bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

EphB6 regulates TFEB-lysosomal pathway and survival of disseminated indolent breast cancer cells

Manuela Zangrossi[1], Probir Chakravarty[2], Patrizia Romani[1], Colin D.H. Ratcliffe[4], Steven Hooper[4], Martina Dori[3], Mattia Forcato[3], Silvio Bicciato[3], Sirio Dupont[1], Erik Sahai[4,*], Marco Montagner[1,*]

1 Department of Molecular Medicine, University of Padua, Viale G. Colombo, 3, 35126 Padua, Italy

2 Bioinformatics Platform, Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK

3 Department of Life Sciences, University of Modena and Reggio Emilia, Via Giuseppe Campi, 287, 41125 Modena, Italy

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

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

Late relapse of disseminated cancer cells is a common feature of breast and prostate tumors. Several intrinsic and extrinsic factors have been shown to affect quiescence and reawakening of disseminated dormant cancer cells (DDCCs); however, the signals and processes sustaining survival of DDCCs in a foreign environment are still poorly understood. We have recently shown that crosstalk with lung epithelial cells promotes survival of DDCCs of estrogen receptor positive (ER+) breast tumors. Here we show that TFEB-lysosomal axis is activated in DDCCs and that it is modulated by the pro-survival ephrin receptor EphB6. TFEB lysosomal direct targets are enriched in DDCCs *in vivo* and correlate with relapse in ER+ breast cancer patients. Direct coculture of DDCCs with alveolar type I-like lung epithelial cells drives lysosomal accumulation and EphB6 induction. EphB6 contributes to TFEB transcriptional activity and lysosome formation in DDCCs *in vitro* and *in vivo*, and supports survival of DDCCs in coculture and *in vivo*. Furthermore, signaling from EphB6 promotes the proliferation of surrounding lung parenchymal cells *in vivo*.

Introduction

The time required to form overt metastases upon dissemination to a secondary organ varies considerably according to tissue of origin and subtype of the tumor (Risson *et al*, 2020; Klein, 2020). Estrogen receptor positive breast cancers are amongst those cancer types whose latency period can reach 15 years (Zhang *et al*, 2013; Pantel & Hayes, 2018), leading to the question of how DDCCs manage to survive in a foreign environment for such a long time. Intertwined cell-intrinsic (Gawrzak *et al*, 2018; Pavlovic *et al*, 2015) and cell-extrinsic mechanisms sustain adaptation of DDCCs in the foreign environment (Aguirre-Ghiso & Sosa, 2018).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Inflammation (De Cock *et al*, 2016; Albrengues *et al*, 2018), stromal cells (Montagner *et al*, 2020; Malladi *et al*, 2016; Lu *et al*, 2011; Chen *et al*, 2011; Sosnoski *et al*, 2015; Ghajar *et al*, 2013; Wang *et al*, 2015; Ombrato & Montagner, 2020), extracellular matrix proteins and architecture (Barkan *et al*, 2010; Ghajar *et al*, 2013; Montagner *et al*, 2020; Montagner & Dupont, 2020; Montagner & Sahai, 2020; Malanchi *et al*, 2011; Lowy & Oskarsson, 2015; Albrengues *et al*, 2018), diffusible ligands (TGFβ1 (Ghajar *et al*, 2013), BMP (Gao *et al*, 2012), Wnt (Oskarsson *et al*, 2011), Notch (Capulli *et al*, 2019)) and hypoxia (Johnson *et al*, 2016; Fluegen *et al*, 2017) have been shown to regulate intracellular sensors that drive the choice between sustained quiescence and proliferation of DDCCs of breast origin (P-ERK/P-p38 ratio (Aguirre-Ghiso & Sosa, 2018); PI3K/Akt/mTOR pathway (Chen *et al*, 2011; Wang *et al*, 2015; Balz *et al*, 2012); integrins (Montagner *et al*, 2020; Barkan *et al*, 2008; Touny *et al*, 2014; Albrengues *et al*, 2018; Barney *et al*, 2019; Carlson *et al*, 2019)). However, the mechanisms supporting sustained quiescence and survival of DDCCs are still unknown.

Because of the inherent asymptomatic nature of this process, isolation of DDCCs from healthy patients is often technically and ethically not possible, thus, we and others developed *in vitro* systems to study DDCCs-stroma crosstalk (Montagner & Sahai, 2020; Montagner *et al*, 2020). A lung organotypic system, employing a defined combination of lung epithelial cells and fibroblasts, allowed us to recapitulate *in vitro* important features observed in DDCCs *in vivo*, such as Sfrp2-dependent development of filopodia-like structures, fibronectin fibrillogenesis and ultimately survival. Importantly, we showed that pro-survival and growth-restrictive signals emanating from alveolar type I cells (AT1) coexist and are modulated by surrounding stromal cell types and biochemical environment. In the view of a therapy that kills

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

DDCCs before their reawakening, the imperative is targeting pro-survival signals, thus we concentrated on the crosstalk between indolent cancer cells and AT1 cells in a condition of low nutrient medium (mitogen low nutrient low, MLNL), where those signals dominate.

Results and Discussion

To understand crosstalk between DDCCs and lung epithelial cells more fully, we performed transcriptomic profiling and Gene Set Enrichment Analysis (GSEA) of AT1-like lung epithelial cells co-cultured with D2.0R breast cancer cells, which exhibit a DDCC phenotype *in vivo*. We additionally compared this analysis with RNA sequencing of DDCCs from lungs (Montagner *et al*, 2020). Remarkably, a highly significant number of pathways and genes were similarly regulated *in vivo* (lung-disseminated) and in coculture with respect to monoculture (Fig. 1A and Fig. EV1, Pearson correlations equal to 0.811 and 0.559, respectively). Importantly, our transcriptomic analysis captured proliferation of AT1-like cells as well (Fig.1B), as previously observed *in vivo* in response to proximity with dormant cells, indicating that our coculture faithfully models processes observed *in vivo* in DDCCs and stromal cells.

Gene signatures related to lysosomal biogenesis and vesicle transport stood out as a process significantly upregulated after dissemination and upon coculture with AT1-like cells (Fig. 1C). In line with this finding, lysosome accumulation was observed in different DDCCs in coculture compared to monoculture, as visualized by staining with Lysotracker, a specific dye for acidic organelles (Fig. 1D and E). Having established an *in vitro* system that mimics lung-specific *in vivo* processes upon dissemination of DDCCs, we tested if functional lysosomes are required for survival

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of DDCCs in this context. Importantly, blocking lysosome acidification with bafilomycin A1 led to reduced surviving DDCCs in the coculture without affecting stromal cells (Fig. 1F and Fig. EV2).

As TFEB, a member of MiT-TFE family of transcription factors, is a master regulator of lysosome biogenesis and associated metabolic processes (Napolitano & Ballabio, 2016; Lawrence & Zoncu, 2019; Perera & Zoncu, 2016), we queried our transcriptomic analysis with a gene list of 471 TFEB direct targets and with a selection of TFEB direct targets with known lysosomal function (Sardiello *et al*, 2009; Palmieri *et al*, 2011). Both signatures were highly enriched in DDCCs *in vivo* (lung-disseminated) and cocultured dormant cells, suggesting that activated TFEB might be responsible for the observed lysosomal accumulation (Fig. 1G). TFEB transcriptional activation upon coculture was confirmed by a TFEB luciferase reporter in DDCCs (Fig. 1H). Interestingly, higher expression of TFEB direct targets in ER+ breast cancer patients was linked with increased number of relapses at delayed time points (Fig. 1I). These data suggest the activation of TFEB-dependent lysosomal biogenesis in DDCCs *in vivo* and in a lung coculture system and support its involvement in survival of disseminated indolent breast cancer cells.

We previously showed that coculture of D2.0R cells with AT1-like cells induced the transcription of several genes associated *in vivo* with the dormant phenotype. Among those, several were related to ECM deposition and EMT (Montagner *et al*, 2020). We then asked whether direct contact among those cell types was required for activation of these processes. Quantitative PCR with reverse transcription (RT-qPCR) revealed that conditioned medium from AT1-like cells was not sufficient to trigger transcription of representative genes that are instead induced by cell-cell contacts following direct coculture (Fig. 2A). We then sought to identify

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

cell surface signaling molecules that might be involved in communication between lung epithelial cells and indolent breast cancer cells.

An *in vivo* loss-of-function screen identified genes required for the survival of breast DDCCs, such as *Sfrp2*, i.e. genes whose depletion caused death of DDCCs upon dissemination to the lungs (Montagner *et al*, 2020). Among those genes we noticed *EphB6*, a member of the Eph family of receptor tyrosine kinase (Fig. 2B). Interestingly, EphB receptors' ligands, ephrin-Bs, are transmembrane proteins and are thus a good candidate to explain contact-mediated processes (Kania & Klein, 2016; Liang *et al*, 2019; Nikas *et al*, 2018). Importantly, we repeated the results from the screening with multiple short hairpin RNA targeting *EphB6*, validating its relevance in the context of persistence of DDCCs (Fig. 2C).

An additional feature pointing to a role for EphB6 in the communication between lung epithelial cells and DDCCs was the observation that *EphB6* mRNA was upregulated in lung-disseminated DDCCs compared to culture on plastic (Fig. 2D and 2E) in indolent breast cancer cells. As lung parenchyma is characterized by an ECM with low stiffness (Young's modulus of approximately 0,5-2 kPa according to (Booth *et al*, 2012; Liu *et al*, 2010)), we hypothesized that a soft microenvironment could contribute to *EphB6* induction in DDCCs. We tested this hypothesis by assessing *EphB6* expression in indolent breast cancer cells cultivated on substrates with different stiffness. First, *EphB6* was induced when cells are cultured on top of a soft naturally-derived 3D ECM scaffold (Matrigel), compared to ECM-coated stiff plastic substrate (Fig. 2F). Second, we cultivated D2.0R cells on synthetic ECM-coated acrylamide hydrogels of controlled stiffness and confirmed *EphB6* induction at low Young's modulus values (Fig. 2G). Notably, *EphB6* was also found upregulated in breast cancer cells from metastases compared to estrogen receptor

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

positive primary breast cancers (Fig. 2H). Next, we investigated the link between EphB6-dependent gene program and human breast cancer. A signature of genes upregulated in *EphB6*-depleted cells (i.e. genes reflecting a low EphB6 activity) were associated with improved distant metastasis-free survival (DMSF) of ER+ subtypes of breast cancers (Fig. 2I). These results support a model whereby EphB6 is induced *in vivo* in soft microenvironments and has a role in the survival of indolent disseminated breast cancer cells.

Eph-ephrin stimulation is bidirectional and signals are propagated in Eph-expressing cells as well as in ephrin-expressing cells (forward and reverse signaling, respectively (Kania & Klein, 2016; Liang *et al*, 2019; Nikas *et al*, 2018)). We then asked whether EphB6 expression in breast cancer cells could influence gene expression in AT1-like cells. RNA sequencing of cocultured AT1-like cells revealed that depletion of EphB6 in D2.0R cells led to downregulation of several cell cycle-related pathways and upregulation of metabolic and other signaling pathways (Fig. 3A). In agreement with this, less proliferating lung stromal cells were observed in the proximity of EphB6-deficient DDCCs *in vivo* (Fig. 3B and C). We then turned our attention to the role of EphB6 in dormant cancer cells. Pearson correlation coefficients highlighted that a large number of pathways and processes are affected by EphB6 depletion (with two independent short interfering RNAs) both *in vivo* and in coculture compared to control cells (Fig. 3D). Strikingly, lysosomal and other vesicle biogenesis signatures were amongst the processes most strongly affected by EphB6 knockdown in coculture and *in vivo* (Fig. 3E). In line with our findings on the role of TFEB and lysosomal biogenesis in DDCCs shown above, we found that TFEB transcriptional activity was significantly reduced in cells with short interfering RNAs against EphB6, suggesting a connection between EphB6 and TFEB (Fig. 3F).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To better understand the link between EphB6 and lysosomal biogenesis, we exploited our coculture system. We showed that EphB6 knockdown decreases TFEB transcriptional activity (Fig. 4A) and consequently lysosomal accumulation in coculture and *in vivo* (Fig. 4B-D). Importantly, EphB6 is required for the survival of indolent breast cancer cells cultivated together with lung epithelial cells (Fig. 4E), as previously observed *in vivo* (Fig. 2C). Collectively this data suggest that downregulation of EphB6 affects TFEB transcriptional activity, lysosome accumulation and survival of DDCCs.

As the leading cause of cancer-related death, the metastatic process has been the object of intense research in the last decades. However, effective prevention or metastases-specific therapies are still an elusive goal. Metastatic dormancy offers a therapeutic window so far unexploited, and yet processes associated with persistence of DDCCs are still largely unknown (Risson *et al*, 2020; Klein, 2020; Massagué & Obenauf, 2016). Our work suggests that EphB6 plays a critical role in the crosstalk of indolent breast cancer cells with alveolar type I and supports survival of DDCCs *in vivo* and *in vitro*. Interestingly, it has been reported that restoration of EphB6 expression in EphB6-null triple negative breast cancer cells (TNBC), induced cell spreading, actin-dependent protrusions and scattered proliferation of tumor initiating cells (Truitt *et al*, 2010; Toosi *et al*, 2018), in agreement with some features observed in ER+ DDCCs. *EphB6* mRNA level is higher in soft natural and synthetic scaffolds as well as in metastatic lesions (Fig. 2F-H). Moreover, *EphB6* mRNA expression depends also on ECM composition, in addition to stiffness, as evidenced in publicly available RNA sequencing analysis of cells plated on collagenI-enriched Matrigel (Vera-Ramirez *et al*, 2018). This observation may suggest that disseminated

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

clones with higher EphB6 expression, having increased fitness, might participate in metastatic outgrowth. In line with bidirectional Ephs-ephrins signaling, *EphB6* knock-down in DDCCs affected proliferation of neighboring AT1 cells *in vitro* and *in vivo* (Fig. 3A-C). EphB6 reduction is accompanied by reduced TFEB-dependent genes transcription in indolent breast cancer cells from lungs and lung-organotypic system (Fig. 3F) and decreased cytoplasmic area with Lysotracker-positive organelles (Fig. 4B).

Lysosomes are the cellular hub that integrates degradation/recycling of cellular components with stress responses allowing dynamic metabolic adaptation, an essential asset for cells disseminated in a foreign microenvironment. Trafficking routes that funnel into lysosomal degradation include clathrin-dependent and independent endocytosis, phagocytosis and macropinocytosis, macroautophagy, chaperone-mediated autophagy, integrin-mediated scavenging (Muranen *et al*, 2017) and entosis (Overholtzer *et al*, 2007). Interestingly, lysosome regulation has been found key in the modulation of quiescence, proliferation and differentiation of hematopoietic stem cells (Liang *et al*, 2020). Although the specific role of lysosomes has not been addressed before in metastatic dormancy, conflicting data on autophagy have been reported, with opposite phenotypes upon knock-down of different autophagy mediators (Vera-Ramirez *et al*, 2018; La Belle Flynn *et al*, 2019). Our data support the view of lysosomal flux as an essential process for the survival of DDCCs in the lung, however future investigations are required to understand whether this is due to specific cargoes converging on a greater number of lysosomes or a more general role for enhanced lysosomal-mediated turnover. Another important aspect that remains to be elucidated is whether this mechanism is shared among other tissues or other lung stromal cells. For example, *in vivo* labelling experiments

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

revealed de-differentiation and proliferation of the lung-epithelial compartment (of AT2 origin) in the metastatic niche of aggressive lung-disseminated breast cancer cells (Ombrato *et al*, 2019). Whether EphB6 is involved in this specific crosstalk is an open question for future experiments. Data from our work describe a novel regulator of breast DDCCs survival, EphB6, that modulates adaptation to lung microenvironment through TFEB-lysosomal axis, providing potential novel liabilities of disseminated dormant breast cancer cells.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Materials and Methods

Cell lines

Alveolar type 1-like cells (TT1 cells) were a gift from J. Downward (The Francis Crick Institute, London) and were originally provided by T. Tetley (Imperial College, London). T47D-DBM cells were a gift from R. Gomis (Institute for Research in Biomedicine, Barcelona). D2.0R and MCF7-GFP cells were a gift from D. Barkan (University of Haifa). All cells were kept in DMEM with 10% FBS (Thermo Fisher Scientific, 41965-039) and routinely screened for mycoplasma at the Cell Services facility at The Francis Crick Institute or with Universal Mycoplasma Detection kit (ATCC, 30–1012 K).

Stable protein expression and gene knock-down

Generation of D2.0R-EGFP and D2.0R-mCherry has been described in (Montagner *et al*, 2020). shRNA expressing cells were generated with lentiviral transduction. pLKO.1-based plasmids (MISSION, Sigma-Aldrich) were transfected into 293T cells together with packaging plasmids (pMD2, psPAX2). After 2 days, supernatants were collected, filtered through a 0.45 µm filter and added to indicated cells for 2 days before selection with puromycin. List of shRNA sequences is provided in Supplementary Table 1.

Tissue dissociation

Lungs were dissociated to single-cell suspension as in (Montagner *et al*, 2020). Briefly, chopped lungs were digested with digestion solution (PBS buffer with 75 µg/ml TM Liberase (Roche, 05401151001), 75 µg/ml TH Liberase (Roche, 05401127001), 12.5 µg/ml DNAse (Sigma-Aldrich, DN25)) for 1 h at 37 °C on a

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

rocker and homogenised by pipetting. After filtration to remove undigested clumps, red blood cells were lysed with Red Blood Cells Lysis Solution (Miltenyi Biotec, 130-094-183) following the manufacturer's protocol. After washing, cells were resuspended in FACS buffer (PBS, 2 mM EDTA and 3% BSA) and labelled with CD45–APC antibody for 30 min (eBiosciences, 30-F11, 1:400) to avoid contamination from leukocytes during sorting.

In vivo assays and quantifications

The study is compliant with all relevant ethical regulations regarding animal research. All protocols were in accordance with UK Home Office regulations under project license PPL80/2368 and subsequently PPL70/8380, which passed ethical review by the LRI Animal Welfare Ethical Review Board in 2014.

For tail vein injections, cells were resuspended in PBS and 150 µl/mouse injected using a 25G needle. At end point, mice were culled by a schedule 1 method. For quantification of disseminated indolent cells after EphB6 knockdown, 5×10^5 D2.0R-mCherry shControl cells (Sigma-Aldrich, SHC016) were injected into the tail vein of 6- to 8-week-old female nude athymic BALB/c mice together with 5×10^5 D2.0R-EGFP shControl cells or 5×10^5 D2.0R-EGFP shEphB6. Lungs were collected and processed as in (Montagner *et al*, 2020). Number of $CD45^-/EGFP^+$ 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).

Sample preparation for RNA sequencing and qPCR analysis. D2.0R-EGFP-shControl ($1x10^6$ cells/sample) or a mix of D2.0R-EGFP-shEphB6 (#31, #34, #35) were injected in the tail vein of nude athymic BALB/c mice. After two weeks, lungs were harvested and digested into a single-cell suspension as

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

described above. CD45$^-$/EGFP$^+$ cells were sorted (Flow Cytometry Facility at Cancer Research UK-LRI and The Francis Crick Institute) directly into lysis buffer and total RNA was extracted with RNeasy Plus Micro Kit (Qiagen) following the manufacturer's instructions.

Lung organotypic system

Samples preparation for survival analysis and imaging. At day 1, 1,36x10^5 TT1 cells/well were plated onto Lumox 24-multiwell plate (Sarstedt, 94.699.00.14) in MLNL medium (low-glucose DMEM/1% FCS, Thermo Fisher Scientific 21885025). At day 2, breast cancer cells (100 cells/well for survival assays, 500 cells/well for imaging) were plated onto TT1 cell layer in MLNL. For quantification, GFP$^+$ cells were manually counted under an inverted fluorescent microscope after replacing the medium with HBSS.

Samples preparation for RNA sequencing. D2.0R-EGFP-shControl, D2.0R-EGFP-shEphB6#31 or D2.0R-EGFP-shEphB6#35 (1,8x10^5 cells/sample) were plated as above and cocultured with AT1-like cells (1,36x10^6 cells/sample) in MLNL medium (in 60 mm dish) for three days before separation by Fluorescence-Activated Cell Sorter (FACS) based on EGFP signal. Total RNA from three biological replicates/sample was extracted using RNeasy Plus Micro Kit (Qiagen). For bioinformatic analysis, replicate samples for D2.0R-EGFP-shEphB#31 and #35 were grouped together and compared against control sample to identify genes that were differentially expressed relative to control that were common between both shRNAs.

Samples preparation for qPCR analysis. Samples were prepared as for RNA sequencing, except that total RNA was extracted from the whole coculture,

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

retrotranscribed, and mouse genes were amplified by using mouse-specific qPCR primers.

Reverse Transcriptase Real Time PCR (RT-qPCR)

Total RNA was retrotranscribed with dT-primed M-MLV Reverse Transcriptase (Thermo Fisher Scientific, 28025013). qPCR analysis was carried out in a QuantStudio 6 Flex Real-Time PCR System (Thermo Fisher Scientific) with Fast SYBR Green Master Mix (Applied Biosystems 4385612). Gene expression values of EphB6 *in vivo* and cultivated on scaffolds were normalized to GAPDH. Gene expression values from cocultured Vs monocultured D2.0R cells were normalized to GFP expression levels (not expressed AT1-like cells). For RT–qPCR analysis of EphB6 gene in disseminated breast cancer cells *in vivo*, cells were isolated from lungs and total RNA was amplified with Arcturus RiboAmp HS PLUS kit to obtain enough cDNA for RT–qPCR analysis. List of primers used in qPCR is provided in Supplementary Table 2.

Cell culture on natural and synthetic scaffolds

For experiments with natural scaffolds, wells were coated with 100% Matrigel (BD Bioscience) for soft substrate. For stiff substrate, 2% Matrigel was used to coat plastic for 1 hour and then removed.

For experiments with synthetic scaffolds, cells were plated on commercial soft (0.2 KPa, SOFTWELL SW12-COL-0.2 PK) or stiff hydrogels (50 KPa, SW12-COL-50 PK). Cells were harvested for total RNA isolation after 24 hours.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Bioinformatics

RNA sequencing. Before analysis, RNA samples were assessed for quantity and integrity using the NanoDrop 8000 spectrophotometer v.2.0 (Thermo Fisher Scientific) and Agilent 2100 Bioanalyser (Agilent Technologies), respectively. Biological replicate libraries were prepared using the polyA KAPA mRNA HyperPrep Kit and sequenced on Illumina HiSeq 4000 platform, generating ~24 million 100bp single-end reads *per* sample. Read quality trimming and adaptor removal was carried out using Trimmomatic (version 0.36). The RSEM package (version 1.3.30)(Li & Dewey, 2011) in conjunction with the STAR alignment algorithm (version 2.5.2a)(Dobin *et al*, 2013) was used for the mapping and subsequent gene-level counting of the sequenced reads with respect to Ensembl mouse GRCm.38.89 version transcriptome. Normalization of raw count data and differential expression analysis was performed with the DESeq2 package (version 1.18.1)(Love *et al*, 2014) within the R programming environment (version 3.4.3)(R: A language and environment for statistical computing (ISBN 3-900051-07-0) – ScienceOpen). Differentially expressed genes were defined as those showing statistically significant differences (FDR <0.05). Differential gene lists ranked by the Wald statistic were used to look for pathway and selected genesets using the Broad's GSEA software (version 2.1.0) with genesets from MSigDB (version 6)(Subramanian *et al*, 2005) and additional published and custom datasets (Supplementary Table 3). Spearman's rank correlation was used to compare the normalized enrichment scores between comparisons from different experiments to determine which pathways were similarly enriched. Scatterplots (generated using the R base graphics package) shows the correlation between the Wald's statistic (gene level differences from DESEQ2) or the normalized enrichment score (pathway level differences from GSEA) when

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

comparing D2.0R lung-disseminated_Vs_monoculture and coculture_vs_monoculture comparisons. Dot plot (generated using R's ggplot2 package) shows the correlation of NES values generated from GSEA between four indicated comparisons, where the color represents the Spearman correlation and size presents the –log10(*p*-value) of the correlation using the cor.test function. Volcano plot was produced using log2FC and adjusted p-value obtained by differential expression analysis exploiting the "ggscatter" function from ggpubr R package (v. 0.2). Balloon plots were made using "ggballoon" function from ggpubr R package (v. 0.2). For enrichment map, GSEA results from D2.0R versus other groups were visualized using Cytoscape (v.3.6.0) and the enrichment map plug-in (Merico *et al*, 2010). The map has been manually annotated to reduce complexity and redundancy.

Analysis of public datasets of primary and metastatic breast cancer samples. To gain insights into the expression of EphB6 in primary breast cancer and metastases, we analyzed publicly available data from microarray (GSE26338,(Harrell *et al*, 2012)). We downloaded from Gene Expression Omnibus the series matrix of samples analyzed using Agilent Human 1A Oligo UNC custom Microarrays (GPL1390; https://ftp.ncbi.nlm.nih.gov/geo/series/GSE26nnn/GSE26338/matrix/GSE26338-GPL1390_series_matrix.txt.gz) and used data as is. Differentially expressed genes were identified using the Significance Analysis of Microarray algorithm coded in the *samr* R package (Tusher *et al*, 2001). In SAM, we estimated the percentage of false positive predictions (i.e., False Discovery Rate, FDR) with 1,000 permutations and identified as differentially expressed those genes with FDR ≤ 5% and absolute fold change larger than a selected threshold (e.g. ≥ 2) in the comparison of primary tumors and metastases, with either paired and unpaired response types.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Survival analysis. Kaplain-Meier were generated with GOBO online tool (http://co.bmc.lu.se/gobo/gsa.pl) which calculates log-rank P value (Mantel–Cox method). EphB6 activity signature has been generated by selecting the most upregulated genes in coculture in cells with EphB6 knock-down (Supplementary Table 4), i.e. genes that are anti-correlated with EphB6.

Immunofluorescence and imaging

Lysosomes visualisation. Cells were plated onto coverslips in MLNL medium and incubated with 50 nM LysoTracker Red DND-99 (ThermoFisher, L7528) for 30 minutes at 37°C prior to fixation (4% PFA for 12 minutes at room temperature and washed three times in PBS). Coverslips were mounted with ProLong Diamond Antifade Mountant with DAPI (Invitrogen, P36962). For quantification, at least 20 fields were acquired for each condition using the same acquisition settings. Images were analyzed with Fiji software. Percentage of Lysotracker$^+$ cytoplasmic area was calculated according to the formula: Lysotracker$^+$ area/(total cell area-nuclear area)*100. Lysotracker$^+$ area was determined with “Analyze particles” tool applying the same threshold for all the images.

Visualisation of mouse lungs with DDCCs. 5×10^5 D2.0R-EGFP expressing shCtrl, shEphB6#31 or shEphB6#35 cells were injected as indicated above (three mice/cell line). After 4 days, mice were culled and left ventricle perfused with 4% PFA to ensure optimal fixation of inner lung tissue. Lungs were then excised, fixed for 3 hours in 4% PFA and immersed in 30% sucrose for 72 hours. After incubation, lungs were embedded in O.C.T. compound (Histo-Line Laboratories, R0030) for rapid freezing with liquid nitrogen vapor. Frozen material was cut in 10 µm sections, fixed in 4% PFA for 10 minutes at room temperature and, after washes, permeabilized for 15 minutes in 0.2% Triton-X 100 in PBS. Blocking step was performed O/N at 4°C

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with 3% BSA, 0.02% Tween-20 in PBS. Primary and secondary antibodies were incubated in blocking buffer at room temperature for 4 and 1 hours, respectively, in a wet chamber; phalloidin was also added to secondary antibodies. The following antibodies and stain were used: Chicken anti-GFP (Abcam, ab13970, 1:200); Rabbit anti-Ki76 (Spring Bioscience Corp, M3062, 1:100); Alexa Fluor 488 Goat anti-chicken (Life technologies, A11039, 1:200); Alexa Fluor 568 goat anti-rabbit (Life Technologies, cat. A11036, 1:300); Alexa Fluor 647 phalloidin (Life Technologies, A22287, 1:100); Alexa Fluor 647 anti-LAMP2 (Santa Cruz Biotechnology, clone H4B4, sc-18822 AF647, 1:30); Hoechst (Sigma-Aldrich, B1155, 10 µg/ml). Slides were mounted with ProLong Diamond Antifade Mountant with DAPI (Invitrogen, P36962). Images of cells and lungs were acquired with Leica TCS SP8 MP confocal microscope employing the LasX software (40x objective). $Ki67^+$ cells in contact with GFP^+ cells were manually counted.

Reporter assay

D2.0R (parental, EGFP-shCtrl and EGFP-shEphB6 #31 and #35) cells were transfected with Lipofectamine 3000 Transfection Reagent (Invitrogen, L3000001) following the manufacturer's instructions. TFEB transcriptional reporter plasmid (RAGD promoter cloned upstream of luciferase gene, gift from Prof. Graziano Martello, University of Padua)(Malta *et al*, 2017) were transfected together with a plasmid with constitutive expression of Renilla luciferase to normalize for transfection efficiency (Montagner *et al*, 2016). After 6h, $1.8x10^4$ transfected cells were plated both on TT1 layer (coculture) and on plastic (monoculture) in 24-well format. 48h after replating, cells were harvested in Luc lysis buffer (25 mM Tris pH 7.8, 2.5 mM EDTA, 10% glycerol, 1% NP-40) and samples on plastic were diluted 1:5 in Luc lysis buffer to balance the Luciferase/Renilla content compared to coculture. Luciferase

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and Renilla activity were determined in a Tecan plate luminometer with freshly reconstituted assay reagents (0.5 mM D-Luciferin (Sigma-Aldrich, L9504), 20 mM tricine, 1 mM $(MgCO_3)_4Mg(OH)_2$, 2.7 mM $MgSO_4$, 0.1 mM EDTA, 33 mM DTT, 0.27 mM CoA, 0.53 mM ATP for Luciferase reaction, and 4 µg/ml coelenterazine (Invitrogen, C2944) in TBS 1X for Renilla reaction). Each sample was transfected at least in three biological duplicates in each experiment.

Statistical methodology

Normal distribution of data was tested with Shapiro-Wilk test for experiments with sample size greater than ten. For sample sizes lower than ten, it is not easy to assess the underlying data distribution and non-parametric tests were preferred. For samples sizes lower than five, we preferred parametric tests owing to the minimum possible P value becoming large in the non-parametric case. For normally distributed samples, we performed Student's two-tailed t-test for single comparisons (paired or unpaired) and ANOVA test (one-way or two-ways) for multiple comparisons. For non-normal data, we performed two-tailed Mann-Whitney test for single comparisons and Dunn's test for multiple comparisons. Statistical analyses were performed with GraphPad Prism Software. Gene expression derived from microarray data of clinical samples was analysed with Significant Analysis of Microarray method (SAM, see Bioinformatics section). For survival plots (Kaplan–Meier analysis), data were analysed with GOBO (http:// co.bmc.lu.se/gobo/gsa.pl) online tool which calculates log-rank P value (Mantel–Cox method). GSEA is generated from the GSEA online tool (http://software.broadinstitute.org/gsea/index.jsp), which also calculates the two primary statistics of the analysis: NES and false discovery rate (FDR). NES is calculated by normalizing enrichment score to gene-set size; FDR represents an estimated likelihood that a gene set with a given NES represents a false positive.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Data availability

RNAseq data have been deposited at GEO Database and will be available concomitant with publication. Read-only access token is available to reviewers upon request. Other data that support the findings are available upon reasonable request from the corresponding authors.

Acknowledgment

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 Sirio Dupont, and Graziano Martello (University of Padua) for thoughtful discussion and reagents. We are indebted to Flow cytometry, Experimental Histopathology, Bioinformatics and Biostatistics, Biological research, Cell services and Advanced sequencing facilities at the Crick Institute for exceptional scientific and technical support throughout the project. We thank Dr. Mattia Arboit for support with data visualization in R. C.R., S.H., & E.S. are supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC010144), the UK Medical Research Council (FC0010144) and the Wellcome Trust (FC010144). M.M. received funding from Marie Curie Actions—Intra-European Fellowships #625496 and BIRD Seed grant from Department of Molecular Medicine (University of Padua). C.D.H.R. was supported by a postdoctoral training award from Fonds de Recherche du Québec – Santé.

Author contributions

M.Z.: conceptualisation, methodology, data curation, writing - review & editing. M.F., M.D., S.B.: formal analysis and visualisation of microarray data, writing - review &

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

editing. P.C.: formal analysis and visualization of RNA sequencing data. S.D.: conceptualisation, funding acquisition, resources, writing - review & editing. C. D. H. R, S. H.: investigation, writing - review & editing. E.S, M.M.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, visualization, writing - original draft.

Conflicts of interest

The authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

Aguirre-Ghiso JA & Sosa MS (2018) Emerging Topics on Disseminated Cancer Cell Dormancy and the Paradigm of Metastasis. *Annual Review of Cancer Biology* 2: 377–393

Albrengues J, Shields MA, Ng D, Park CG, Ambrico A, Poindexter ME, Upadhyay P, Uyeminami DL, Pommier A, Küttner V, *et al* (2018) Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice. *Science* 361: eaao4227

Balz LM, Bartkowiak K, Andreas A, Pantel K, Niggemann B, Zänker KS, Brandt BH & Dittmar T (2012) The interplay of HER2/HER3/PI3K and EGFR/HER2/PLC-γ1 signalling in breast cancer cell migration and dissemination. *J Pathol* 227: 234–244

Barkan D, El Touny LH, Michalowski AM, Smith JA, Chu I, Davis AS, Webster JD, Hoover S, Simpson RM, Gauldie J, *et al* (2010) Metastatic growth from dormant cells induced by a Col-I-enriched fibrotic environment. *Cancer Research* 70: 5706–5716

Barkan D, Kleinman H, Simmons JL, Asmussen H, Kamaraju AK, Hoenorhoff MJ, Liu Z, Costes SV, Cho EH, Lockett S, *et al* (2008) Inhibition of metastatic outgrowth from single dormant tumor cells by targeting the cytoskeleton. *Cancer research* 68: 6241–6250

Barney LE, Hall CL, Schwartz AD, Parks AN, Sparages C, Galarza S, Platt MO, Mercurio AM & Peyton SR (2019) Tumor cell-organized fibronectin is required to maintain a dormant breast cancer population. *bioRxiv*: 686527

Booth AJ, Hadley R, Cornett AM, Dreffs AA, Matthes SA, Tsui JL, Weiss K, Horowitz JC, Fiore VF, Barker TH, *et al* (2012) Acellular normal and fibrotic human lung matrices as a culture system for in vitro investigation. *American journal of respiratory and critical care medicine* 186: 866–876

Capulli M, Hristova D, Valbret Z, Carys K, Arjan R, Maurizi A, Masedu F, Cappariello A, Rucci N & Teti A (2019) Notch2 pathway mediates breast cancer cellular dormancy and mobilisation in bone and contributes to haematopoietic stem cell mimicry. *British Journal of Cancer* 121: 157–171

Carlson P, Dasgupta A, Grzelak CA, Kim J, Barrett A, Coleman IM, Shor RE, Goddard ET, Dai J, Schweitzer EM, *et al* (2019) Targeting the perivascular niche sensitizes disseminated tumour cells to chemotherapy. *Nat Cell Biol* 21: 238–250

Chen Q, Zhang XH & Massagué J (2011) Macrophage binding to receptor VCAM-1 transmits survival signals in breast cancer cells that invade the lungs. *Cancer cell* 20: 538–549

De Cock JM, Shibue T, Dongre A, Keckesova Z, Reinhardt F & Weinberg RA (2016) Inflammation Triggers Zeb1-Dependent Escape from Tumor Latency. *Cancer Res* 76: 6778–6784

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21

Fluegen G, Alvaro A-V, Wang Y, Padgen MR, Williams JK, Nobre ARR, Calvo V, Cheung JF, J, Bravo-Cordero J, Entenberg D, *et al* (2017) Phenotypic heterogeneity of disseminated tumour cells is preset by primary tumour hypoxic microenvironments. *Nature Cell Biology* 19: 120–132

Gao H, Chakraborty G, Ai L-L, Mo Q, Decker M, Vonica A, Shen R, Brogi E,

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Brivanlou AH & Giancotti FG (2012) The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. *Cell* 150: 764–779

Gawrzak S, Rinaldi L, Gregorio S, Arenas EJ, Salvador F, Urosevic J, Figueras-Puig C, Rojo F, Barrantes I del B, Cejalvo JM, *et al* (2018) MSK1 regulates luminal cell differentiation and metastatic dormancy in ER + breast cancer. *Nat Cell Biol* 20: 211–221

Ghajar CM, Peinado H, Mori H, Matei IR, Evason KJ, Brazier H, Almeida D, Koller A, Hajjar KA, Stainier DYR, *et al* (2013) The perivascular niche regulates breast tumour dormancy. *Nature Cell Biology* 15: 807–817

Harrell JC, Prat A, Parker JS, Fan C, He X, Carey L, Anders C, Ewend M & Perou CM (2012) Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat* 132: 523–535

Johnson RW, Finger EC, Olcina MM, Vilalta M, Aguilera T, Miao Y, Merkel AR, Johnson JR, Sterling JA, Wu JY, *et al* (2016) Induction of LIFR confers a dormancy phenotype in breast cancer cells disseminated to the bone marrow. *Nature cell biology* 18: 1078–1089

Kania A & Klein R (2016) Mechanisms of ephrin–Eph signalling in development, physiology and disease. *Nature Reviews Molecular Cell Biology* 17: 240–256

Klein CA (2020) Cancer progression and the invisible phase of metastatic colonization. *Nature Reviews Cancer* 20: 681–694

La Belle Flynn A, Calhoun BC, Sharma A, Chang JC, Almasan A & Schiemann WP (2019) Autophagy inhibition elicits emergence from metastatic dormancy by inducing and stabilizing Pfkfb3 expression. *Nature Communications* 10: 3668

Lawrence RE & Zoncu R (2019) The lysosome as a cellular centre for signalling, metabolism and quality control. *Nature Cell Biology* 21: 133–142

Li B & Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12: 323

Liang L-Y, Patel O, Janes PW, Murphy JM & Lucet IS (2019) Eph receptor signalling: from catalytic to non-catalytic functions. *Oncogene* 38: 6567–6584

Liang R, Arif T, Kalmykova S, Kasianov A, Lin M, Menon V, Qiu J, Bernitz JM, Moore K, Lin F, *et al* (2020) Restraining Lysosomal Activity Preserves Hematopoietic Stem Cell Quiescence and Potency. *Cell Stem Cell* 26: 359-376.e7

Liu F, Mih JD, Shea BS, Kho AT, Sharif AS, Tager AM & Tschumperlin DJ (2010) Feedback amplification of fibrosis through matrix stiffening and COX-2 suppression. *J Cell Biol* 190: 693–706

Love MI, Huber W & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15: 550

Lowy CM & Oskarsson T (2015) Tenascin C in metastasis: A view from the invasive front. *Cell Adhesion and Migration* 9: 112–124

Lu X, Mu E, Wei Y, Riethdorf S, Yang Q, Yuan M, Yan J, Hua Y, Tiede BJ, Lu X, *et al* (2011) VCAM-1 promotes osteolytic expansion of indolent bone micrometastasis of breast cancer by engaging α4β1-positive osteoclast progenitors. *Cancer cell* 20: 701–714

Malanchi I, Albert S-M, Susanto E, Peng H, Lehr H-A, Delaloye J-F & Huelsken J (2011) Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature* 481: 85–89

Malladi S, Macalinao DG, Jin X, He L, Basnet H, Zou Y, de Stanchina E & Massagué J (2016) Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. *Cell* 165: 45–60

Malta CD, Siciliano D, Calcagni A, Monfregola J, Punzi S, Pastore N, Eastes AN,

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Davis O, Cegli RD, Zampelli A, *et al* (2017) Transcriptional activation of RagD GTPase controls mTORC1 and promotes cancer growth. *Science* 356: 1188–1192

Massagué J & Obenauf AC (2016) Metastatic colonization by circulating tumour cells. *Nature* 529: 298–306

Merico D, Isserlin R, Stueker O, Emili A & Bader GD (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *{PLoS} {ONE}* 5: e13984

Montagner M, Bhome R, Hooper S, Chakravarty P, Qin X, Sufi J, Bhargava A, Ratcliffe CDH, Naito Y, Pocaterra A, *et al* (2020) Crosstalk with lung epithelial cells regulates Sfrp2 -mediated latency in breast cancer dissemination. *Nature Cell Biology*: 1–8

Montagner M & Dupont S (2020) Mechanical Forces as Determinants of Disseminated Metastatic Cell Fate. *Cells* 9: 250

Montagner M, Martello G & Piccolo S (2016) Monitoring smad activity in vivo using the xenopus model system. *Methods Mol Biol* 1344: 245–259

Montagner M & Sahai E (2020) In vitro Models of Breast Cancer Metastatic Dormancy. *Front Cell Dev Biol* 8

Muranen T, Iwanicki MP, Curry NL, Hwang J, DuBois CD, Coloff JL, Hitchcock DS, Clish CB, Brugge JS & Kalaany NY (2017) Starved epithelial cells uptake extracellular matrix for survival. *Nature Communications* 8: 1–12

Napolitano G & Ballabio A (2016) TFEB at a glance. *J Cell Sci* 129: 2475–2481

Nikas I, Ryu HS & Theocharis S (2018) Viewing the Eph receptors with a focus on breast cancer heterogeneity. *Cancer Letters* 434: 160–171

Ombrato L & Montagner M (2020) Technical Advancements for Studying Immune Regulation of Disseminated Dormant Cancer Cells. *Front Oncol* 10

Ombrato L, Nolan E, Kurelac I, Mavousian A, Bridgeman VL, Heinze I, Chakravarty P, Horswell S, Gonzalez-Gualda E, Matacchione G, *et al* (2019) Metastatic-niche labelling reveals parenchymal cells with stem features. *Nature* 572: 603–608

Oskarsson T, Acharyya S, Zhang XHFF, Vanharanta S, Tavazoie SF, Morris PG, Downey RJ, Manova-Todorova K, Brogi E, Massagué J, *et al* (2011) Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nature medicine* 17: 867–874

Overholtzer M, Mailleux AA, Mouneimne G, Normand G, Schnitt SJ, King RW, Cibas ES & Brugge JS (2007) A Nonapoptotic Cell Death Process, Entosis, that Occurs by Cell-in-Cell Invasion. *Cell* 131: 966–979

Palmieri M, Impey S, Kang H, di Ronza A, Pelz C, Sardiello M & Ballabio A (2011) Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways. *Hum Mol Genet* 20: 3852–3866

Pantel K & Hayes DF (2018) Disseminated breast tumour cells: biological and clinical meaning. *Nat Rev Clin Oncol* 15: 129–131

Pavlovic M, Arnal-Estapé A, Rojo F, Bellmunt A, Tarragona M, Guiu M, Planet E, Garcia-Albéniz X, Morales M, Urosevic J, *et al* (2015) Enhanced MAF Oncogene Expression and Breast Cancer Bone Metastasis. *J Natl Cancer Inst* 107

Perera RM & Zoncu R (2016) The Lysosome as a Regulatory Hub. *Annual Review of Cell and Developmental Biology* 32: 223–253

R: A language and environment for statistical computing (ISBN 3-900051-07-0) – ScienceOpen

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Risson E, Nobre AR, Maguer-Satta V & Aguirre-Ghiso JA (2020) The current paradigm and challenges ahead for the dormancy of disseminated tumor cells. *Nature Cancer*: 1–9

Sardiello M, Palmieri M, Ronza A di, Medina DL, Valenza M, Gennarino VA, Malta CD, Donaudy F, Embrione V, Polishchuk RS, *et al* (2009) A Gene Network Regulating Lysosomal Biogenesis and Function. *Science* 325: 473–477

Sosnoski DM, Norgard RJ, Grove CD, Foster SJ & Mastro AM (2015) Dormancy and growth of metastatic breast cancer cells in a bone-like microenvironment. *Clinical and Experimental Metastasis* 32: 335–344

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, *et al* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci {USA}* 102: 15545–15550

Toosi BM, El Zawily A, Truitt L, Shannon M, Allonby O, Babu M, DeCoteau J, Mousseau D, Ali M, Freywald T, *et al* (2018) EPHB6 augments both development and drug sensitivity of triple-negative breast cancer tumours. *Oncogene* 37: 4073–4093

Touny LHHE, Vieira A, Mendoza A, Khanna C, Hoenerhoff MJJ & Green JEE (2014) Combined SFK/MEK inhibition prevents metastatic outgrowth of dormant tumor cells. *Journal of Clinical Investigation* 124: 156–168

Truitt L, Freywald T, John D, Sharfe N & Freywald A (2010) The EphB6 receptor cooperates with c-Cbl to regulate the behavior of breast cancer cells. *Cancer Res* 70: 1141–53

Tusher VG, Tibshirani R & Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *PNAS* 98: 5116–5121

Vera-Ramirez L, Vodnala SK, Nini R, Hunter KW, Green JE, Vera-Ramirez L, Vodnala SK, Nini R, Hunter KW & Green JE (2018) Autophagy promotes the survival of dormant breast cancer cells and metastatic tumour recurrence. *Nat Commun* 9: 1944

Wang H, Yu C, Gao X, Welte T, Muscarella AM, Tian L, Zhao H, Zhao Z, Du S, Tao J, *et al* (2015) The Osteogenic Niche Promotes Early-Stage Bone Colonization of Disseminated Breast Cancer Cells. *Cancer Cell* 27: 193–210

Zhang XH-F, Giuliano M, Trivedi MV, Schiff R & Osborne CK (2013) Metastasis Dormancy in Estrogen Receptor–Positive Breast Cancer. *Clin Cancer Res* 19: 6389–6397

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure legends

Figure 1. Disseminated dormant breast cancer cells activate TFEB-lysosomal axis *in vivo* and in lung-coculture.

A. Scatterplot shows the correlation between the Normalized Enrichment Score of genesets included in Gene Set Enrichment Analysis (GSEA) of D2.0R-EGFP cells disseminated *in vivo* (lungs), in coculture or monoculture as indicated.

B. Enrichment map for lung AT1-like cells upon coculture with indolent breast cancer cells. The map shows gene-set enrichment results of AT1-like cells cocultured with D2.0R-EGFP cells compared with AT1-like cells in monoculture. Node size, genes in pathway; node colour, enrichment score (red indicates enrichment in cocultured AT1-like cells, blue indicates enrichment monocultures of AT1-like cells); edge width, overlap size between connected nodes.

C. Profile of the running ES score for KEGG_LYSOSOME geneset after GSEA of D2.0R-EGFP cells disseminated *in vivo*, in coculture or monoculture as indicated.

D. Percentage of the cytoplasm with positive Lysotracker signal in the indicated indolent breast cancer cells lines upon monoculture or coculture with AT1-like cells. n=20-24 cells for D2.0R; n=21-29 cells for MCF7-EGFP; n=22-24 cells for T47D-DBM-EGFP. Mann-Whitney test. Whisker plot: midline, median; box, 25–75th percentile; whisker, minimum to maximum.

E. Representative images of Lysotracker staining quantified in **D**. Dashed box: area magnified in the middle images. Scale bar: 20 μm.

F. Relative cell number of the indicated indolent breast cancer cells upon coculture with AT1-like cells, after treatment with 4 μM Bafilomycin A for 4 days. n=4 for D2.0R, n=3 for MCF7 and T47D-DBM cells (independent experiments). Paired two tailed t-test, mean with SEM.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

G. Profile of the running ES score for genesets including TFEB direct targets (all targets or subselection of lysosomal genes, as in (Palmieri *et al*, 2011)) after GSEA of D2.0R-EGFP cells disseminated *in vivo*, in coculture or monoculture as indicated.

H. Relative induction of transfected TFEB-luciferase reporter in D2.0R-EGFP cells cocultured with AT1-like cells compared to monoculture. n=3 independent experiments, ratio paired two-tailed t-test, mean with SEM.

I. Kaplan-Meier curve showing Relapse-Free Survival of ER+ breast cancer patients derived from the database at http://co.bmc.lu.se/gobo/gsa.pl, stratified according to the TFEB signature.

Figure 2. EphB6 supports survival of disseminated dormant breast cancer cells.

A. RT-qPCR analysis of representative genes of the dormancy response. D2.0R-EGFP cells have been cultured alone, cocoltured with AT1-like cells (Contact) or with conditioned medium from AT1-like cells (CM). n=3. Two-way ANOVA test, multiple comparisons. Mean with SEM.

B. Left, workflow of loss-of-function screen *in vivo* for survival genes in DDCCs (Montagner *et al*, 2020). Right, Representation scores for each gene included in the screen, calculated from the fold change of representation of each shRNA relative to pre-injection abundance. On the light blue side, there are genes whose downregulation leads to reduced representation of the clones, such as EphB6.

C. D2.0R-EGFP cells stably expressing the indicating shRNA were injected intravenously together with an equal amount of D2.0R-mCherry-shCtrl cells as internal control. After 3 weeks the amount of surviving D2.0R cells was measured and the ratio EGFP/mCherry calculated. n=5 mice for shCtrl cells, n=4 mice for each shEphB6 sequence. One-way ANOVA test. Mean with SD.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

D. Transcriptome analysis (RNA sequencing) of D2.0R-EGFP cells upon dissemination in the lung Vs D2.0R-EGFP in monoculture. DOWN-regulated (Log2 fold-change < -1 and adjusted p-value < 0.01) and UP-regulated (Log2 fold-change > 1 and adjusted p-value < 0.01) genes are indicated in blue and red respectively.

E. qPCR of EphB6 gene in D2.0R-EGFP cells upon dissemination in the lung (n=3 mice) or after monoculture (n=3 samples). Unpaired t-test. Mean with SD.

F. Relative expression of EphB6 gene in D2.0R-EGFP cells cultivated on coated plastic or on Matrigel. n=3 independent experiments, ratio paired two-tailed t-test, mean with SEM.

G. Relative expression of EphB6 gene in D2.0R-EGFP cells cultivated on synthetic hydrogels with indicated stiffness. n=7 samples merged from n=3 independent experiments, unpaired two-tailed t-test.

H. EphB6 expression in ER+ primary breast cancers and metastases from publicly available databases (Harrell *et al*, 2012). q-value after unpaired Significance Analysis of Microarray.

I. Kaplan-Meier curves showing Distant Metastasis-Free Survival of indicated breast cancer patients with indicated cancer subtypes derived from the database at http://co.bmc.lu.se/gobo/gsa.pl, stratified according to genes repressed by EphB6. Black line indicates patients with lower expression of those genes, i.e. with higher EphB6 activity, that is correlated to increased likelihood of distant relapses.

Figure 3. EphB6 is required for the transcriptional responses of disseminated dormant breast cancer cells and lung stromal cells.

A. Balloon plot summarizing GSEA results of AT1-like cells in monoculture and cocultured with D2.0R-EGFP-shCtrl cells or shEphB6 cells (results from two independent shEphB6 sequences, #31 and #35). Balloon size represents the

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

statistical significance (-log10 FDR), while colour indicates the fold-enrichment for each term (NES).

B. Representative images of Ki67$^+$ lung stromal cells surrounding GFP$^+$ disseminated indolent breast cancer cells with control shRNA or shRNAs targeting EphB6. Scale bar, 10 µm.

C. Quantification of images in **B**. Percentage of D2.0R-EGFP cells in contact with the indicated Ki67$^+$ stromal cells. n = 165, 119 and 147 cells for shCtrl, shEphB6#31 and shEphB6#35, respectively, merged from 3 mice/sample. Data showed as parts of a whole.

D. Dot plot shows the correlation of NES values generated from GSEA between four indicated comparisons, where the color represents the Spearman correlation and size presents the –log10(*p*-value) of the correlation.

E. Balloon plot summarizing GSEA results of the indicated comparisons for each indicated genesets. The plot was manually curated to help visualization and shows genesets with higher coherent enrichment in the different conditions. Balloon size represents the statistical significance (-log10 FDR), while colour indicates the fold-enrichment for each term.

F. Profile of the running ES score for genesets including TFEB direct targets (all targets or subselection of lysosomal genes) after GSEA of D2.0R-EGFP cells with shCtrl or shEphB6 either disseminated *in vivo* or in coculture with AT1-like cells.

Figure 4. EphB6 regulates TFEB-lysosomal axis.

A. Relative induction (coculture with AT1-like cells Vs monoculture) of transfected TFEB-luciferase reporter in D2.0R-EGFP cells stably expressing the indicated

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

shRNAs. n=5 (shCtrl and shEphB6#31) or 4 (shEphB6#35) independent experiments. Dunn's multiple comparisons test. Mean with SEM.

B. Percentage of the cytoplasm with positive Lysotracker signal in D2.0R-EGFP cells with indicated shRNAs upon coculture with AT1-like cells, or in monoculture. n = 46 cells for shCtrl on monoculture; n = 44 cells for shCtrl on coculture; n = 38 cells for shEphB6#31; n = 35 cells for shEphB6#35. Dunn's multiple comparisons test. Whisker plot: midline, median; box, 25–75th percentile; whisker, minimum to maximum.

C. Relative mean intensity of lung disseminated shCtrl-, or shEphB6-, D2.0R-EGFP cells. n = 29 cells for shCtrl, n = 13 or 14 cells for shEphB6 across 2 mice/sample. One-way ANOVA. Whisker plot: midline, median; box, 25–75th percentile; whisker, minimum to maximum.

D. Representative images of lysosomal membrane protein LAMP2 in control D2.0R-EGFP cells or cells with EphB6 knockdown after dissemination to lung parenchyma quantified in **C**. Scale bar, 20 μm.

E. Variation of cell number of D2.0R-EGFP upon shEphB6 knock-down in coculture with AT1-like cells. n = 3 independent experiments. One-way ANOVA. Mean with SEM.

Figure Extended View 1.

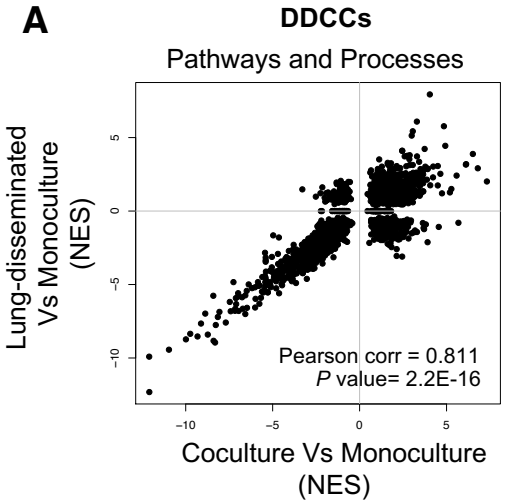
Scatterplot shows the correlation between the Wald's statistic (gene level differences from DESEQ2) of D2.0R-EGFP cells disseminated *in vivo* (lungs), in coculture or monoculture as indicated.

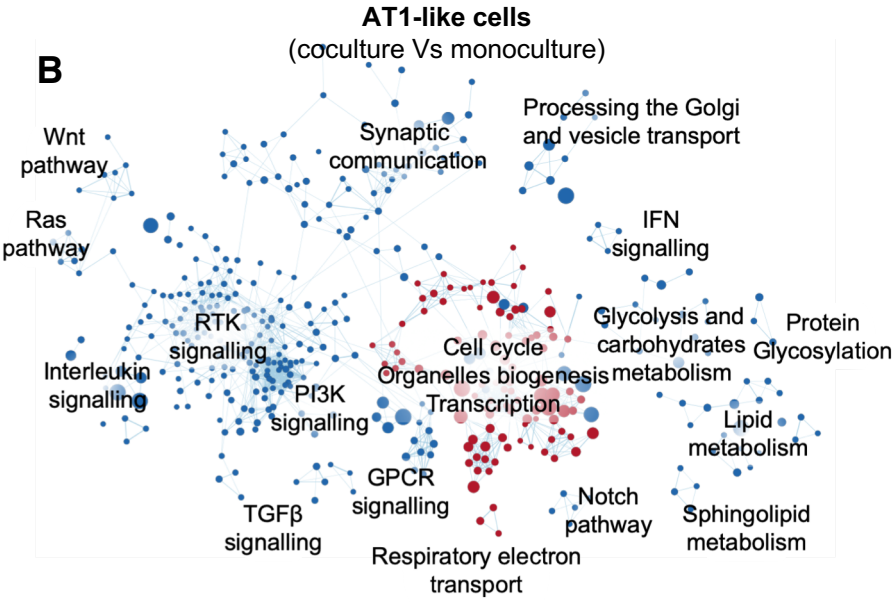
bioRxiv preprint doi: https://doi.org/10.1101/2020.11.10.376186; this version posted November 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

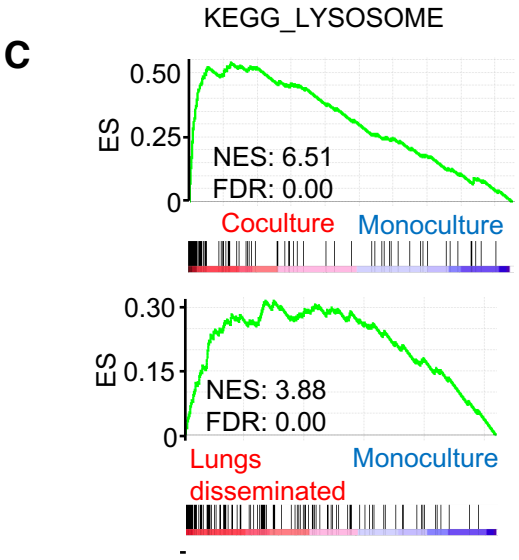
Figure Extended View 2.

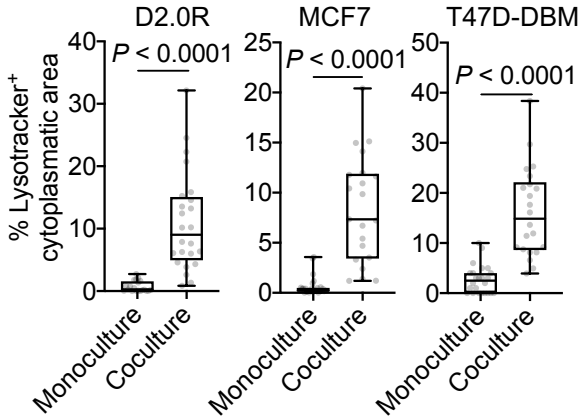
Cell viability, assayed by resazurin staining, of AT1-like cells treated or not with 4 μM Bafilomycin A for 4 days.

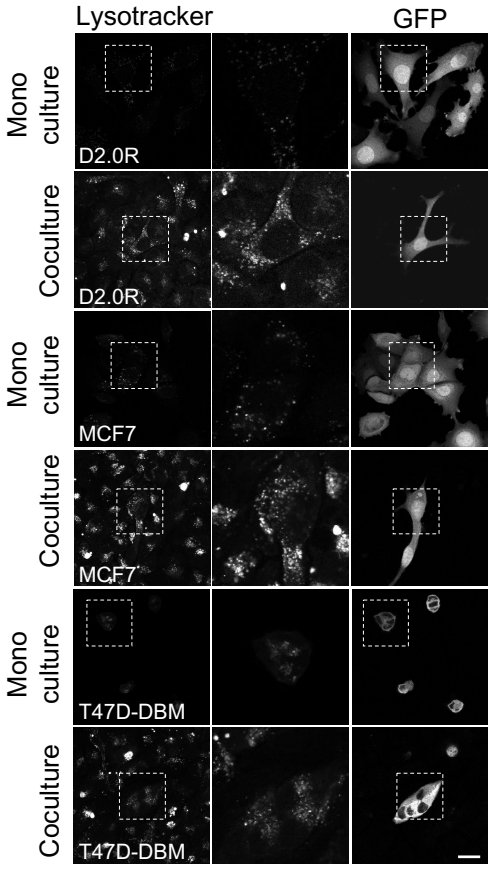
Figure 1

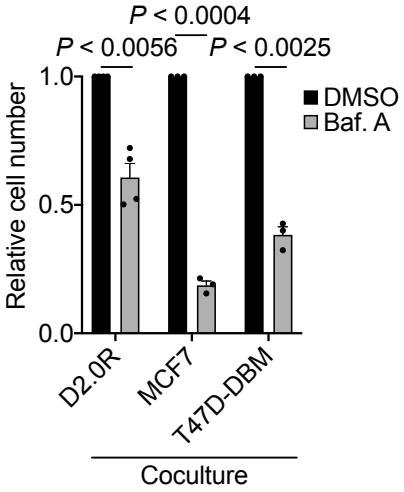

A
DDCCs
Pathways and Processes
Lung-disseminated Vs Monoculture (NES)
Coculture Vs Monoculture (NES)
Pearson corr = 0.811
P value= 2.2E-16

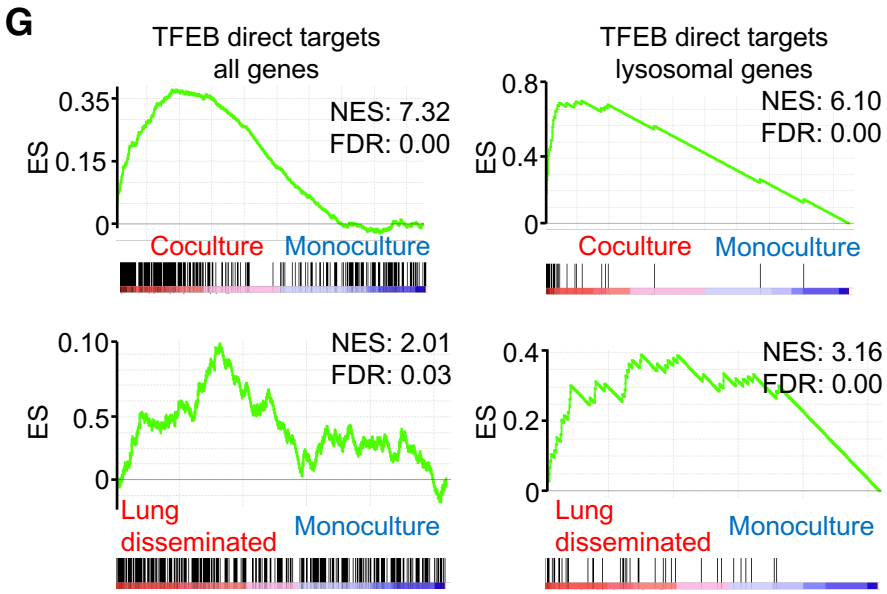

B
AT1-like cells
(coculture Vs monoculture)
Wnt pathway
Ras pathway
Synaptic communication
Processing the Golgi and vesicle transport
IFN signalling
RTK signalling
Interleukin signalling
PI3K signalling
Cell cycle
Organelles biogenesis
Transcription
Glycolysis and carbohydrates metabolism
Protein Glycosylation
Lipid metabolism
GPCR signalling
TGFβ signalling
Notch pathway
Sphingolipid metabolism
Respiratory electron transport

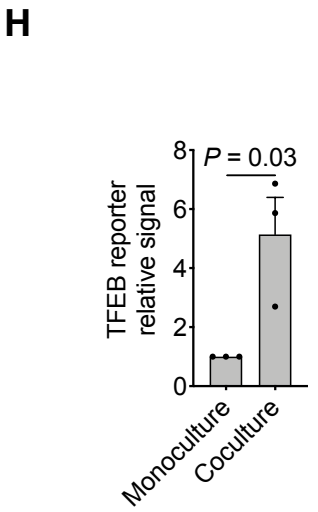

C
KEGG_LYSOSOME
ES
NES: 6.51
FDR: 0.00
Coculture
Monoculture
NES: 3.88
FDR: 0.00
Lungs disseminated
Monoculture


D
D2.0R
MCF7
T47D-DBM
P < 0.0001
% Lysotracker+ cytoplasmatic area
Monoculture
Coculture


E
Lysotracker
GFP
Mono culture
Coculture
D2.0R
MCF7
T47D-DBM


F
P < 0.0004
P < 0.0056
P < 0.0025
Relative cell number
DMSO
Baf. A
D2.0R
MCF7
T47D-DBM
Coculture


G
TFEB direct targets
all genes
ES
0.35
0.15
0
NES: 7.32
FDR: 0.00
Coculture
Monoculture
TFEB direct targets
lysosomal genes
0.8
0.4
0
NES: 6.10
FDR: 0.00
Coculture
Monoculture
0.10
0.5
0
NES: 2.01
FDR: 0.03
Lung
disseminated
Monoculture
0.4
0.2
0
NES: 3.16
FDR: 0.00
Lung
disseminated
Monoculture


H
TFEB reporter
relative signal
8
6
4
2
0
P = 0.03
Monoculture
Coculture

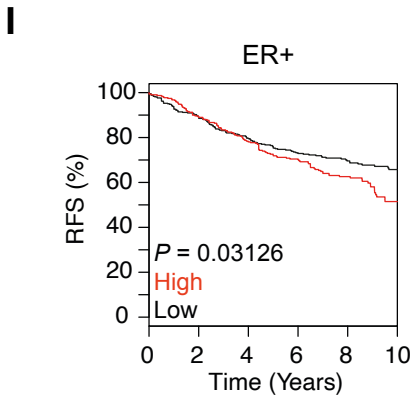

I
ER+
RFS (%)
100
80
60
40
20
0
P = 0.03126
High
Low
0
2
4
6
8
10
Time (Years)

Figure 2

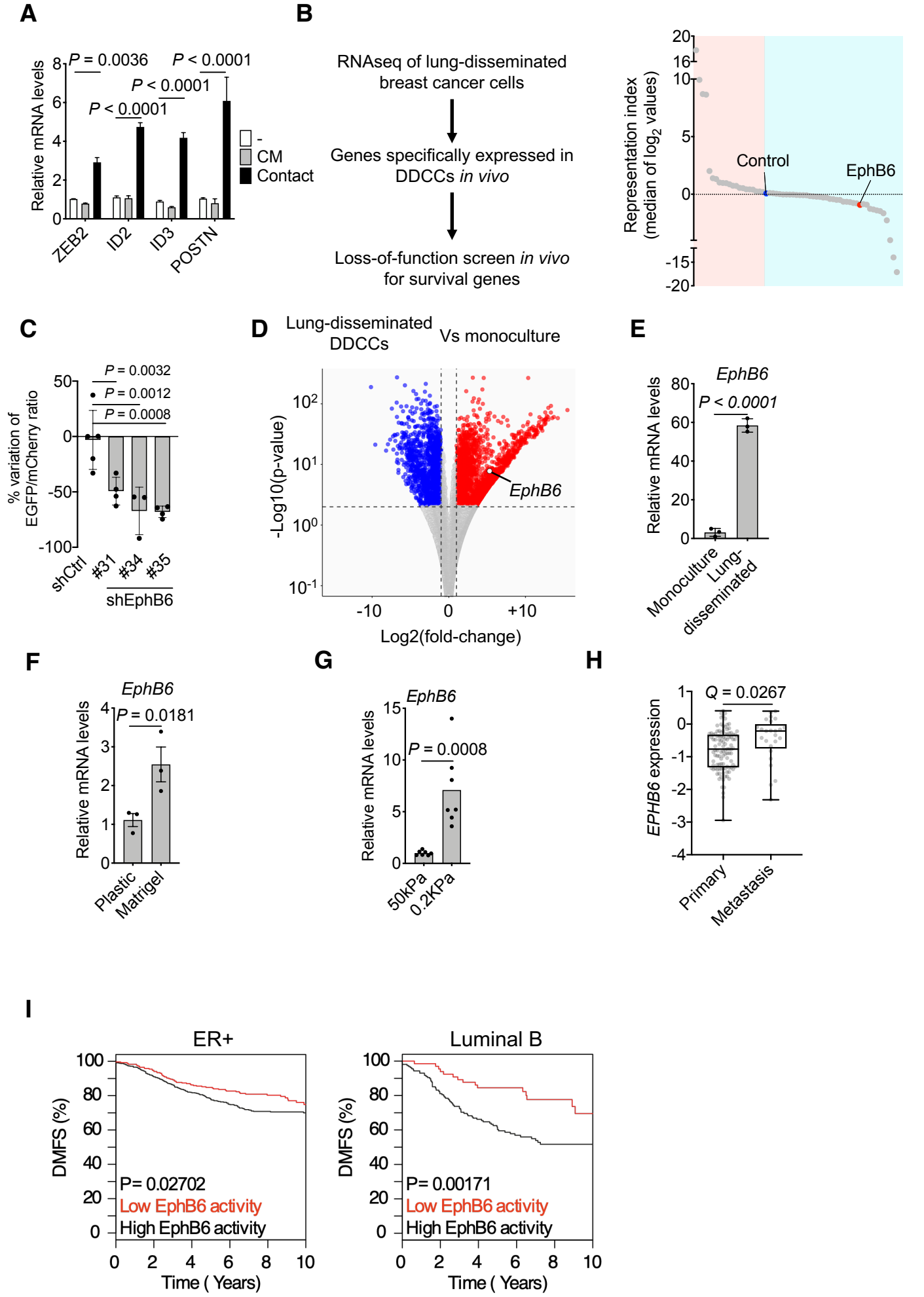
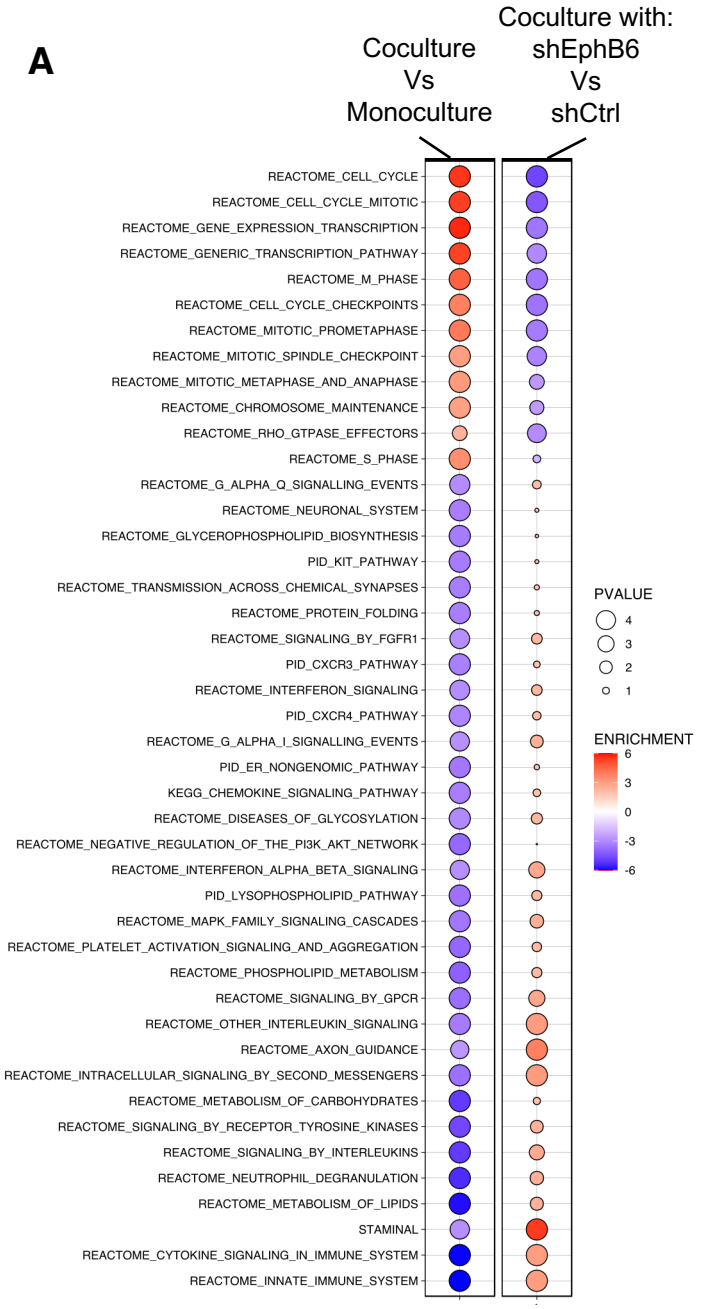

A
Relative mRNA levels
P = 0.0036
P < 0.0001
P < 0.0001
P < 0.0001
-
CM
Contact
ZEB2
ID2
ID3
POSTN
B
RNAseq of lung-disseminated breast cancer cells
Genes specifically expressed in DDCCs in vivo
Loss-of-function screen in vivo for survival genes
Representation index (median of log2 values)
Control
EphB6
C
% variation of EGFP/mCherry ratio
P = 0.0032
P = 0.0012
P = 0.0008
shCtrl
#31
#34
#35
shEphB6
D
Lung-disseminated DDCCs
Vs monoculture
-Log10(p-value)
EphB6
Log2(fold-change)
E
EphB6
P < 0.0001
Relative mRNA levels
Monoculture
Lung-disseminated
F
EphB6
P = 0.0181
Relative mRNA levels
Plastic
Matrigel
G
EphB6
P = 0.0008
Relative mRNA levels
50kPa
0.2KPa
H
Q = 0.0267
EPHB6 expression
Primary
Metastasis
I
ER+
DMFS (%)
P= 0.02702
Low EphB6 activity
High EphB6 activity
Time (Years)
Luminal B
DMFS (%)
P= 0.00171
Low EphB6 activity
High EphB6 activity
Time (Years)

Figure 3

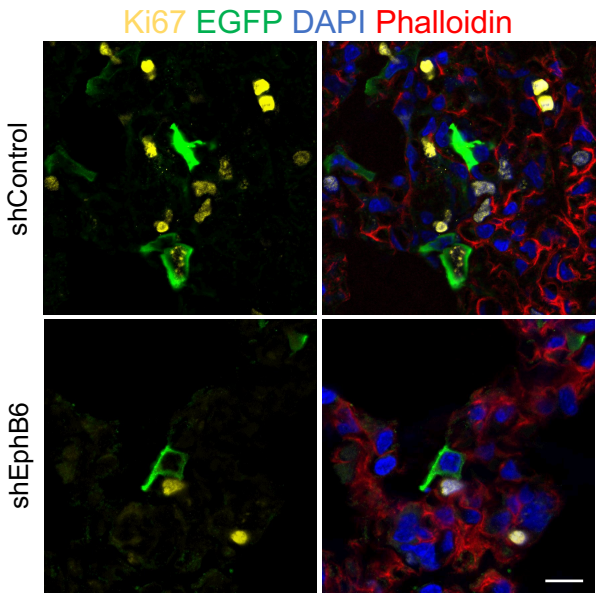
AT1-like cells

A

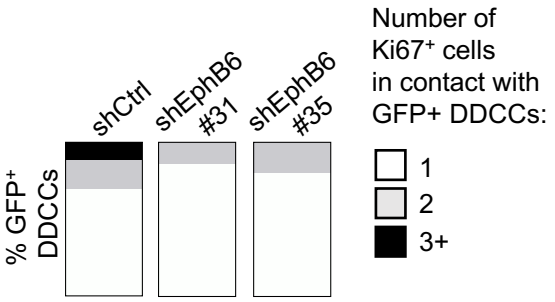


B

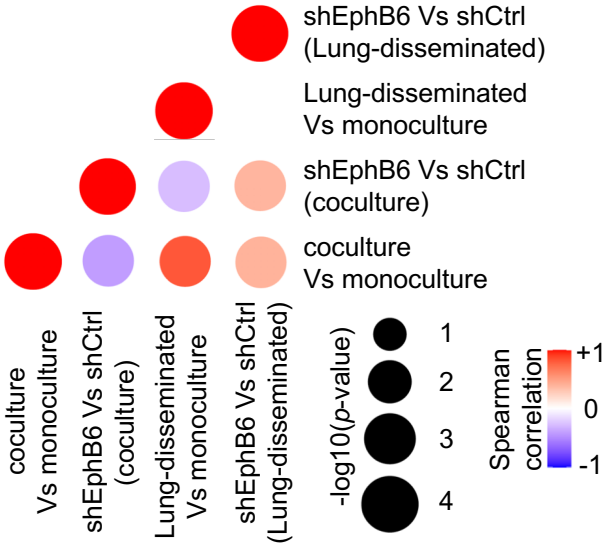
Lung-disseminated



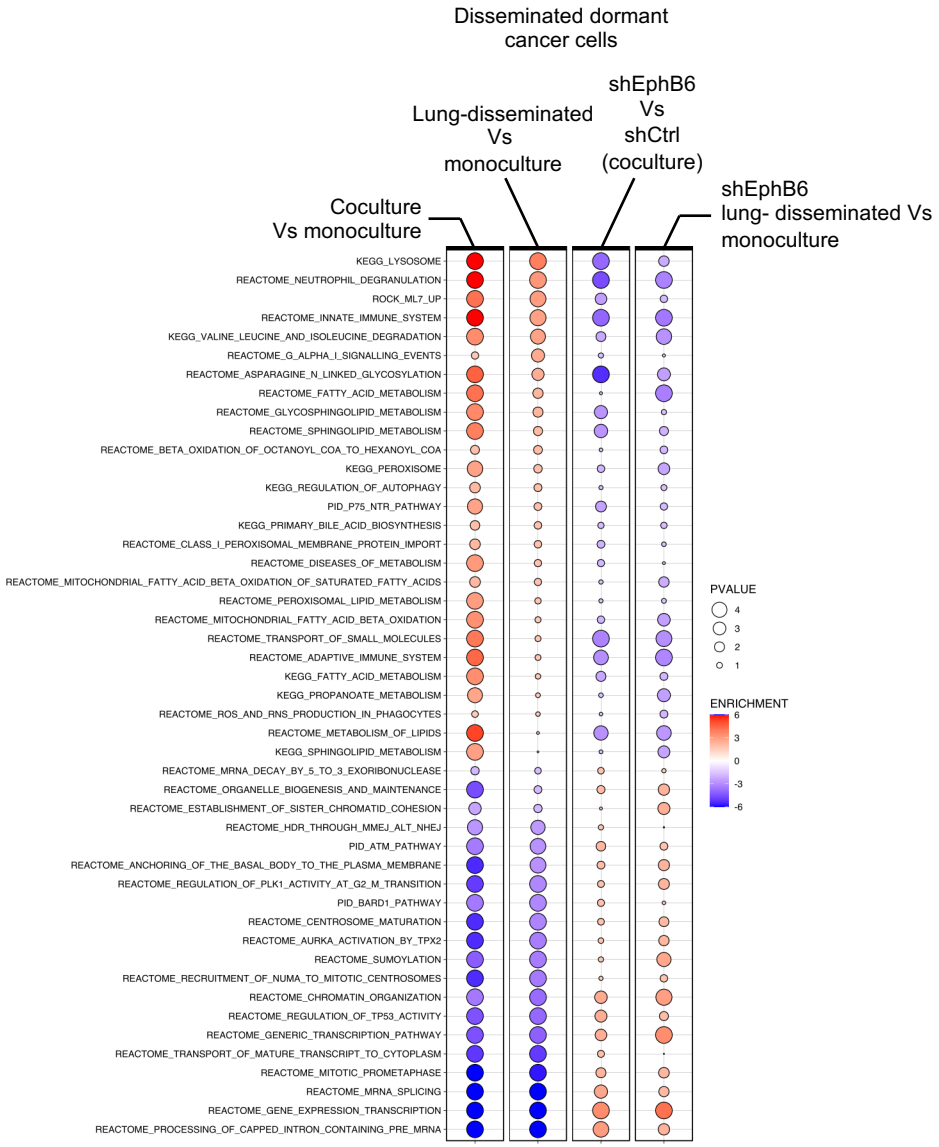
C



D



E


Disseminated dormant cancer cells
Coculture Vs monoculture
Lung-disseminated Vs monoculture
shEphB6 Vs shCtrl (coculture)
shEphB6 lung- disseminated Vs monoculture
KEGG_LYSOSOME
REACTOME_NEUTROPHIL_DEGRANULATION
ROCK_ML7_UP
REACTOME_INNATE_IMMUNE_SYSTEM
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION
REACTOME_G_ALPHA_I_SIGNALLING_EVENTS
REACTOME_ASPARAGINE_N_LINKED_GLYCOSYLATION
REACTOME_FATTY_ACID_METABOLISM
REACTOME_GLYCOSPHINGOLIPID_METABOLISM
REACTOME_SPHINGOLIPID_METABOLISM
REACTOME_BETA_OXIDATION_OF_OCTANOYL_COA_TO_HEXANOYL_COA
KEGG_PEROXISOME
KEGG_REGULATION_OF_AUTOPHAGY
PID_P75_NTR_PATHWAY
KEGG_PRIMARY_BILE_ACID_BIOSYNTHESIS
REACTOME_CLASS_I_PEROXISOMAL_MEMBRANE_PROTEIN_IMPORT
REACTOME_DISEASES_OF_METABOLISM
REACTOME_MITOCHONDRIAL_FATTY_ACID_BETA_OXIDATION_OF_SATURATED_FATTY_ACIDS
REACTOME_PEROXISOMAL_LIPID_METABOLISM
REACTOME_MITOCHONDRIAL_FATTY_ACID_BETA_OXIDATION
REACTOME_TRANSPORT_OF_SMALL_MOLECULES
REACTOME_ADAPTIVE_IMMUNE_SYSTEM
KEGG_FATTY_ACID_METABOLISM
KEGG_PROPANOATE_METABOLISM
REACTOME_ROS_AND_RNS_PRODUCTION_IN_PHAGOCYTES
REACTOME_METABOLISM_OF_LIPIDS
KEGG_SPHINGOLIPID_METABOLISM
REACTOME_MRNA_DECAY_BY_5_TO_3_EXORIBONUCLEASE
REACTOME_ORGANELLE_BIOGENESIS_AND_MAINTENANCE
REACTOME_ESTABLISHMENT_OF_SISTER_CHROMATID_COHESION
REACTOME_HDR_THROUGH_MMEJ_ALT_NHEJ
PID_ATM_PATHWAY
REACTOME_ANCHORING_OF_THE_BASAL_BODY_TO_THE_PLASMA_MEMBRANE
REACTOME_REGULATION_OF_PLK1_ACTIVITY_AT_G2_M_TRANSITION
PID_BARD1_PATHWAY
REACTOME_CENTROSOME_MATURATION
REACTOME_AURKA_ACTIVATION_BY_TPX2
REACTOME_SUMOYLATION
REACTOME_RECRUITMENT_OF_NUMA_TO_MITOTIC_CENTROSOMES
REACTOME_CHROMATIN_ORGANIZATION
REACTOME_REGULATION_OF_TP53_ACTIVITY
REACTOME_GENERIC_TRANSCRIPTION_PATHWAY
REACTOME_TRANSPORT_OF_MATURE_TRANSCRIPT_TO_CYTOPLASM
REACTOME_MITOTIC_PROMETAPHASE
REACTOME_MRNA_SPLICING
REACTOME_GENE_EXPRESSION_TRANSCRIPTION
REACTOME_PROCESSING_OF_CAPPED_INTRON_CONTAINING_PRE_MRNA
PVALUE
4
3
2
1
ENRICHMENT
6
3
0
-3
-6

F

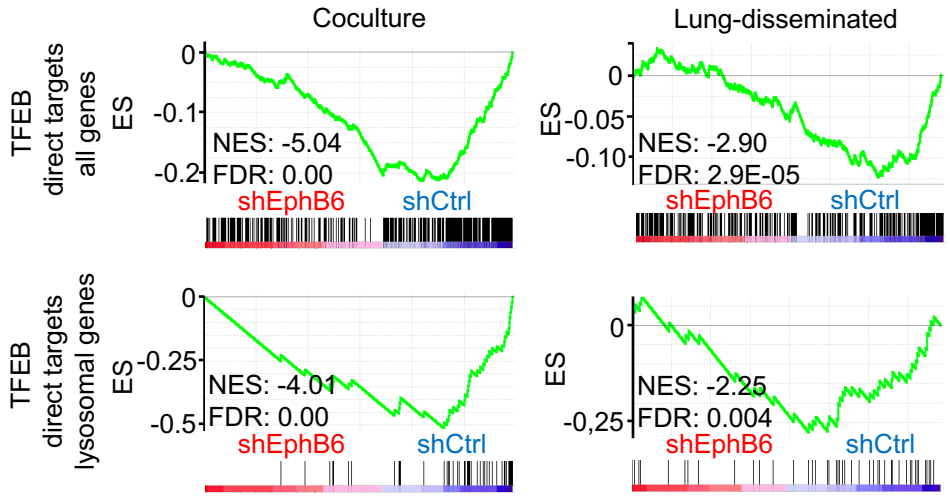

Coculture
Lung-disseminated
TFEB direct targets all genes
ES
0
-0.1
-0.2
NES: -5.04
FDR: 0.00
shEphB6
shCtrl
0
-0.05
-0.10
NES: -2.90
FDR: 2.9E-05
shEphB6
shCtrl
TFEB direct targets lysosomal genes
ES
0
-0.25
-0.5
NES: -4.01
FDR: 0.00
shEphB6
shCtrl
0
-0,25
NES: -2.25
FDR: 0.004
shEphB6
shCtrl

Figure 4

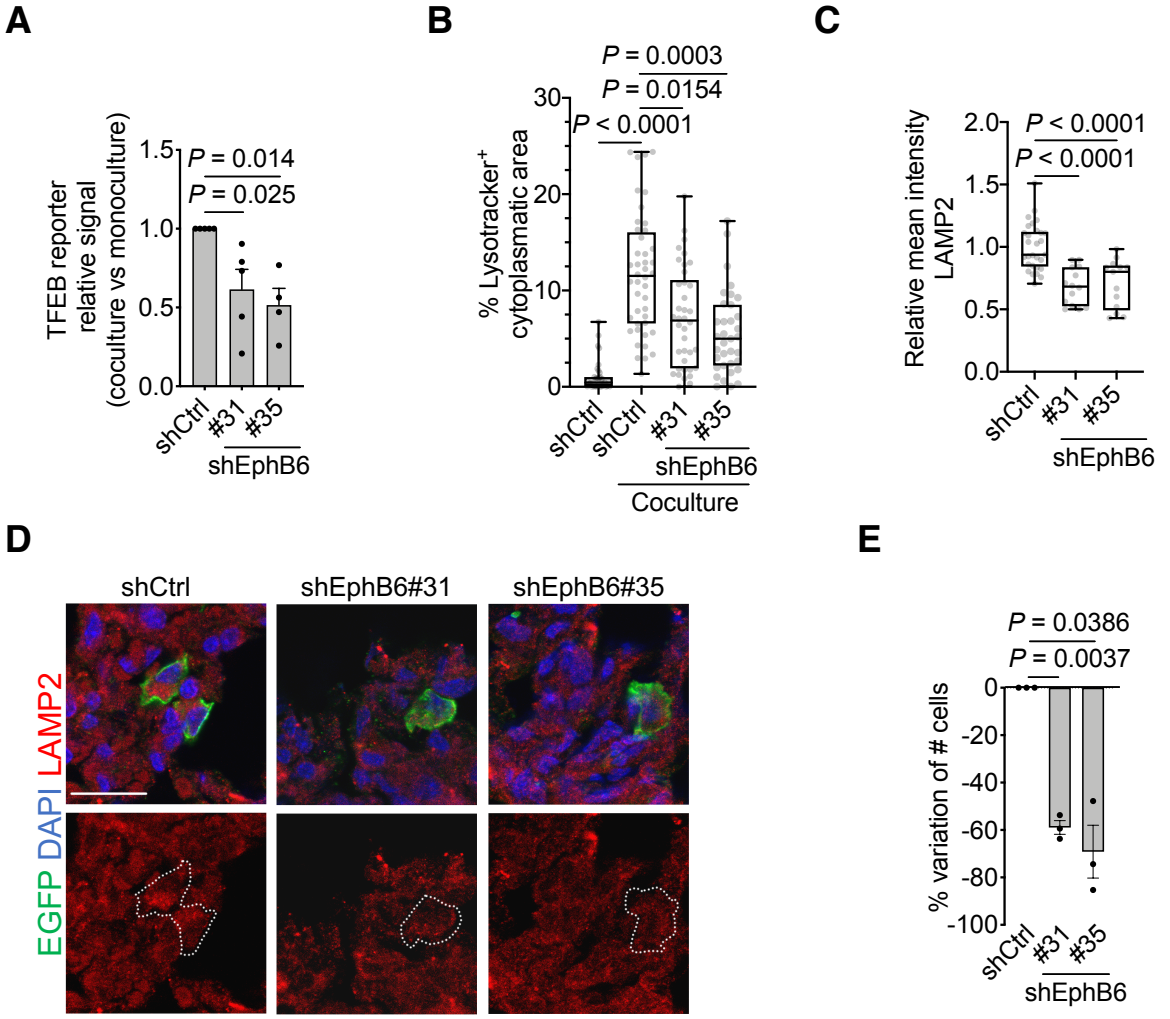

A
TFEB reporter relative signal (coculture vs monoculture)
1.5
1.0
0.5
0.0
P = 0.014
P = 0.025
shCtrl
#31
#35
shEphB6
B
% Lysotracker+ cytoplasmatic area
30
20
10
0
P = 0.0003
P = 0.0154
P < 0.0001
shCtrl
shCtrl
#31
#35
shEphB6
Coculture
C
Relative mean intensity LAMP2
2.0
1.5
1.0
0.5
0.0
P < 0.0001
P < 0.0001
shCtrl
#31
#35
shEphB6
D
shCtrl
shEphB6#31
shEphB6#35
EGFP DAPI LAMP2
E
% variation of # cells
0
-20
-40
-60
-80
-100
P = 0.0386
P = 0.0037
shCtrl
#31
#35
shEphB6

Figure EV1

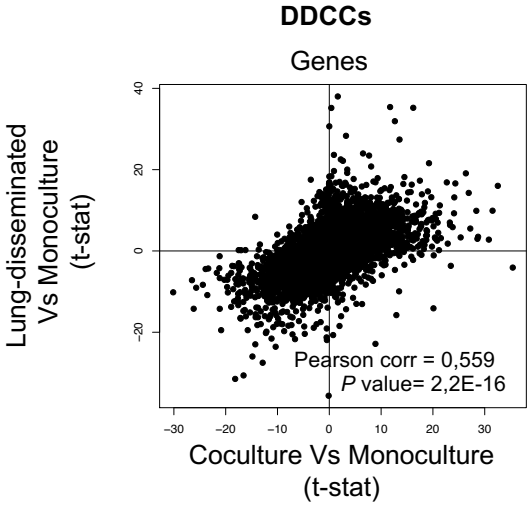
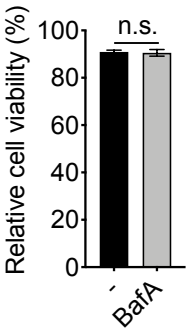



Figure EV2