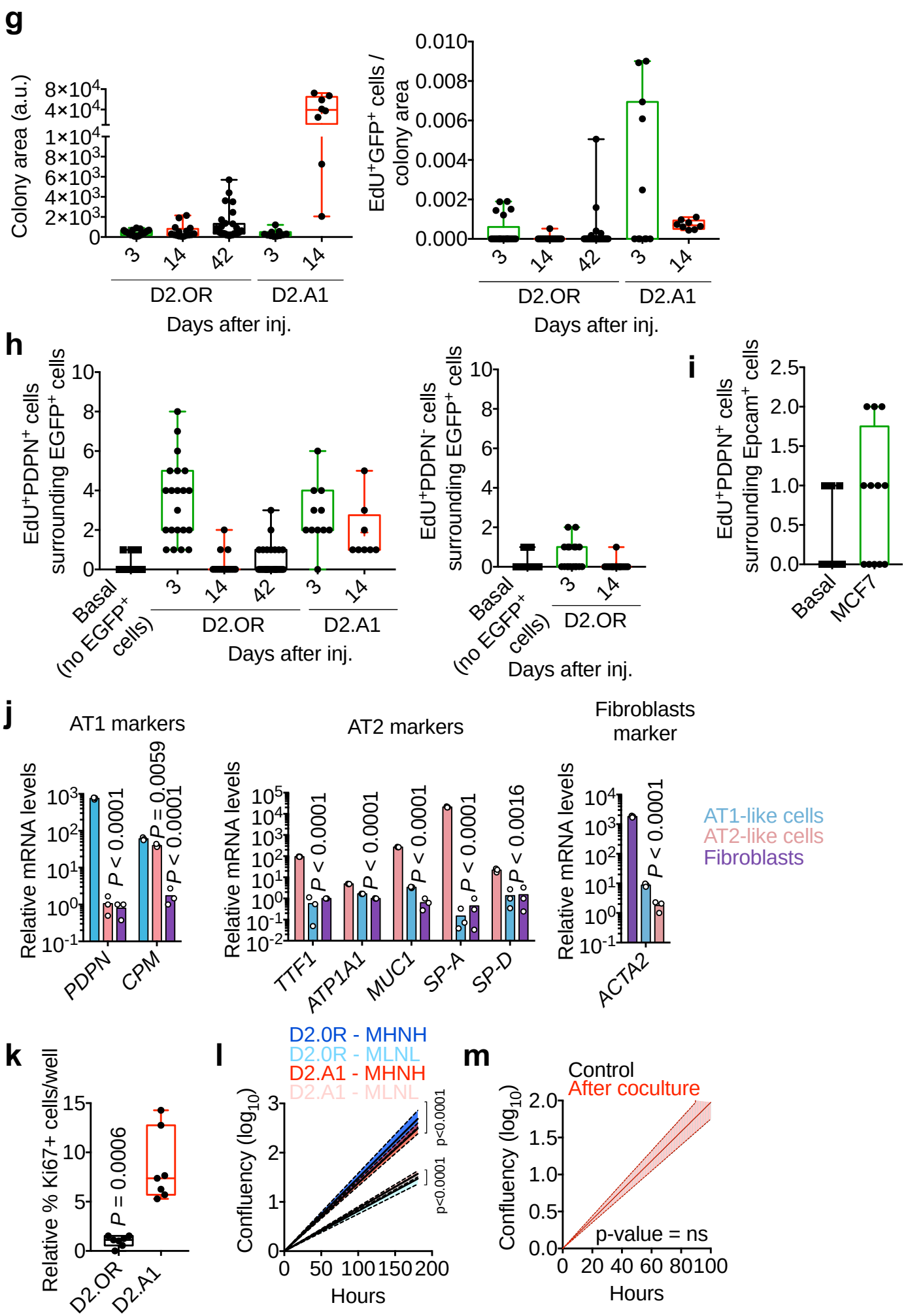
Figure 1


a
Ctrl lung
F-actin GFP
xy yz
D2.OR 2 days
xy yz
D2.OR 2 weeks
xy yz
b
D2.OR 2 days
F-actin
GFP
PDPN
c
D2.OR EdU PDPN AQP5
i
ii
iii
d
MCF7 EdU, AQP
e
Gas Exchange
Alveolar type2-like cells
Lung fibroblasts
Alveolar type1-like cells
f
D2.OR
D2.A1
D2.OR
D2.A1
g
P < 0.0001
Relative number of cells/mL
20
15
10
5
0
D2.OR
D2.A1
h
Restrictive MLNL medium
n.s.
P = 0.0248
P = 0.0001
Cell number (relative to start)
15
10
5
0
-
+
+
+
AT1 like
AT2 like
Fibr.
i
D2.OR alone
F-actin FN D2.OR
D2.OR + AT1-like


j
D2.OR
P = 0.0019
Cells with protrusions (%)
1.0
0.8
0.6
0.4
0.2
0.0
Control
+ AT1 like
k
pSRC GFP F-actin
D2.OR alone
D2.OR + AT1-like
l
P < 0.0001
Relative number of cells/well
1.6
1.2
0.8
0.4
0.0
Control
Cilengitide
m
D2.OR
P = 0.0002
Cells with protrusions (%)
1.0
0.8
0.6
0.4
0.2
0.0
Control
+ Cilengitide
+ AT1-like
n
D2.OR
in vivo
P = 0.0169
Colony area (a.u.)
2.5×10^5
2.0×10^5
1.5×10^5
1.0×10^5
5.0×10^4
0.0
Control
Cilengitide

Supplementary Figure 1




g
Colony area (a.u.)
8×10^4
4×10^4
1×10^4
8×10^3
6×10^3
4×10^3
2×10^3
0
3
14
42
3
14
D2.OR
D2.A1
Days after inj.
EdU+GFP+ cells / colony area
0.010
0.008
0.006
0.004
0.002
0.000
3
14
42
3
14
D2.OR
D2.A1
Days after inj.
h
EdU+PDPN+ cells surrounding EGFP+ cells
10
8
6
4
2
0
Basal (no EGFP+ cells)
3
14
42
3
14
D2.OR
D2.A1
Days after inj.
EdU+PDPN- cells surrounding EGFP+ cells
10
8
6
4
2
0
Basal (no EGFP+ cells)
3
14
D2.OR
Days after inj.
i
EdU+PDPN+ cells surrounding Epcam+ cells
2.5
2.0
1.5
1.0
0.5
0.0
Basal
MCF7
j
AT1 markers
AT2 markers
Fibroblasts marker
Relative mRNA levels
10^3
10^2
10^1
10^0
10^-1
P < 0.0001
P = 0.0059
P < 0.0001
PDPN
CPM
Relative mRNA levels
10^5
10^4
10^3
10^2
10^1
10^0
10^-1
10^-2
P < 0.0001
P < 0.0001
P < 0.0001
P < 0.0001
P < 0.0016
TTF1
ATP1A1
MUC1
SP-A
SP-D
Relative mRNA levels
10^4
10^3
10^2
10^1
10^0
10^-1
P < 0.0001
ACTA2
AT1-like cells
AT2-like cells
Fibroblasts
k
Relative % Ki67+ cells/well
15
10
5
0
P = 0.0006
D2.OR
D2.A1
l
D2.0R - MHNH
D2.0R - MLNL
D2.A1 - MHNH
D2.A1 - MLNL
Confluency (log10)
3
2
1
0
p<0.0001
p<0.0001
0
50
100
150
200
Hours
m
Control
After coculture
Confluency (log10)
2.0
1.5
1.0
0.5
0.0
p-value = ns
0
20
40
60
80
100
Hours

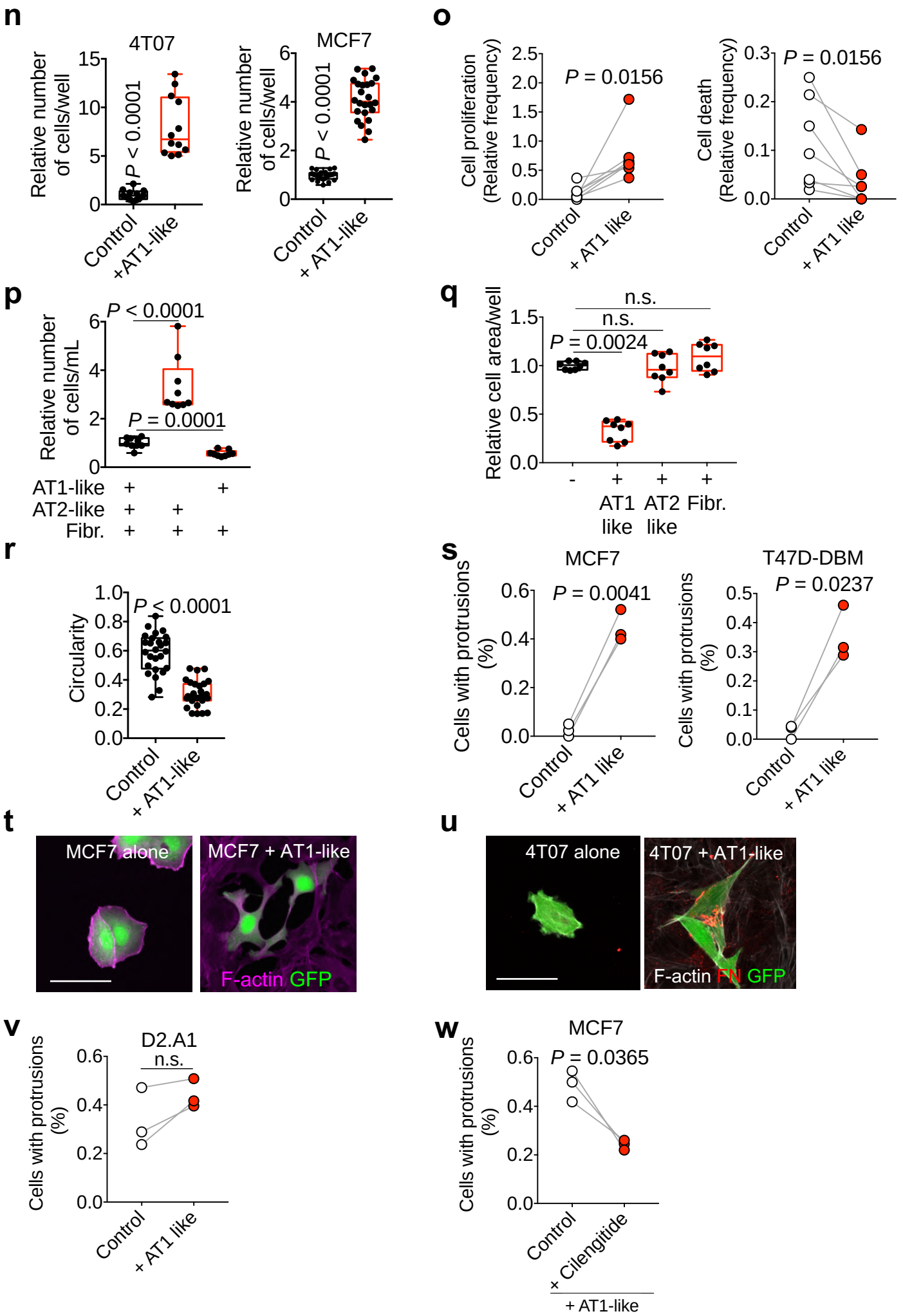
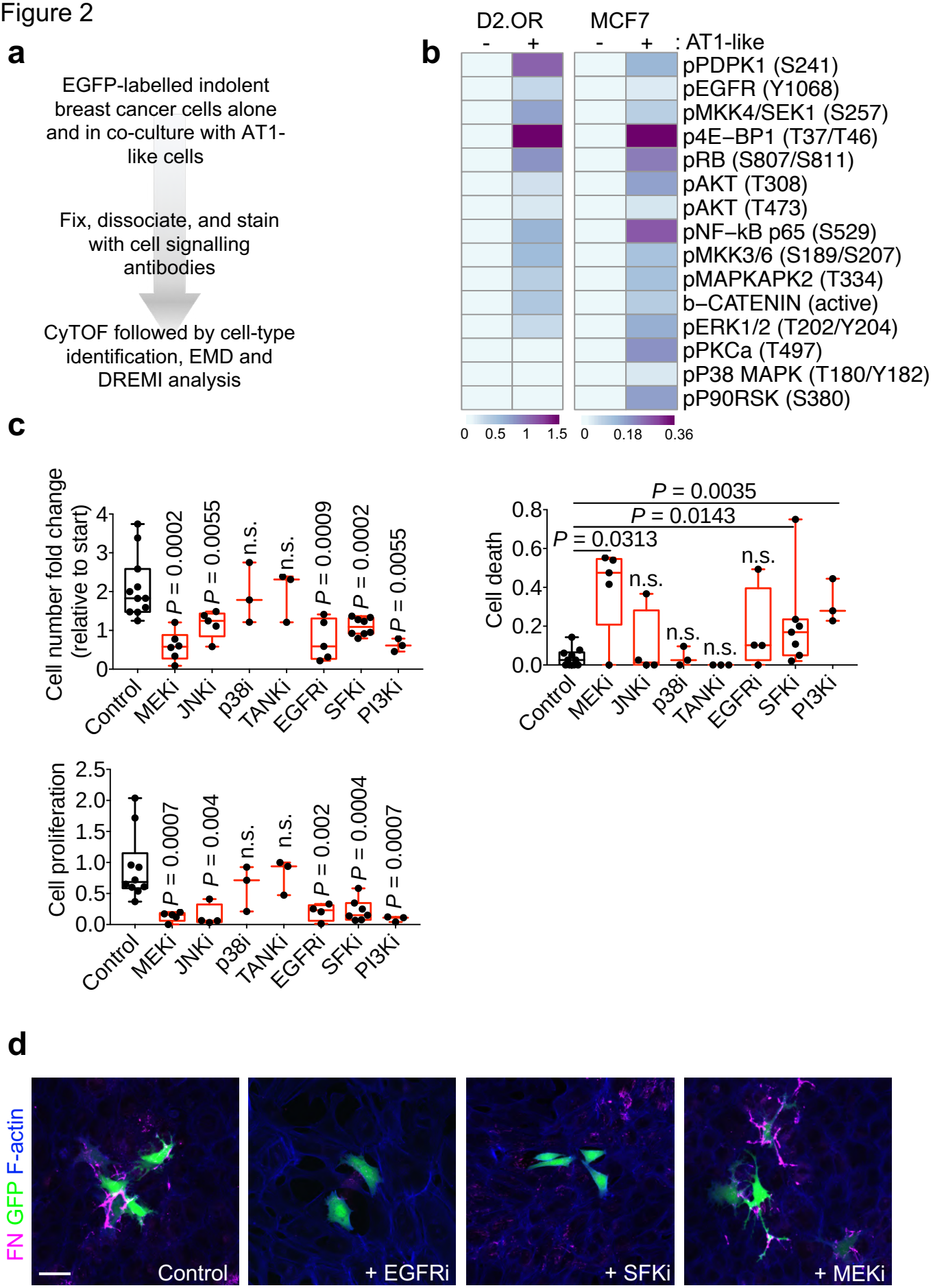
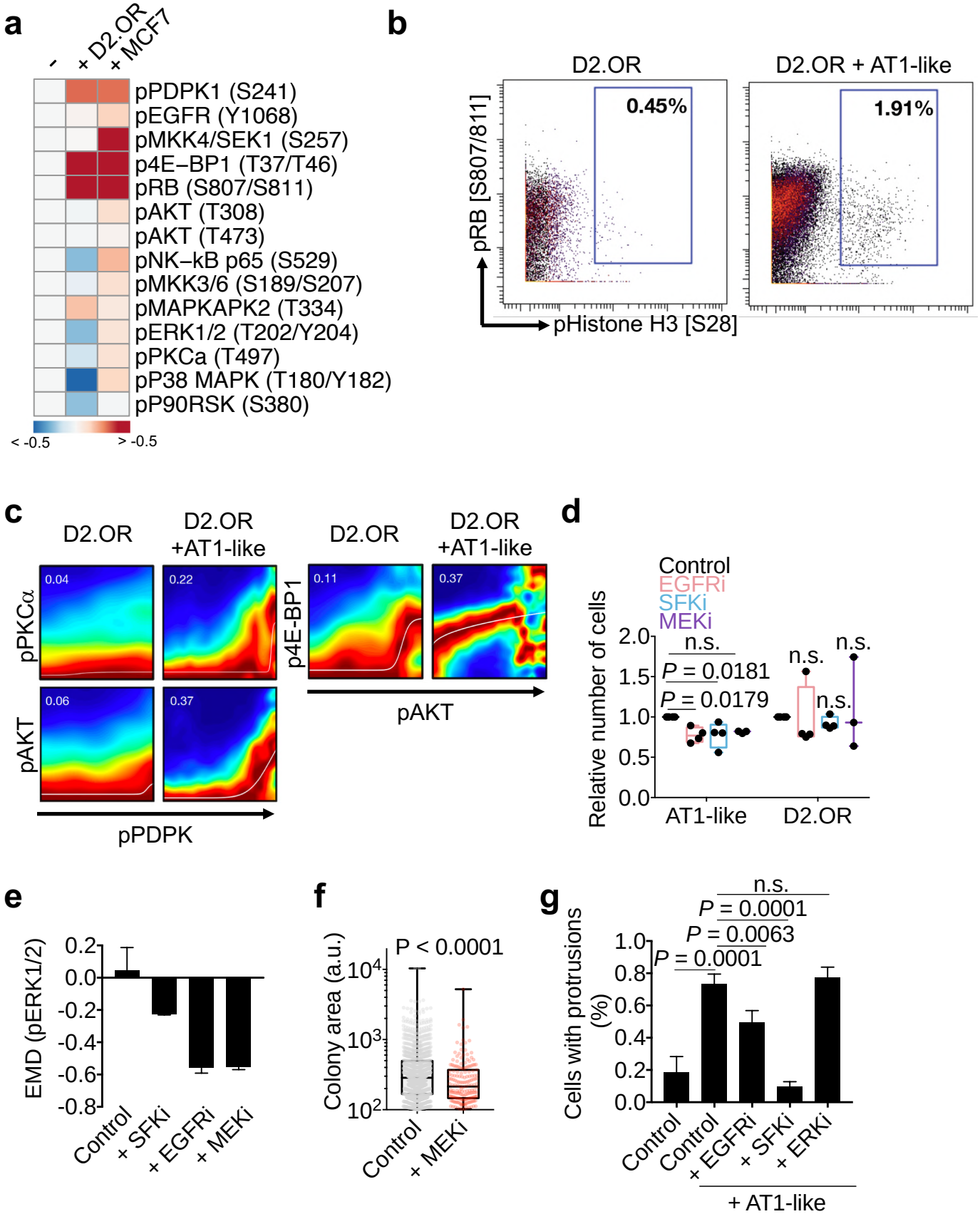

n
4T07
MCF7
Relative number of cells/well
P < 0.0001
P < 0.0001
Control
+AT1-like
+ AT1-like
o
Cell proliferation (Relative frequency)
P = 0.0156
Cell death (Relative frequency)
P = 0.0156
Control
+ AT1 like
p
Relative number of cells/mL
P < 0.0001
P = 0.0001
AT1-like
AT2-like
Fibr.
q
Relative cell area/well
n.s.
n.s.
P = 0.0024
AT1 like
AT2 like
Fibr.
r
Circularity
P < 0.0001
Control
+ AT1-like
s
MCF7
T47D-DBM
Cells with protrusions (%)
P = 0.0041
P = 0.0237
Control
+ AT1 like
t
MCF7 alone
MCF7 + AT1-like
F-actin GFP
u
4T07 alone
4T07 + AT1-like
F-actin FN GFP
v
D2.A1
n.s.
Cells with protrusions (%)
Control
+ AT1 like
w
MCF7
P = 0.0365
Cells with protrusions (%)
Control
+ Cilengitide
+ AT1-like

Figure 2


a
EGFP-labelled indolent breast cancer cells alone and in co-culture with AT1-like cells
Fix, dissociate, and stain with cell signalling antibodies
CyTOF followed by cell-type identification, EMD and DREMI analysis
b
D2.OR
MCF7
- + - + : AT1-like
pPDPK1 (S241)
pEGFR (Y1068)
pMKK4/SEK1 (S257)
p4E–BP1 (T37/T46)
pRB (S807/S811)
pAKT (T308)
pAKT (T473)
pNF–kB p65 (S529)
pMKK3/6 (S189/S207)
pMAPKAPK2 (T334)
b–CATENIN (active)
pERK1/2 (T202/Y204)
pPKCa (T497)
pP38 MAPK (T180/Y182)
pP90RSK (S380)
0 0.5 1 1.5
0 0.18 0.36
c
Cell number fold change (relative to start)
P = 0.0002
P = 0.0055
n.s.
n.s.
P = 0.0009
P = 0.0002
P = 0.0055
Control MEKi JNKi p38i TANKi EGFRi SFKi PI3Ki
Cell death
P = 0.0035
P = 0.0143
P = 0.0313
n.s.
n.s.
n.s.
n.s.
Cell proliferation
P = 0.0007
P = 0.004
n.s.
n.s.
P = 0.002
P = 0.0004
P = 0.0007
d
FN GFP F-actin
Control
+ EGFRi
+ SFKi
+ MEKi

Supplementary Figure 2



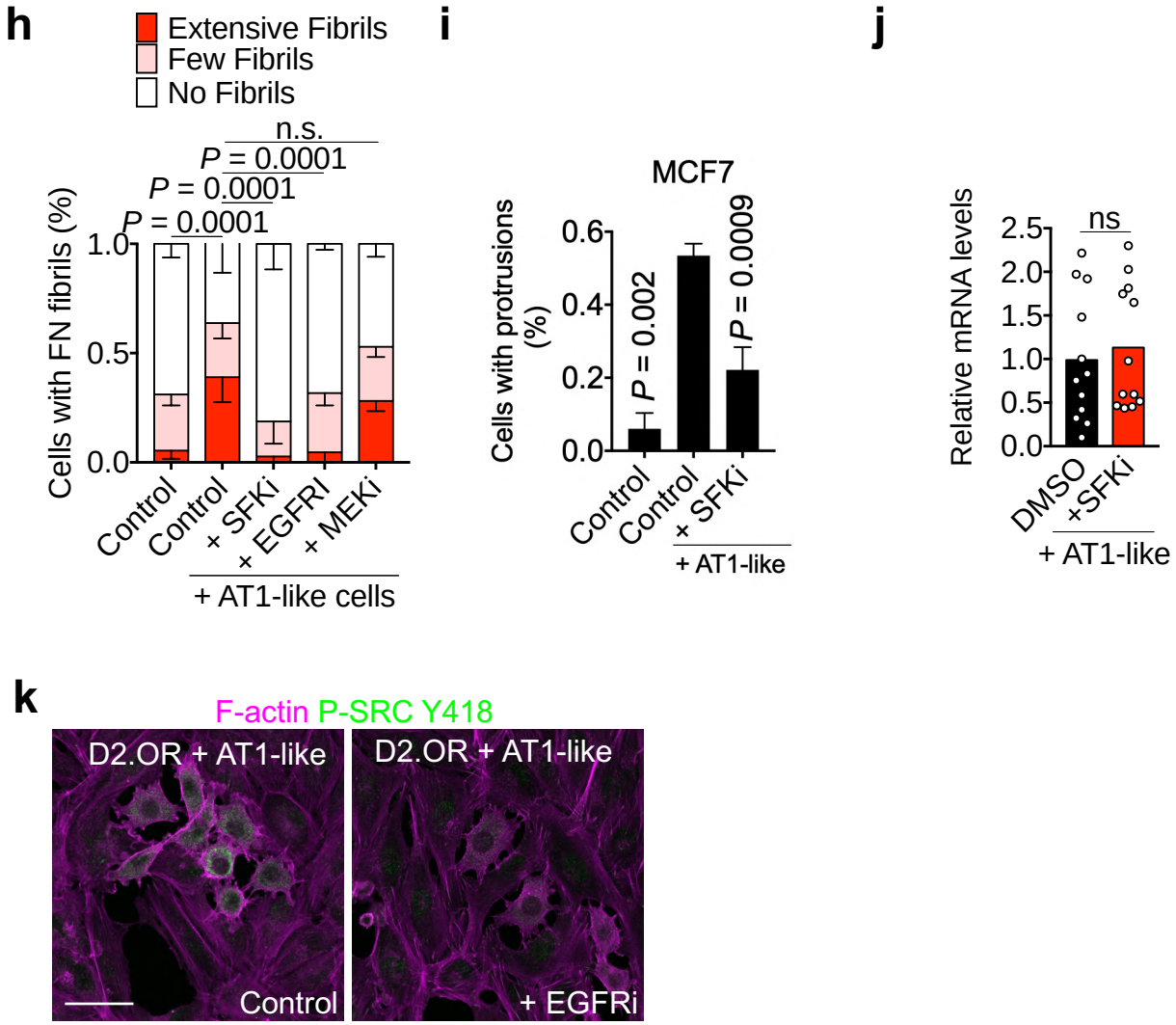
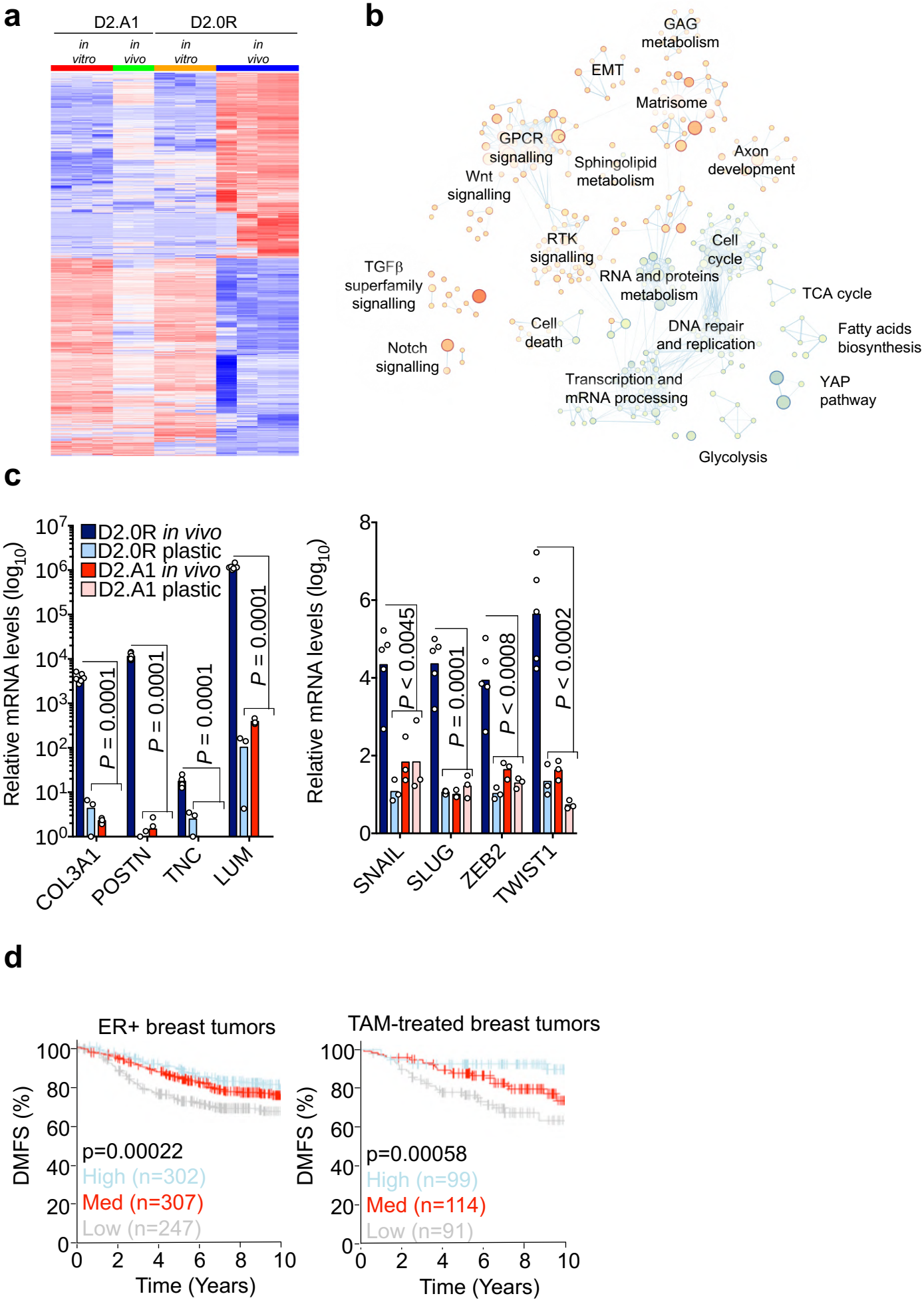
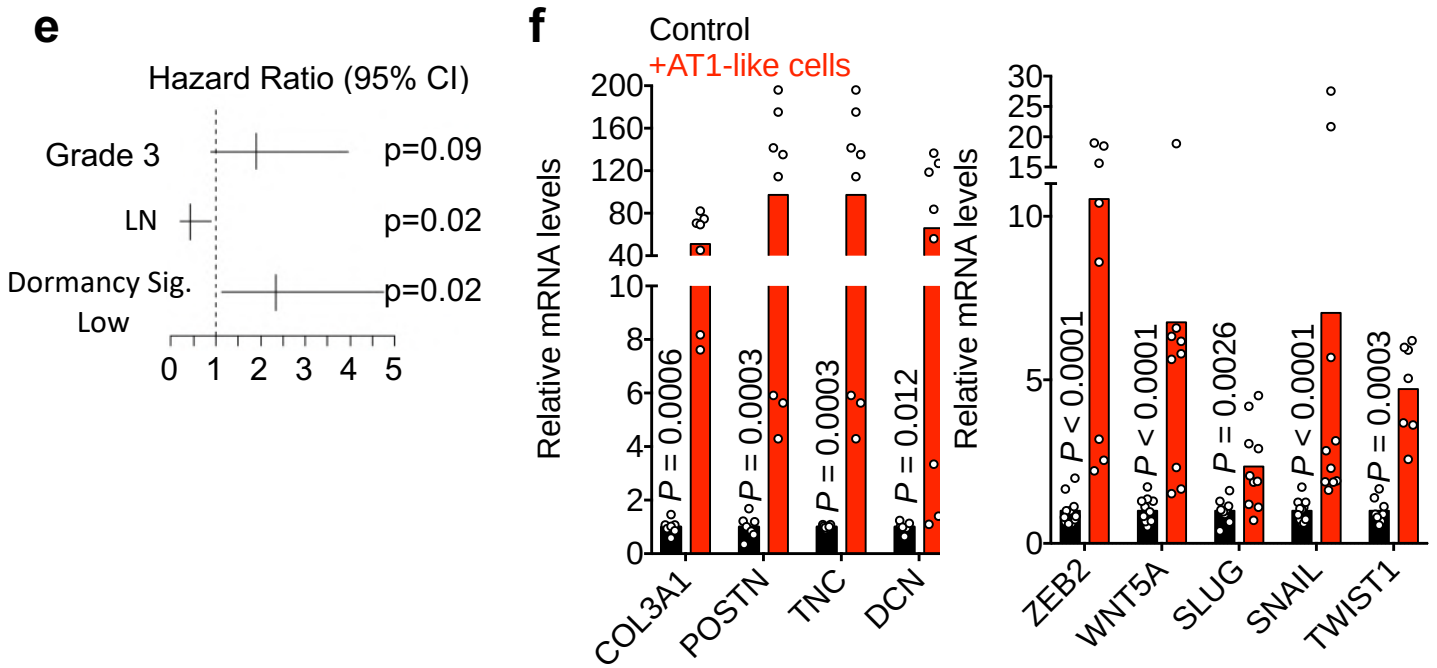
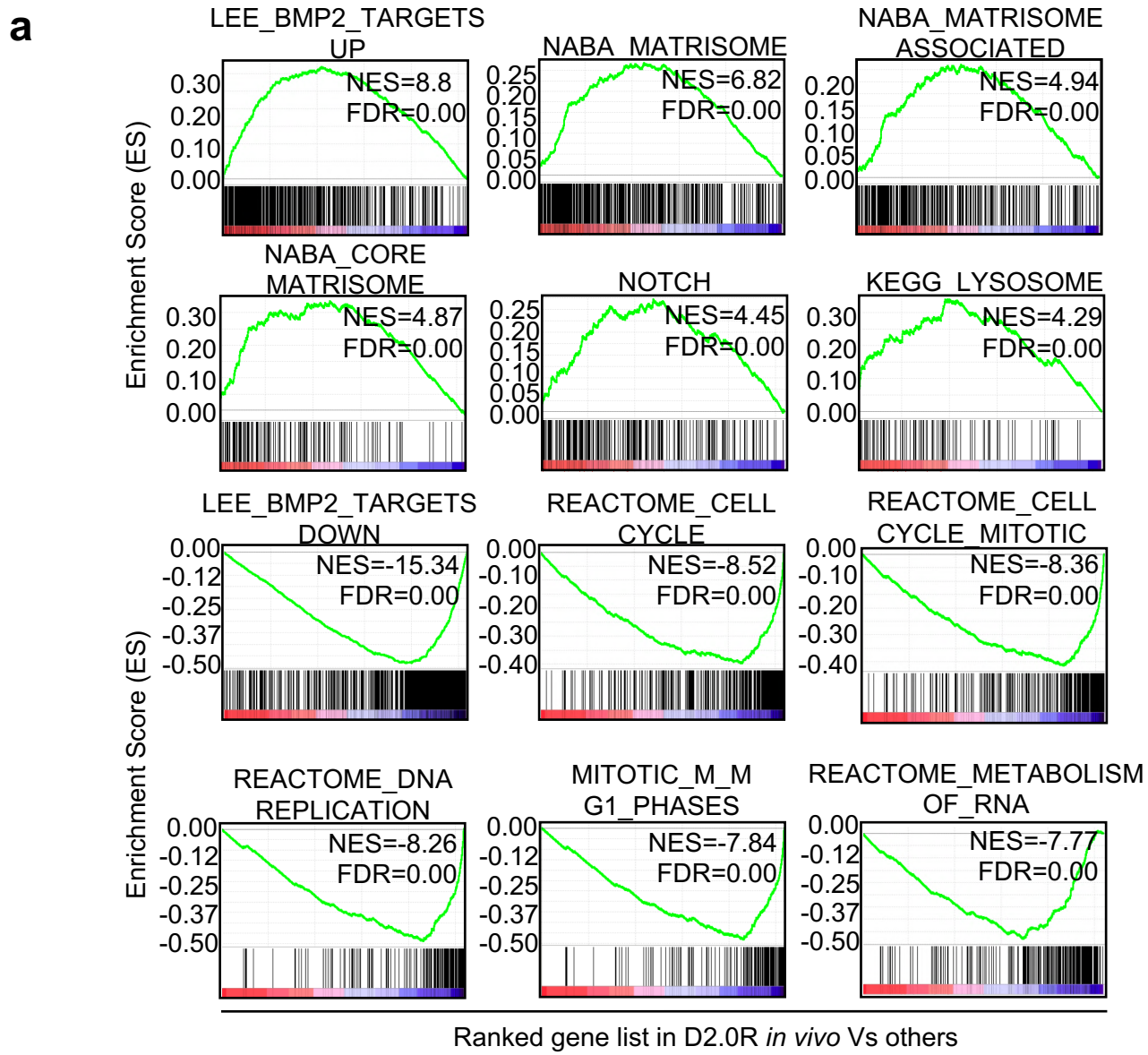

h
Extensive Fibrils
Few Fibrils
No Fibrils
n.s.
P = 0.0001
P = 0.0001
P = 0.0001
Cells with FN fibrils (%)
1.0
0.5
0.0
Control
Control
+ SFKi
+ EGFRI
+ MEKi
+ AT1-like cells
i
MCF7
Cells with protrusions (%)
0.6
0.4
0.2
0.0
P = 0.002
P = 0.0009
Control
Control
+ SFKi
+ AT1-like
j
Relative mRNA levels
2.5
2.0
1.5
1.0
0.5
0.0
ns
DMSO
+SFKi
+ AT1-like
k
F-actin P-SRC Y418
D2.OR + AT1-like
Control
D2.OR + AT1-like
+ EGFRi

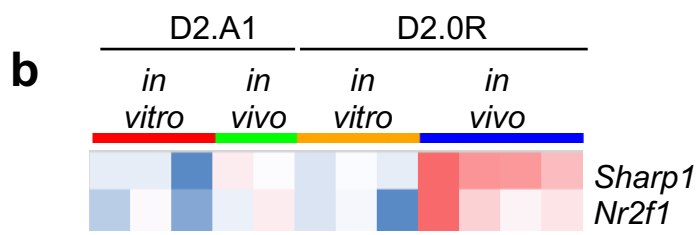
Figure 3

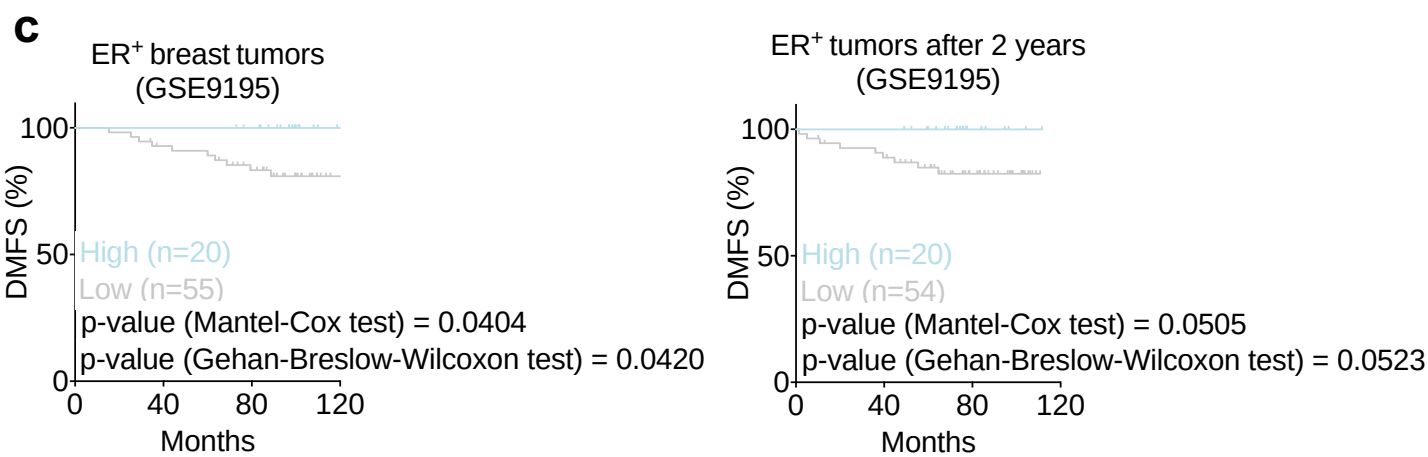

a
D2.A1
D2.0R
in vitro
in vivo
in vitro
in vivo
b
GAG metabolism
EMT
Matrisome
GPCR signalling
Sphingolipid metabolism
Axon development
Wnt signalling
RTK signalling
Cell cycle
TGFβ superfamily signalling
RNA and proteins metabolism
TCA cycle
Cell death
DNA repair and replication
Fatty acids biosynthesis
Notch signalling
Transcription and mRNA processing
YAP pathway
Glycolysis
c
D2.0R in vivo
D2.0R plastic
D2.A1 in vivo
D2.A1 plastic
Relative mRNA levels (log10)
P = 0.0001
P = 0.0001
P = 0.0001
P = 0.0001
COL3A1
POSTN
TNC
LUM
P < 0.0045
P = 0.0001
P < 0.0008
P < 0.0002
SNAIL
SLUG
ZEB2
TWIST1
d
ER+ breast tumors
TAM-treated breast tumors
DMFS (%)
p=0.00022
High (n=302)
Med (n=307)
Low (n=247)
p=0.00058
High (n=99)
Med (n=114)
Low (n=91)
Time (Years)


e
Hazard Ratio (95% CI)
Grade 3
p=0.09
LN
p=0.02
Dormancy Sig.
Low
p=0.02
0 1 2 3 4 5
f
Control
+AT1-like cells
Relative mRNA levels
200
160
120
80
40
10
8
6
4
2
0
P = 0.0006
P = 0.0003
P = 0.0003
P = 0.012
COL3A1
POSTN
TNC
DCN
Relative mRNA levels
30
25
20
15
10
5
0
P < 0.0001
P < 0.0001
P = 0.0026
P < 0.0001
P = 0.0003
ZEB2
WNT5A
SLUG
SNAIL
TWIST1

Supplementary Figure 3







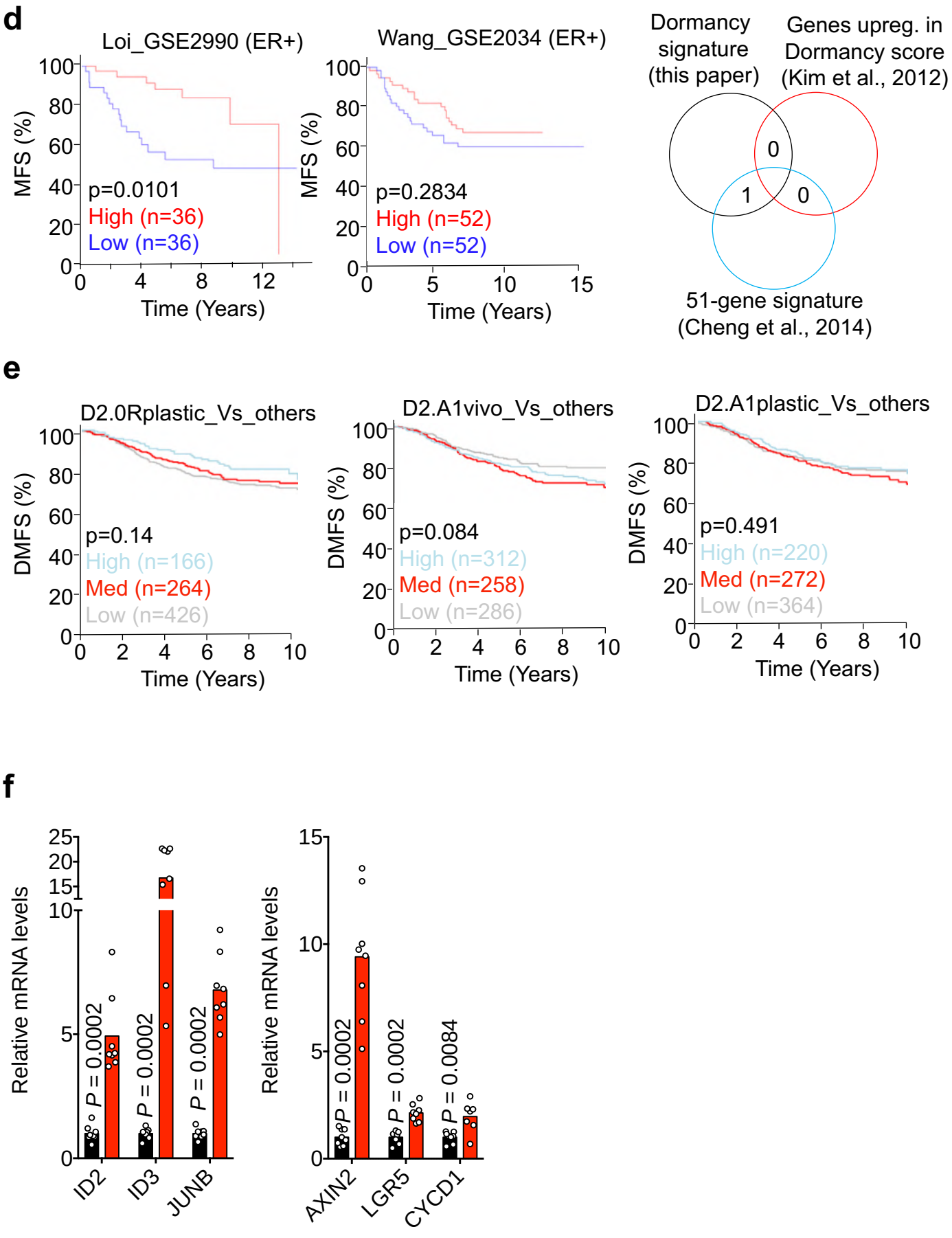
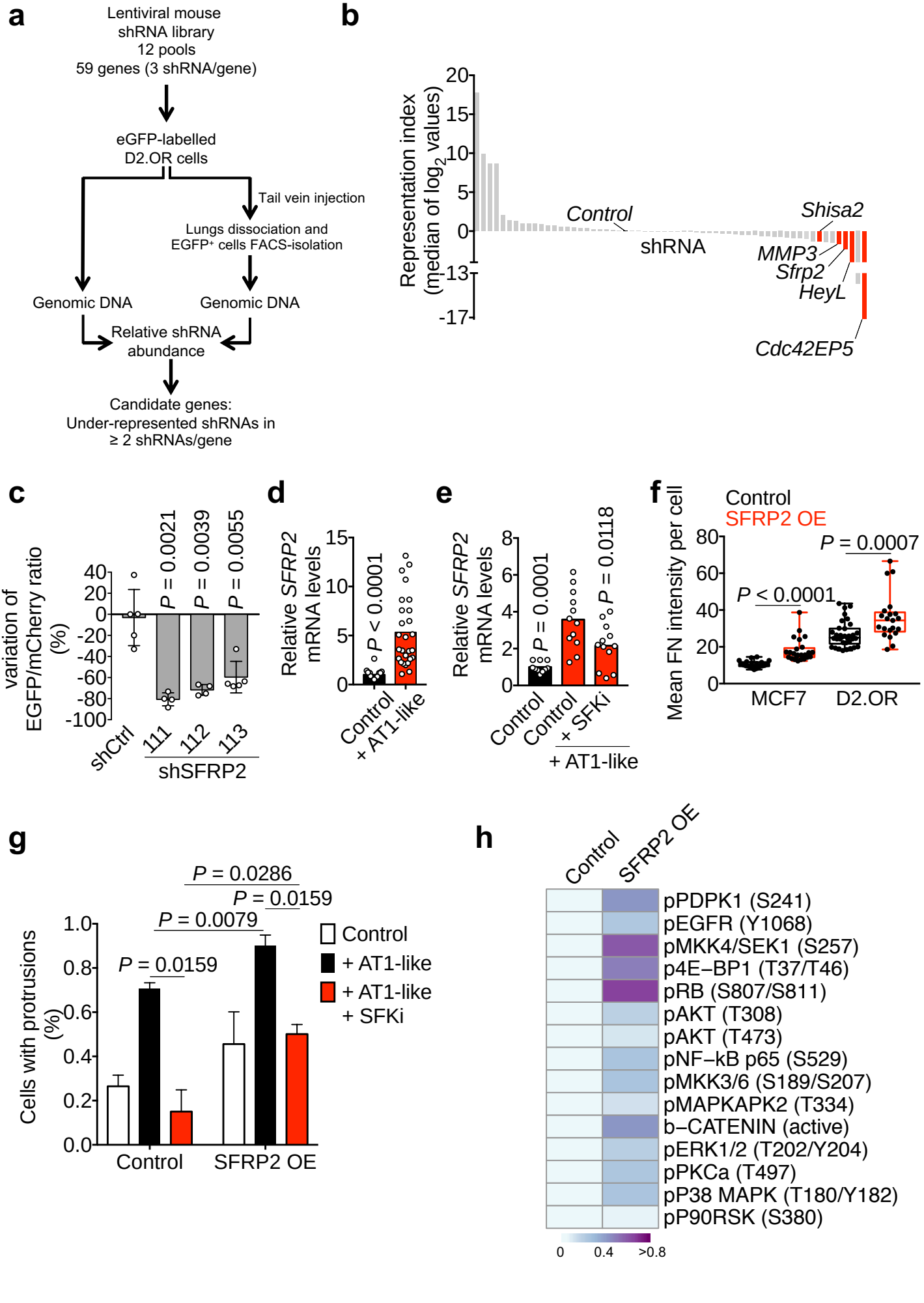
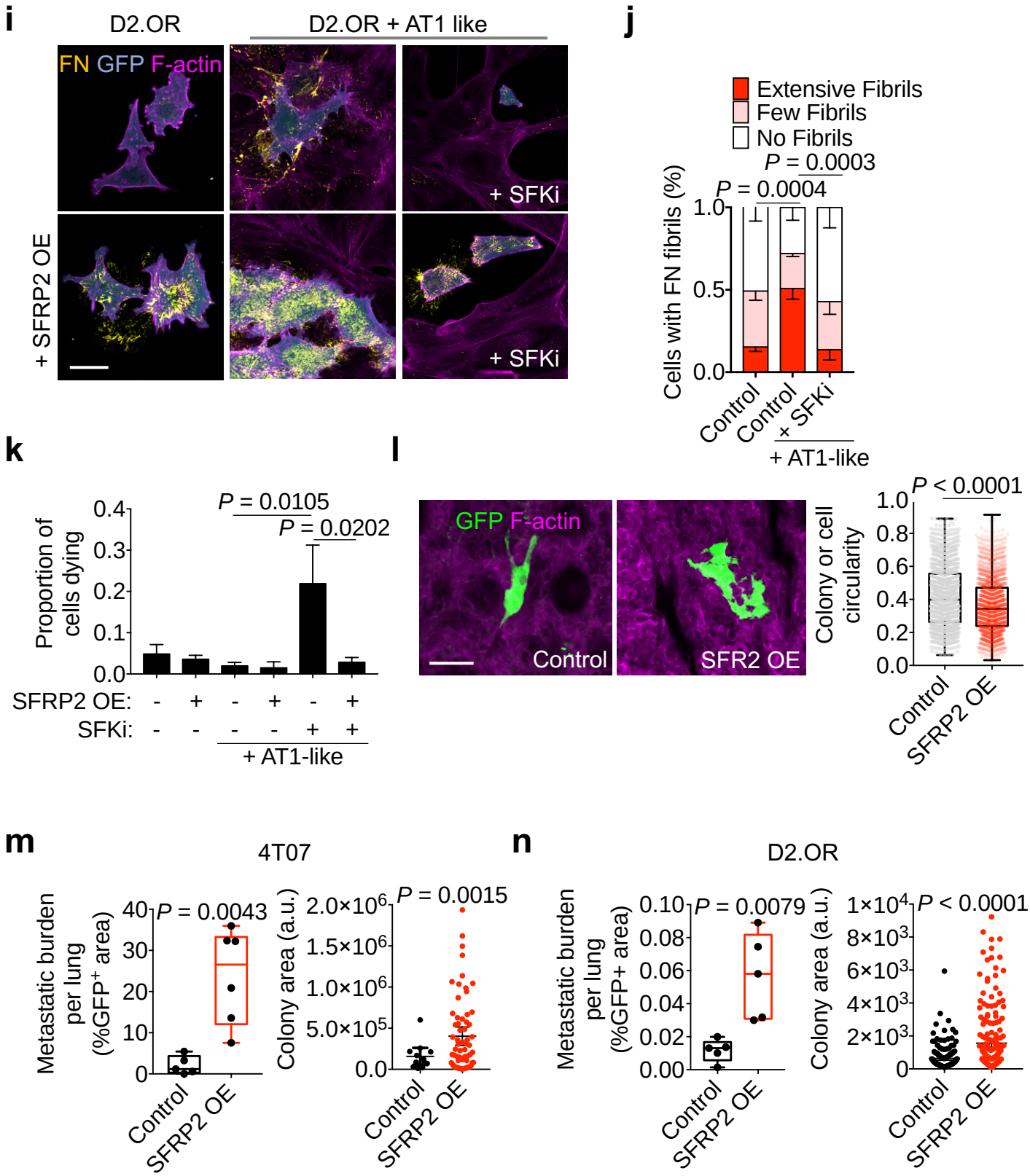
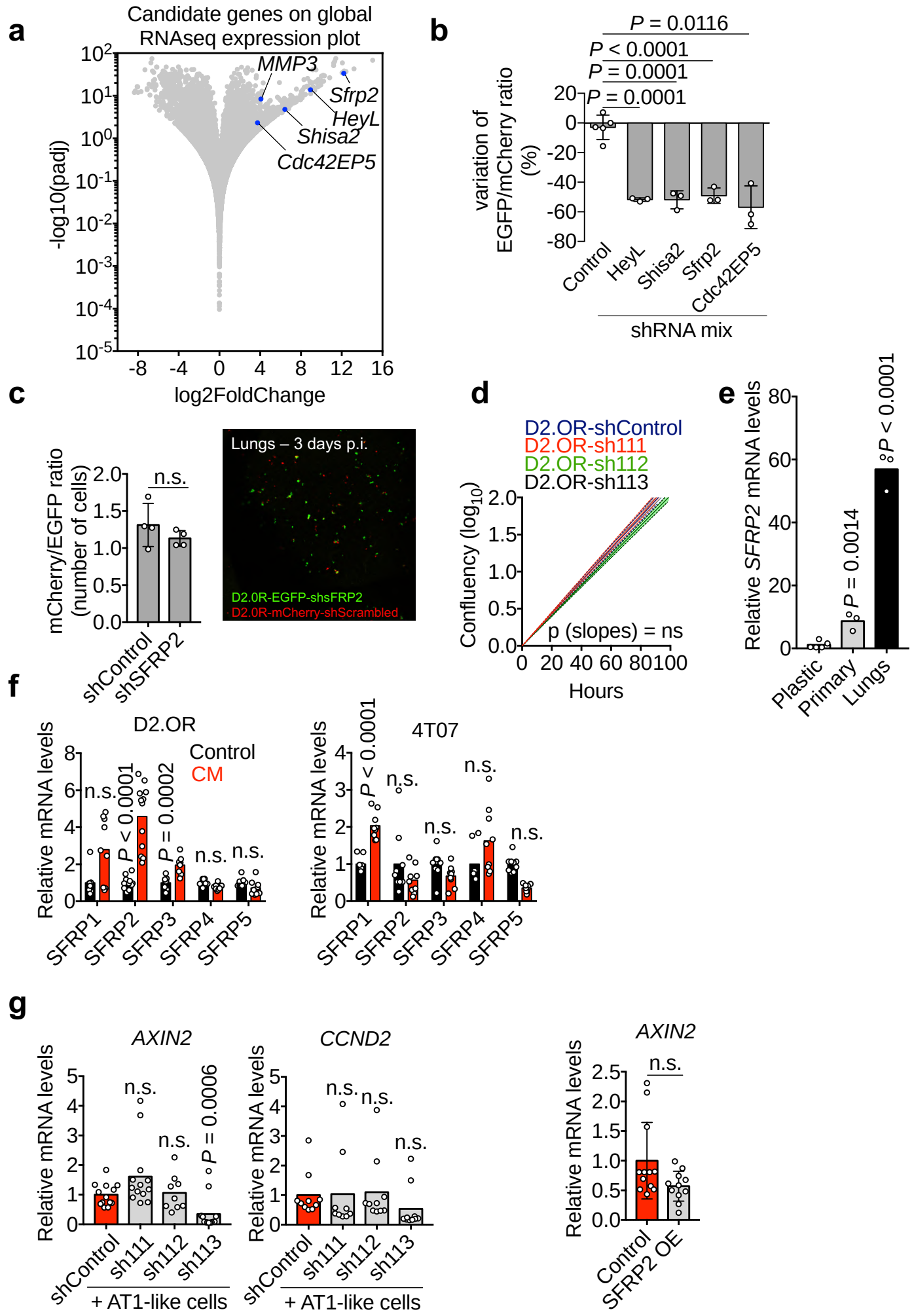

d
Loi_GSE2990 (ER+)
MFS (%)
p=0.0101
High (n=36)
Low (n=36)
Time (Years)
Wang_GSE2034 (ER+)
p=0.2834
High (n=52)
Low (n=52)
Dormancy signature (this paper)
Genes upreg. in Dormancy score (Kim et al., 2012)
0
1
0
51-gene signature (Cheng et al., 2014)
e
D2.0Rplastic_Vs_others
DMFS (%)
p=0.14
High (n=166)
Med (n=264)
Low (n=426)
D2.A1vivo_Vs_others
p=0.084
High (n=312)
Med (n=258)
Low (n=286)
D2.A1plastic_Vs_others
p=0.491
High (n=220)
Med (n=272)
Low (n=364)
f
Relative mRNA levels
P = 0.0002
P = 0.0002
P = 0.0002
ID2
ID3
JUNB
P = 0.0002
P = 0.0002
P = 0.0084
AXIN2
LGR5
CYCD1

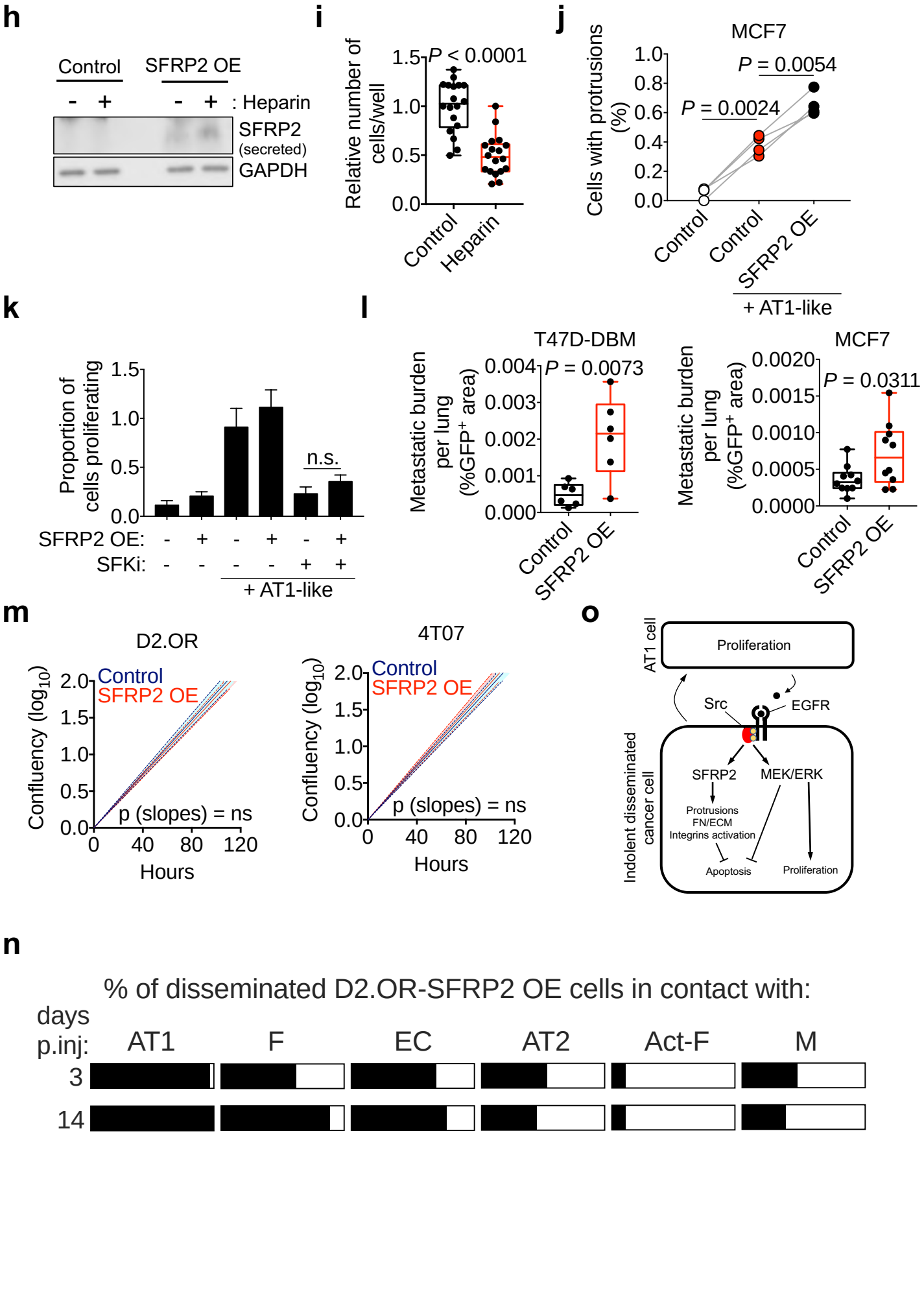
Figure 4


a
Lentiviral mouse shRNA library
12 pools
59 genes (3 shRNA/gene)
eGFP-labelled D2.OR cells
Tail vein injection
Lungs dissociation and EGFP+ cells FACS-isolation
Genomic DNA
Genomic DNA
Relative shRNA abundance
Candidate genes: Under-represented shRNAs in ≥ 2 shRNAs/gene
b
Representation index (median of log2 values)
20
15
10
5
0
-13
-17
Control
shRNA
Shisa2
MMP3
Sfrp2
HeyL
Cdc42EP5
c
variation of EGFP/mCherry ratio (%)
40
20
0
-20
-40
-60
-80
-100
P = 0.0021
P = 0.0039
P = 0.0055
shCtrl
111
112
113
shSFRP2
d
Relative SFRP2 mRNA levels
15
10
5
0
P < 0.0001
Control
+ AT1-like
e
Relative SFRP2 mRNA levels
8
6
4
2
0
P = 0.0001
P = 0.0118
Control
Control + SFKi
+ AT1-like
f
Control
SFRP2 OE
Mean FN intensity per cell
80
60
40
20
0
P < 0.0001
P = 0.0007
MCF7
D2.OR
g
P = 0.0286
P = 0.0159
P = 0.0079
P = 0.0159
Cells with protrusions (%)
1.0
0.8
0.6
0.4
0.2
0.0
Control
+ AT1-like
+ AT1-like + SFKi
Control
SFRP2 OE
h
Control
SFRP2 OE
pPDPK1 (S241)
pEGFR (Y1068)
pMKK4/SEK1 (S257)
p4E–BP1 (T37/T46)
pRB (S807/S811)
pAKT (T308)
pAKT (T473)
pNF–kB p65 (S529)
pMKK3/6 (S189/S207)
pMAPKAPK2 (T334)
b–CATENIN (active)
pERK1/2 (T202/Y204)
pPKCa (T497)
pP38 MAPK (T180/Y182)
pP90RSK (S380)
0
0.4
>0.8


i
D2.OR
D2.OR + AT1 like
FN GFP F-actin
+ SFKi
+ SFRP2 OE
+ SFKi
j
Extensive Fibrils
Few Fibrils
No Fibrils
P = 0.0003
P = 0.0004
Cells with FN fibrils (%)
1.0
0.5
0.0
Control
Control
+ SFKi
+ AT1-like
k
P = 0.0105
P = 0.0202
Proportion of cells dying
0.4
0.3
0.2
0.1
0.0
SFRP2 OE: - + - + - +
SFKi: - - - - + +
+ AT1-like
l
GFP F-actin
Control
SFR2 OE
P < 0.0001
Colony or cell circularity
1.0
0.8
0.6
0.4
0.2
0.0
Control
SFRP2 OE
m
4T07
Metastatic burden per lung (%GFP+ area)
P = 0.0043
40
30
20
10
0
Control
SFRP2 OE
Colony area (a.u.)
P = 0.0015
2.0×10^6
1.5×10^6
1.0×10^6
5.0×10^5
0.0
Control
SFRP2 OE
n
D2.OR
Metastatic burden per lung (%GFP+ area)
P = 0.0079
0.10
0.08
0.06
0.04
0.02
0.00
Control
SFRP2 OE
Colony area (a.u.)
P < 0.0001
1×10^4
8×10^3
6×10^3
4×10^3
2×10^3
0
Control
SFRP2 OE

Supplementary Figure 4




h
Control
SFRP2 OE
- + - + : Heparin
SFRP2
(secreted)
GAPDH
i
Relative number of cells/well
P < 0.0001
Control
Heparin
j
MCF7
Cells with protrusions (%)
P = 0.0054
P = 0.0024
Control
Control
SFRP2 OE
+ AT1-like
k
Proportion of cells proliferating
n.s.
SFRP2 OE:
SFKi:
+ AT1-like
l
T47D-DBM
Metastatic burden per lung (%GFP+ area)
P = 0.0073
Control
SFRP2 OE
MCF7
P = 0.0311
m
D2.OR
4T07
Control
SFRP2 OE
Confluency (log10)
p (slopes) = ns
Hours
o
AT1 cell
Proliferation
Src
EGFR
SFRP2
MEK/ERK
Protrusions
FN/ECM
Integrins activation
Apoptosis
Proliferation
Indolent disseminated cancer cell
n
% of disseminated D2.OR-SFRP2 OE cells in contact with:
days p.inj:
AT1
F
EC
AT2
Act-F
M
3
14