ARTICLES
https://doi.org/10.1038/s41551-019-0400-9
nature biomedical engineering

A microfluidic assay for the quantification of the metastatic propensity of breast cancer specimens

Christopher L. Yankaskas[1,2], Keyata N. Thompson[3], Colin D. Paul[1,2], Michele I. Vitolo[3,4,5], Panagiotis Mistriotis[1,2], Ankit Mahendra[6], Vivek K. Bajpai[7], Daniel J. Shea[1], Kristen M. Manto[1], Andreas C. Chai[1], Navin Varadarajan[6], Aikaterini Kontrogianni-Konstantopoulos[3,8], Stuart S. Martin[3,4,5] and Konstantinos Konstantopoulos[1,2,9,10]*

The challenge of predicting which patients with breast cancer will develop metastases leads to the overtreatment of patients with benign disease and to the inadequate treatment of aggressive cancers. Here, we report the development and testing of a microfluidic assay that quantifies the abundance and proliferative index of migratory cells in breast cancer specimens, for the assessment of their metastatic propensity and for the rapid screening of potential antimetastatic therapeutics. On the basis of the key roles of cell motility and proliferation in cancer metastasis, the device accurately predicts the metastatic potential of breast cancer cell lines and of patient-derived xenografts. Compared with unsorted cancer cells, highly motile cells isolated by the device exhibited similar tumourigenic potential but markedly increased metastatic propensity in vivo. RNA sequencing of the highly motile cells revealed an enrichment of motility-related and survival-related genes. The approach might be developed into a companion assay for the prediction of metastasis in patients and for the selection of effective therapeutic regimens.

Cancer metastasis is responsible for the vast majority of cancer-related deaths[1]. Localized breast cancer has a 99% five-year relative survival rate, which drops to 85% in patients where the disease has spread regionally, and to 27% in patients with distant metastasis[1,2]. In 2018, approximately 266,000 women will be diagnosed with breast cancer in the United States[2]. Current estimates reveal that 20–30% of patients with early stage breast cancer will eventually experience metastatic recurrence. Exposure of patients at low risk of developing metastasis to aggressive treatments, such as radiotherapy, may compromise the patients' ability to tolerate further treatment that may be necessary to combat de novo cancer in the future[3]. It is estimated that 13,000 women, corresponding to 5% of new diagnoses, will develop de novo metastatic breast cancer in 2018[2]. Thus, it is critical to identify which patients are at risk of developing metastatic disease to provide them with effective treatment while also minimizing the overtreatment of patients who are not at risk with potentially harmful and costly therapies.

Current technologies for the prediction or early detection of breast cancer metastasis are limited to gene expression profiling[4] and the quantification within the patient's bloodstream of circulating tumour cells (CTCs)[5] or of circulating tumour DNA (ctDNA) shed by cancer cells[6]. Gene expression profiles, such as Oncotype DX, measure the expression levels of a subset of genes and use this pattern to predict prognosis, and in some cases, likelihood of responding to treatment. However, it is unlikely that one panel will be effective for all patients because breast cancer progression can be caused by mutations in different pathways, at different levels within the same pathway or even at different loci in the same gene[6]. Owing to the high cost of these tests (>US$3,000), questions remain about the cost-effectiveness of their use in the clinic[7]. Detection of CTCs using the Food and Drug Administration-approved CellSearch system has prognostic value for predicting disease-free survival and overall survival. However, current implementation of the technology still suffers from low sensitivity and specificity for predicting patient outcomes[8]. Detection of ctDNA is typically performed by sequencing primary tumour DNA and then developing PCR probes for unique characteristics of the tumour genome (for example, somatic mutations or chromosomal rearrangement). This approach has been applied in early studies to predict the recurrence or metastasis of breast cancer with high specificity but has been limited by its sensitivity to detect ctDNA (31–80%)[9]. Detection of CTCs or ctDNA is minimally invasive and has the potential to monitor a patient's response to treatment after it is administered, but neither approach can predict whether a patient will respond to specific therapeutic regimens[6]. Improved sensitivity, lead time in prediction of metastasis before its clinical detection and ability to screen therapeutic regimens for patient-specific effectiveness will improve patient outcomes.

Cells within a tumour are heterogeneous; it is believed that only a tiny fraction of cells within a primary tumour is capable of forming metastases[10]. The identification and isolation of these metastasis-initiating cells would enable the prediction of a patient's risk of developing metastasis and the design of optimal, personalized therapeutic treatments. Metastatic cells are bestowed with a repertoire of distinct abilities that enable them to separate from the primary tumour, locally invade the surrounding stroma, intravasate and

[1]Department of Chemical and Biomolecular Engineering, The Johns Hopkins University, Baltimore, MD, USA. [2]Johns Hopkins Institute for NanoBioTechnology, The Johns Hopkins University, Baltimore, MD, USA. [3]Marlene and Stewart Greenebaum National Cancer Institute Comprehensive Cancer Center, University of Maryland School of Medicine, Baltimore, MD, USA. [4]Graduate Program in Molecular Medicine, University of Maryland School of Medicine, Baltimore, MD, USA. [5]Department of Physiology, University of Maryland School of Medicine, Baltimore, MD, USA. [6]Department of Chemical and Biomolecular Engineering, University of Houston, Houston, TX, USA. [7]Department of Chemical and Systems Biology, Stanford University, Stanford, CA, USA. [8]Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore, MD, USA. [9]Department of Biomedical Engineering, The Johns Hopkins University, Baltimore, MD, USA. [10]Department of Oncology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA. *e-mail: konstant@jhu.edu

survive in the circulatory system, roll and arrest on a vessel wall, extravasate, locally invade and form a pro-metastatic niche, and finally proliferate to colonize a distant organ. Clearly, cell migration and survival/proliferation represent integral components of the metastatic cascade[1,11]. In vivo, metastasizing cells migrate through pores in the stromal extracellular matrix (ECM), longitudinal tracks created by ECM degradation (by cancer-associated stromal cells or cancer cells themselves) or formed between the basement membrane and tissue, and along blood vessels or ECM fibrils[11]. The cross-sectional area of these in vivo pores/tracks ranges[12] from 10 to 300 μm^2. Following migration in confinement and arrival at a distant site, metastatic cells must proliferate to generate a secondary colony. We herein developed a Microfluidic Assay for quantification of Cell Invasion (MAqCI), which enables us to concurrently measure the relative abundance of migratory cells and their proliferation state in an effort to identify cells capable of forming metastatic lesions. MAqCI consists of two parallel seeding and collection channels connected by Y-shaped microchannels that mimic aspects of the complexity and variety of cross-sectional areas observed in vivo (Fig. 1a). The Y-shaped channels have a relatively large feeder channel (with a width of 20 μm and a height of 10 μm) to maximize the number of cells that enter the microchannels for study. Cells that reach the far end of the feeder channel encounter a bifurcation region and must choose between two narrower branch channels (with a width of 10 μm or 3 μm and a fixed height of 10 μm). The MAqCI assay uses time-lapse microscopy to perform high-throughput screening of numerous cells in less than 24 h. The assay requires only 50,000 cells to accurately and reproducibly detect the rare subpopulation of metastasis-initiating cells, which may constitute less than 0.1% of the tumour population (≤50 out of 50,000 cells)[13]. Furthermore, migratory cells can be physically isolated from a heterogeneous tumour specimen for molecular and genetic characterization. Altogether, MAqCI has the advantages of requiring a small sample, delivering rapid results, screening the effect of therapeutics on highly motile metastasis-initiating cells and physically isolating these cells for further characterization. We herein demonstrate the potential of MAqCI for diagnosis and precision care in breast cancer.

Results

Prediction of the metastatic potential of breast cancer cells. To determine the ability of MAqCI to distinguish between breast cancer cell lines with high versus low metastatic propensity or normal-like breast epithelial cells, we first examined the migratory potential of a large panel of established cell lines (Supplementary Table 1). Specifically, we analysed cells that entered the feeder channel and classified them into two categories: as non-migratory if their locomotion was limited only to the feeder channel, or as migratory if they reached the bifurcation region and entered one of the branch channels (Fig. 1a and Supplementary Videos 1 and 2). Highly metastatic breast cancer cell lines contain a larger fraction of migratory cells than breast cancer cell lines with low metastatic potential or normal-like breast epithelial cells (Fig. 1b and Supplementary Table 1). These data indicate that the relative abundance of migratory cells in a heterogeneous population correlates with its metastatic potential.

In light of these observations, we sought to determine a threshold percentage of migratory cells that separates cell populations with low versus high metastatic potential. Since the percentage of migratory cells is a function of time (Supplementary Fig. 1a), we optimized both the threshold percentage and experiment duration to maximize the sensitivity, specificity and accuracy of MAqCI. MAqCI has 100% sensitivity to detect cell populations with high metastatic potential at low threshold percentages and long durations (Fig. 1c). At early time points, sensitivity is reduced because migratory cells have not had sufficient time to reach and enter the branch channels. Conversely, 100% specificity is achieved at higher threshold percentages and shorter times (Fig. 1c). At later time points, cells from cell lines with low metastatic potential that struggle to reach the bifurcation and/or squeeze into the narrower branch channels have increased chance to succeed and be classified as migratory. The total accuracy of MAqCI is maximized in the middle of these parameter ranges (Fig. 1c). With a narrow range of threshold percentage (7–9%) and experiment duration (12–14 h), we obtained optimal accuracy (96%) with high sensitivity (89%) and specificity (100%) (Fig. 1c and Supplementary Fig. 1b). This corresponds to a positive predictive value of 100% and a negative predictive value of 96% (Supplementary Fig. 1c,d). This analysis produced one false-negative result, occurring for the SUM149 cell line. The MAqCI assay's receiver operating characteristic (ROC) is plotted for 13 h in Fig. 1d and has 94% area under the curve (AUC), demonstrating the ability to achieve high sensitivity without compromising specificity.

To compare the performance of the MAqCI assay to that of conventional migration assays, we also evaluated this panel of cell lines in a transwell-migration assay using the xCELLigence RTCA DP Instrument with CIM-Plate 16 chambers. These plates have chambers that are very similar to Boyden chambers with an average pore diameter of 8 μm and are equipped with electrodes that measure cell migration via changes in electrical impendance, thereby allowing for measurements in real time. The 8-μm-diameter pores were selected for two reasons: (1) these pores have a cross-sectional area of approximately 50 μm^2, which is within the range of 30–100 μm^2 of MAqCI's branch channels; (2) cell passage through 3 μm, but not 8 μm, pores causes significant DNA damage, which portends genome variation[14,15], thereby potentially altering the molecular signature and phenotype of cells. The readout of cell migration from this instrument, cell index, was not strikingly different between breast cancer cell lines with low versus high metastatic potential (Supplementary Fig. 1e). Analysis of the experiment duration and cell index threshold value to optimize the sensitivity, specificity and accuracy of this assay is shown in Supplementary Fig. 1f,g. The ability of transwell migration to predict metastatic potential had a significantly reduced accuracy of 72% (Supplementary Fig. 1f,g) compared with 96% for migration alone in MAqCI. This reduced accuracy corresponded to similar specificity (94–100% for transwell compared with 100% for MAqCI), but markedly reduced sensitivity (22–33% for transwell compared with 89% for MAqCI) (Supplementary Fig. 1f,g). The trade-off between optimizing sensitivity versus specificity is greater for the transwell-migration assay, corresponding to the reduced AUC (70%) of the ROC (Supplementary Fig. 1h) compared with 94% for MAqCI. Altogether, our data reveal that the MAqCI assay is better suited to identify cell lines with low versus high metastatic potential.

To further improve the predictive power of the MAqCI assay, we examined the benefit of incorporating an additional index distinct from migration into our analysis. Ki-67 is a protein found in the nucleus of actively proliferating cells that is already used clinically to evaluate the prognosis of patients with breast cancer[16,17]. Because proliferation is necessary for the establishment of new metastatic colonies, we examined the percentage of Ki-67-positive cells in MAqCI using immunofluorescence (Supplementary Fig. 1i) and found that cell lines with high metastatic potential tend to have a higher percentage of Ki-67-positive cells (Fig. 1e and Supplementary Table 1). The percentage of Ki-67 cells has satisfactory predictive power (Fig. 1f), with a maximum accuracy of 88% (Supplementary Fig. 1j). Although the majority of cell lines were cultured in Dulbecco's modified Eagle medium (DMEM), a few were cultured in other media as recommended by the American Type Culture Collection. However, neither the percentage of Ki-67-positive cells nor the percentage of migratory cells was altered by the different media (Supplementary Fig. 1k,l). In view of these findings, we next evaluated the combination of Ki-67 as a proliferation index

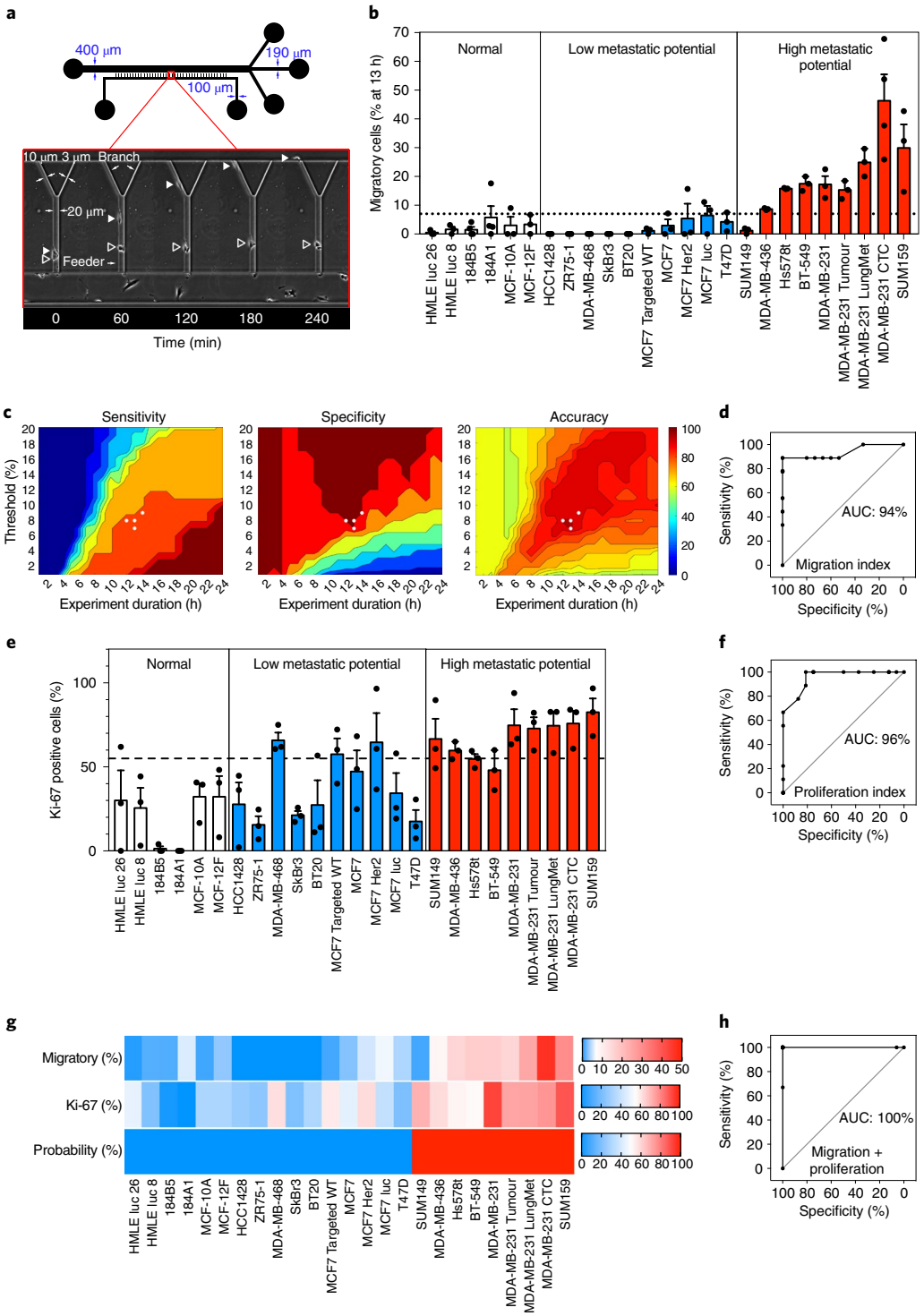

Fig. 1 | Use of MAqCI for prediction of metastatic potential of breast epithelial and breast cancer cell lines with high accuracy, sensitivity and specificity. a, Schematic of MAqCI. Inset: migratory (closed triangle) and non-migratory (open triangle) MDA-MB-231 cells. **b**, Percentage of migratory cells from normal-like breast epithelial and breast cancer cell lines at 13 h. Each data point represents the percentage from one experiment. Column and error bars represent mean ± s.e.m of $n \geq 3$ independent experiments. Dotted line designates 7% of migratory cells. **c**, Sensitivity, specificity and accuracy (%) of prediction of metastatic potential of established cell lines based on the mean percentage of migratory cells. Metrics are calculated as a function of experiment duration (0–24 h) and threshold (0–20%) above which a cell line is predicted to have high metastatic potential. White circles represent the points corresponding to maximum accuracy (96%) for the motility index. **d**, ROC of metastatic potential predictions based only on migration index. AUC is indicated as a percentage. **e**, Percentage of Ki-67-positive cells observed in MAqCI. Each data point represents the percentage from one experiment. Column and error bars represent mean ± s.e.m of $n \geq 3$ independent experiments. Dashed line designates 55% of Ki-67-positive cells. **f**, ROC based only on proliferation index. **g**, Probability (%) of each cell line possessing high metastatic potential calculated using logistic regression and the percentages of migratory and Ki-67-positive cells. Baseline value (shown in white) for each heat map is set to the threshold value for each predictor (7% migratory cells and 55% Ki-67-positive cells). **h**, ROC of MAqCI using the combined migration and proliferation indices.

with the percentage of migratory cells as a motility index. We incorporated both indices as predictors in a logistic regression formula to predict the probability of a cell line having high metastatic potential (Fig. 1g). This probability correctly identifies the metastatic potential of every cell line from the panel (Fig. 1g), corresponding to a ROC curve with 100% AUC (Fig. 1h). Taken together, these results indicate that the MAqCI assay accurately predicts the metastatic potential of breast epithelial and breast cancer cell lines.

To determine the ability of MAqCI to identify migratory cells within heterogeneous populations, we mixed at various ratios (1:1, 1:4 and 1:9) aggressive MDA-MB-231 breast cancer cells with non-aggressive MCF7 breast cancer cells pre-labelled with two spectrally distinct fluorophores and seeded a total of 50,000 cells in MAqCI. The percentage of migratory cells as a function of time was quantified for the mixed population, as well as for the two cell lines individually within each device (Supplementary Fig. 2a–c). Even when MDA-MB-231 cells are diluted tenfold, about 20% of these MDA-MB-231 cells are still migratory. Along these lines, the percentage of MCF7 migratory cells is below the threshold for all of the different ratios examined in this study.

Metastatic potential of migratory versus unsorted cells. As the relative abundance of migratory cells is integral to the prediction of metastatic potential, we hypothesized that these cells have an elevated capacity to form metastases compared with the unsorted cell population. It is noteworthy that isolated migratory MDA-MB-231 cells lose their migration advantage over the heterogeneous unsorted population after cell culture in vitro for 14 days (Supplementary Fig. 3a). To avoid genetic and phenotypic changes caused by in vitro culture on stiff, two-dimensional vessels[18], we performed in vivo assessment of tumour formation and spontaneous metastasis of migratory cells versus unsorted MDA-MB-231 cells directly after isolation using MAqCI. Isolated migratory or unsorted cells were subcutaneously injected into the fourth mammary fat pad of non-obese diabetic/severe combined immunodeficient (NOD/SCID) mice. Bioluminescent imaging revealed that both cell populations formed tumours that grew at similar rates (Fig. 2a and Supplementary Fig. 3b).

Bioluminescent imaging analysis revealed that four out of eight mice injected with migratory cells developed metastases in the bone after eight weeks, whereas no mice injected with cells from the unsorted population developed metastasis in this tissue (Fig. 2b,c and Supplementary Fig. 3c). Moreover, seven out of eight mice injected with migratory cells formed metastases in the lung and liver, as opposed to six out of eight or five out of eight mice injected with the unsorted cell population, respectively (Fig. 2b,c and Supplementary Fig. 3c). Intriguingly, quantitative image analysis revealed an eightfold increase in the metastatic burden of the lung and liver of mice injected with migratory cells relative to those injected with the unsorted cell population. To validate the differences in metastatic burden observed in the lung and liver, DNA was isolated from samples of these tissues and the amount of human DNA was quantified via quantitative PCR (qPCR) using primers specific for human long interspersed nuclear elements[19] (Supplementary Fig. 3d). Mice injected with migratory as opposed to unsorted cells had a 2.5- and 4.5-fold increase in the amount of human DNA in their lung and liver, respectively (Fig. 2d,e). Metastasis was also detected in the axillary lymph nodes of three out of eight animals injected with migratory cells and only two out of eight animals injected with unsorted cells. Tissue samples from representative specimens were also processed for immunohistochemistry against human mitochondria, as well as hematoxylin and eosin staining (Fig. 2f). Representative images confirm the presence of metastatic human cells in the lung, liver and lymph node of mice. Bone marrow was not intact after thin sectioning, and thus images are not shown. Taken together, these data indicate that while migratory and unsorted cell populations both form tumours that grow at similar rates, migratory cells have a markedly enhanced ability to form spontaneous metastases.

Characterization of migratory-cell phenotype and genotype. To understand which factors contribute to the increased metastatic potential of migratory cells, we first compared the phenotype of migratory and non-migratory MDA-MB-231 cells in MAqCI. Migratory cells move with higher velocity (net displacement over time), speed (average displacement over each 20 min time interval) and persistence (net displacement over total distance travelled) (Fig. 3a–c) than non-migratory cells. Migratory cells also spread over a larger area and have a higher aspect ratio, indicating their intrinsic ability to elongate in the direction in which they move (Fig. 3d–e). They also form more protrusions, as evidenced by their lower solidity (cell area/convex area) relative to non-migratory cells (Fig. 3f). Typically, these protrusions occur at the leading edge of the migratory cells, which may serve to probe the local microenvironment. Taken together, migratory cells are more elongated and more protrusive, and move faster and more persistently than non-migratory cells in vitro. These phenotypic differences may contribute to their enhanced metastatic potential in vivo.

To elucidate potential gene expression differences contributing to the migratory cell's increased motility in vitro and metastasis in vivo, we isolated RNA from equal numbers of migratory and unsorted MDA-MB-231 cells and performed genome-wide transcription analysis using RNA sequencing (RNA-seq). RNA-seq analysis identified 1,433 differentially expressed genes (DEGs, adjusted P value < 0.1) between the migratory and unsorted cells, consisting of 582 upregulated and 851 downregulated genes (Supplementary Datasets 1 and 2). Of note, these gene expression changes are not due to DNA damage induced by cell entry and migration inside the narrower branch channels, as evidenced by immunostaining showing similar levels of DNA damage marker phospho-H2A.X between migratory and non-migratory cells (Supplementary Fig. 4). To understand the function of the DEGs, we performed pathway analysis and gene ontology clustering. We found that migratory cells had gene expression changes in multiple signalling pathways (Fig. 3g), including RAS/mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3K)-AKT, tumour necrosis factor (TNF), FOXO and several pathways related to metabolism. Gene ontology clustering (Fig. 3h) revealed that migratory cells have differentially expressed genes related to cell migration, regulation of apoptosis, metabolism, angiogenesis and nitric oxide biosynthesis. Select genes from relevant pathways were verified for fold expression changes using qPCR (Fig. 3i). The PI3K pathway is activated in more than 70% of patients with invasive breast cancer[20]. While activating mutations in the canonical RAS/MAPK pathway occur at a lower rate (2–10%) in breast cancer[21], this pathway's activity is linked to breast cancer metastasis, and its aberrant activity may be induced by overexpression of upstream receptor tyrosine kinases such as human epidermal growth factor receptors 1 and 2 (also known as HER1 and HER2), which is common in breast cancer[22]. As the PI3K and RAS/MAPK pathways are implicated by the RNA-seq data and play a role in breast cancer progression, we sought to determine whether activation of these two pathways was sufficient to confer metastatic ability in non-tumourigenic and non-metastatic cells in vivo.

Prediction of the metastatic potential of mutated epithelial cells. The mammary epithelial cell line, MCF-10A, is non-tumourigenic and non-metastatic. Within this genetic background, we activated the PI3K pathway by knockout of the tumour suppressor, phosphatase and tensin homologue (PTEN), activated the RAS/MAPK pathway through overexpression of activated KRAS, KRAS(G12V), and created a double mutant cell line with both interventions ($PTEN^{-/-}$KRAS(G12V))[23]. The metastatic potential of all four cell

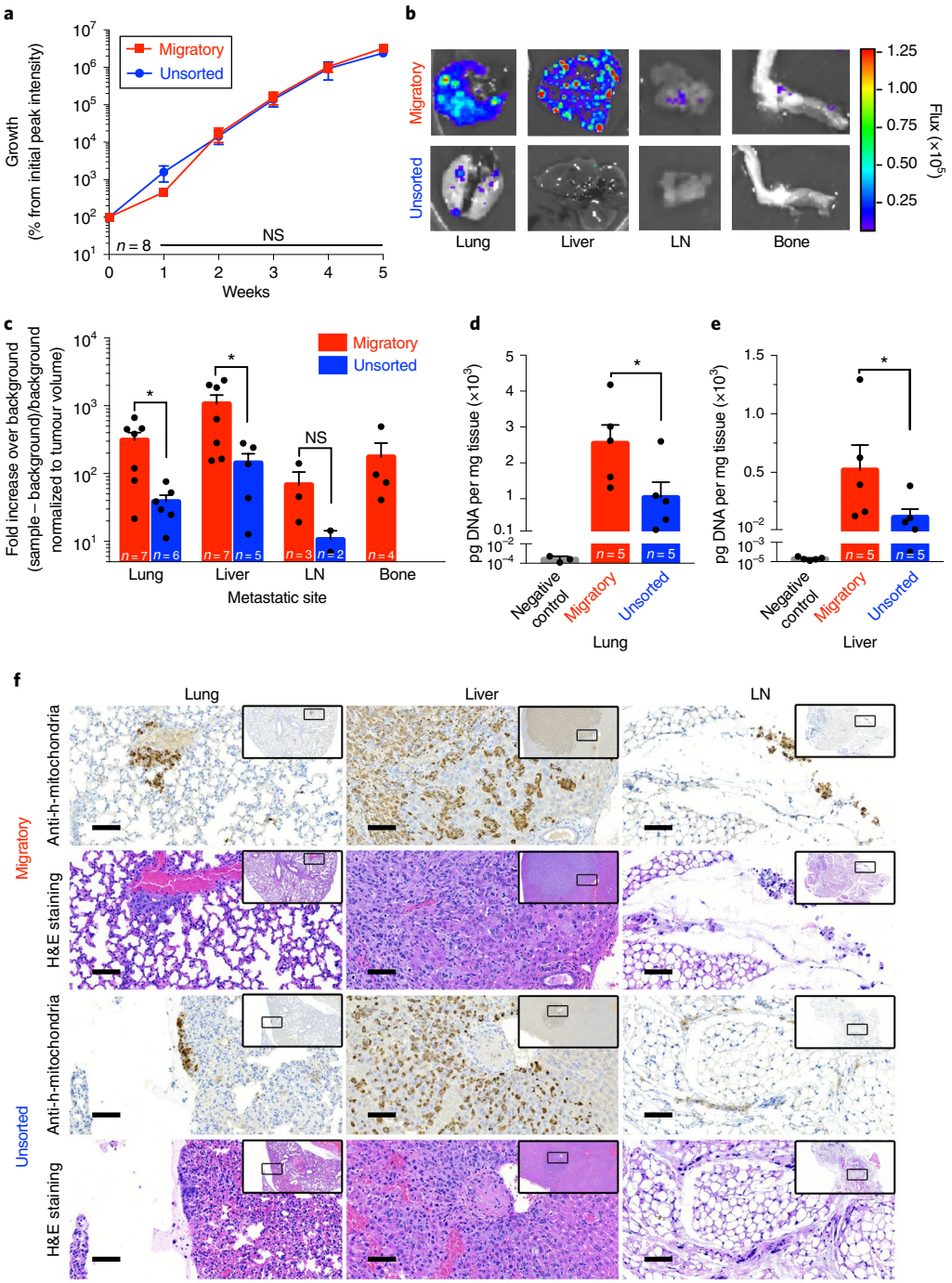

Fig. 2 | Migratory cells have similar tumourigenic but markedly enhanced metastatic potential in vivo than unsorted breast cancer cells. a, Bioluminescent signal from migratory and unsorted MDA-MB-231 cell xenografts. Percent growth was determined by subtracting background from the peak signal on the appropriate day and normalizing to the initial background subtracted reading for the same mouse. Data represent mean ± s.e.m. from $n = 8$ mice per group. NS, $P \geq 0.33$ as assessed by a two-tailed Mann–Whitney test without adjustment for multiple comparisons. **b**, Representative bioluminescent images of the lung, liver, axillary lymph node (LN) and bone of mice injected with migratory or unsorted cells (eight images acquired per condition, see also Supplementary Fig. 3c). **c**, Quantification of bioluminescent signal from lung (*$P = 0.022$), liver (*$P = 0.030$), LN (NS, $P = 0.20$) and bone. Statistical comparison made using a two-tailed Mann–Whitney test. To prevent false-positive readings, each site was considered positive if its peak bioluminescent signal minus background was ≥50× background. This fold increase over background was normalized to tumour volume to control for variation between specimens. Each data point represents the signal from one mouse. Column and error bars represent mean ± s.e.m. from mice with detectable metastases (n mice annotated in each bar). **d**,**e**, Amount of human DNA detected by qPCR following isolation from lungs (**d**) or liver (**e**) of mice injected with migratory or unsorted cells. Each data point represents the reading from one mouse. Data represent mean ± s.e.m. from $n = 5$ mice per group. Lungs/liver from mice not injected with tumour cells served as negative controls. **f**, Representative 20× images of immunohistochemistry for human mitochondria and hematoxylin and eosin (H&E) staining of adjacent sections from the lung, liver and LN of mice injected with migratory or unsorted cells (for each organ and group, the animal with median bioluminescent signal was chosen for histology). Scale bar, 200 μm. Insets show the surrounding area at 4×, and indicate the area displayed at 20×.

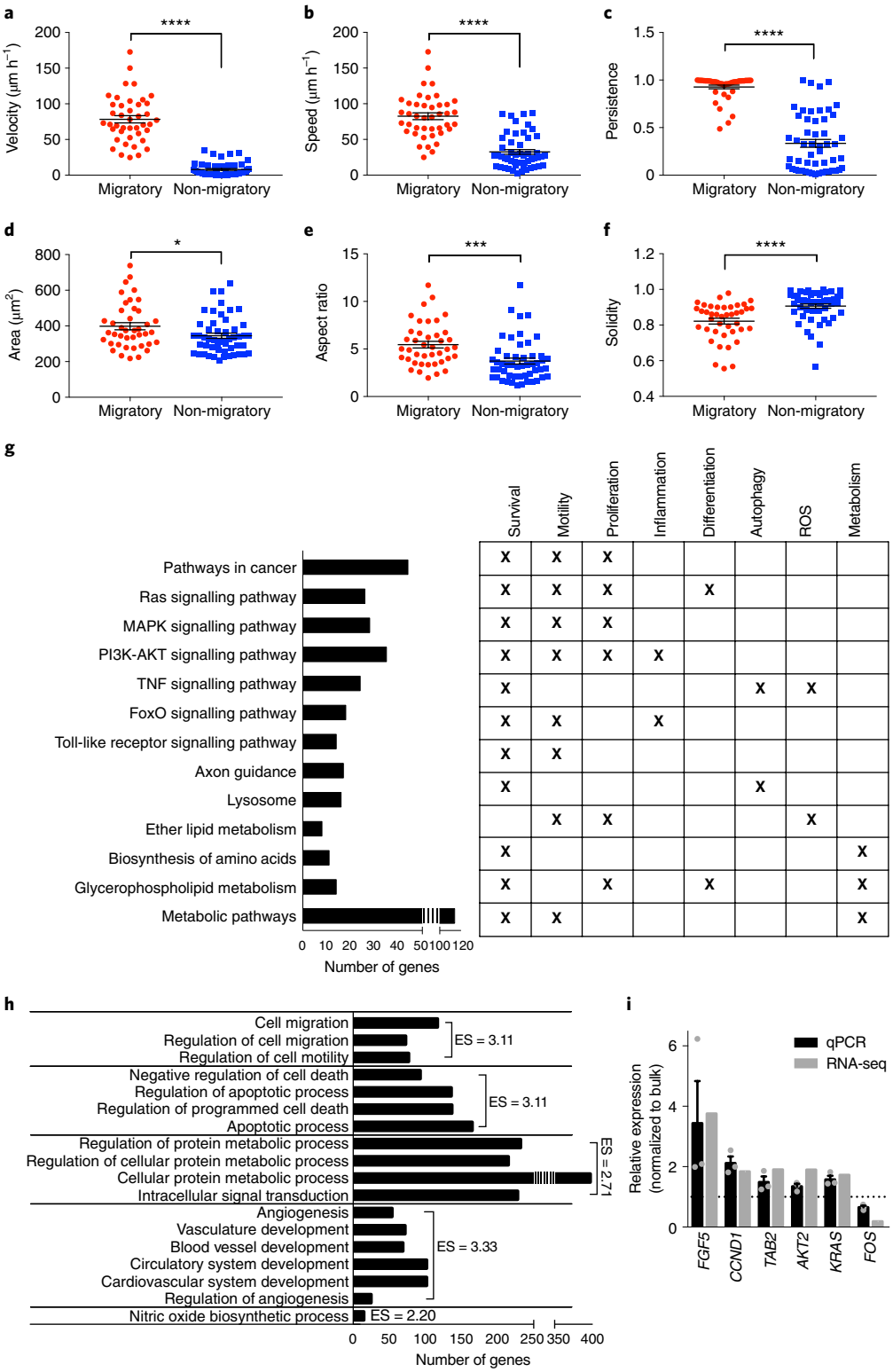

Fig. 3 | Characterization of phenotype and genotype of migratory cells. a–**f**, Velocity (**a**), speed (**b**), persistence (**c**), spread area (**d**), aspect ratio (**e**) and solidity (**f**) of migratory and non-migratory MDA-MB-231 cells while in the feeder channel. Each data point represents the metric value for one cell averaged over its time in the channel for ≥50 cells from $n = 3$ independent experiments. Line and error bars represent mean ± s.e.m. **g**,**h**, Pathway analysis (**g**) and gene ontology annotation (**h**) for DEGs in migratory cells. ROS, reactive oxygen species; ES, enrichment score. **i**, Validation of select DEGs identified by RNA-seq using qPCR. Each data point represents the relative expression from one experiment. Column and error bars represent mean ± s.e.m. from $n = 3$ independent experiments. *$P = 0.030$, ***$P < 0.001$, ****$P < 0.0001$ by two-tailed Student's t-test.

lines was first evaluated in MAqCI. The PTEN$^{-/-}$KRAS(G12V) cell line displayed high percentages of migratory (24% at 13 h) and Ki-67-positive cells (86%) (Fig. 4a,b). In comparison, parental MCF-10A cells and PTEN$^{-/-}$ cells exhibit low proportions of migratory (0% at 13 h) and proliferating cells (33–40%). Importantly, this phenotypic behaviour of PTEN$^{-/-}$ cells is not due to clonal

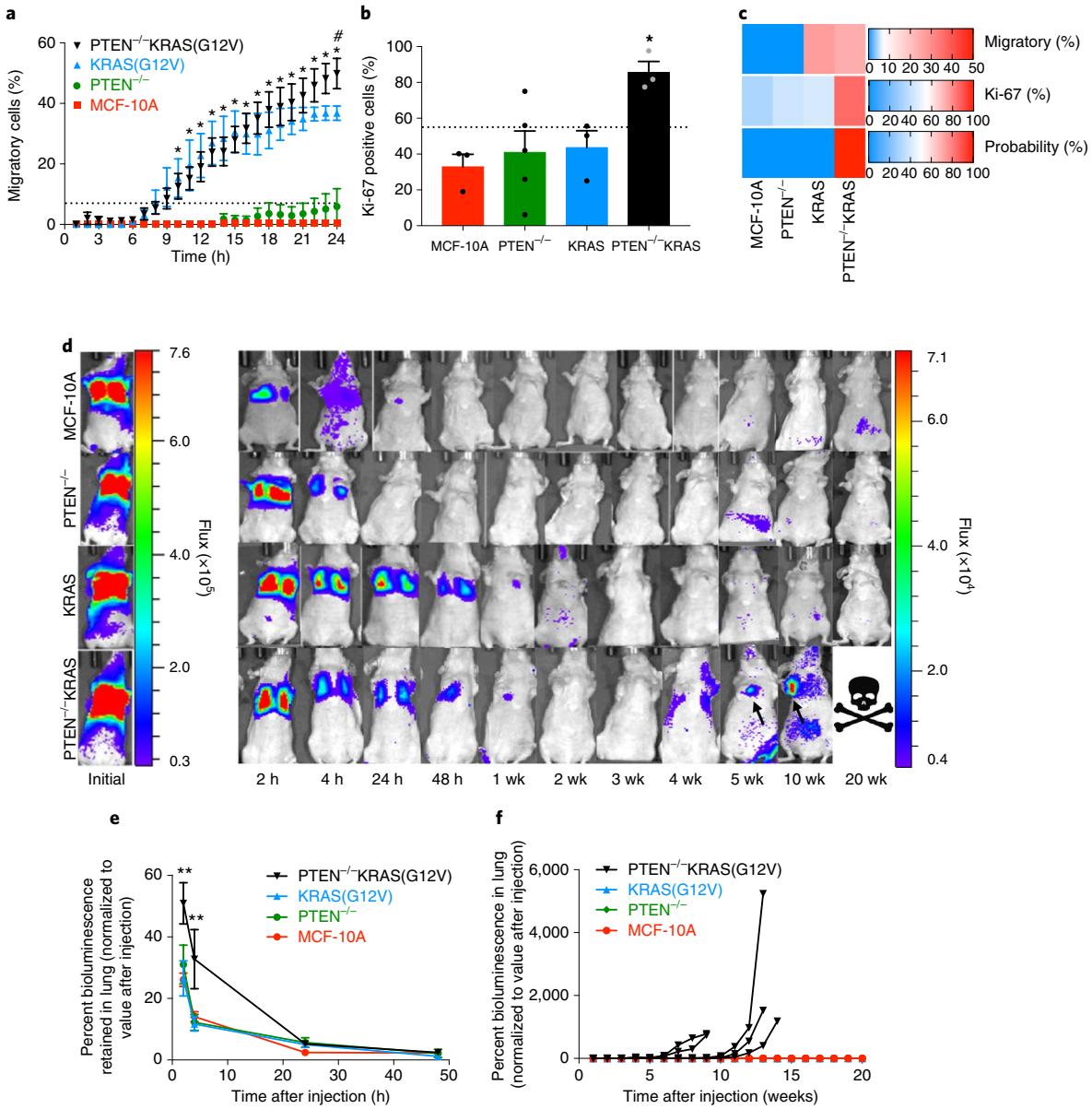

Fig. 4 | MAqCI predicts metastatic potential conferred by activation of PI3K and RAS/MAPK pathways in breast epithelial cells. a, Percentage of migratory cells in PTEN$^{-/-}$KRAS(G12V), KRAS(G12V), PTEN$^{-/-}$ and parental MCF-10A cells. Data represent mean ± s.e.m. from $n \geq 3$ independent experiments. $^{*}P < 0.05$ for PTEN$^{-/-}$KRAS(G12V) and KRAS(G12V) compared with MCF-10A and PTEN$^{-/-}$; $^{\#}P < 0.05$ for PTEN$^{-/-}$KRAS(G12V) compared with KRAS(G12V) at 24 h. *P* values calculated by two-way ANOVA followed by Tukey's multiple comparison test. Exact *P* values are listed in Supplementary Dataset 3. Dotted line designates 7% of migratory cells. **b**, Percentage of Ki-67-positive cells in MAqCI. Each data point represents the percentage from one experiment. Column and error bars represent mean ± s.e.m. $n \geq 3$ independent experiments. $^{*}P = 0.025$ for PTEN$^{-/-}$KRAS compared with MCF-10A and PTEN$^{-/-}$, $P = 0.026$ compared with KRAS, as calculated by one-way ANOVA followed by Tukey's multiple comparisons test. **c**, Probability (%) of each cell line possessing high metastatic potential calculated using logistic regression and the percentages of migratory and Ki-67-positive cells. Baseline value (shown in white) for each heat map is set to the threshold value for each predictor (7% migratory cells and 55% Ki-67-positive cells). **d**, Representative bioluminescent images of mice following tail-vein injection with 10^6 cells (5 animals imaged per group, see also Supplementary Fig. 5b). Scale is shown to the right of each set of images. PTEN$^{-/-}$KRAS(G12V) cells form tumours (arrows) causing ethical endpoint before 20 weeks; the example shown for this condition corresponds to the third curve from the left in **f**. **e**, Percentage of bioluminescent signal retained in the lung of each animal for 48 h post-injection. For each time point, the background was subtracted from the peak signal and the difference was normalized to the initial value for that sample ($n = 5$ per group). Data represent mean ± s.e.m. $^{**}P < 0.01$ for PTEN$^{-/-}$KRAS(G12V) compared with MCF-10A, PTEN$^{-/-}$ and KRAS(G12V) at 2 and 4 h. *P* values calculated by two-way ANOVA followed by Tukey's multiple comparison test. Exact *P* values are listed in Supplementary Dataset 3. **f**, Percentage bioluminescent signal in the lung for 20 weeks post-injection. Each curve represents one mouse ($n = 5$ per group).

selection, as three different clones display minimal migratory potential (Supplementary Fig. 5a). Interestingly, KRAS(G12V) cells exhibit similar motility (27% at 13 h) but markedly lower proliferation (44%) than PTEN$^{-/-}$KRAS(G12V) cells (Fig. 4a,b). The logistic regression formula established from our panel of breast cancer and breast epithelial cell lines (Supplementary Table 2) predicts that only the combined activation of the PI3K and RAS/MAPK pathways will confer high metastatic potential on MCF-10A cells (Fig. 4c).

To test this hypothesis, equal numbers (10^6) of parental, PTEN$^{-/-}$, KRAS(G12V) and PTEN$^{-/-}$KRAS(G12V) cells were injected into the tail vein of mice and monitored with bioluminescent imaging (Fig. 4d). The tail-vein injection model was selected because three of these cell lines do not form or rarely form subcutaneously implanted tumours[23], thus making comparison of their spontaneous metastasis from a primary tumour unrealistic. Following tail vein injection, the initial ability of cells to survive in the bloodstream and reattach in the lung capillaries was determined by imaging from 2 to 48 h. PTEN$^{-/-}$KRAS(G12V) cells exhibit a higher bioluminescent signal in the lungs for 2–4 h after injection, suggesting that they were able to firmly adhere within the lung capillaries (Fig. 4d,e). After 24–48 h, the bioluminescent signal in the lung decreases to a similar level in each cell line, indicating that the elevated signal initially observed in the PTEN$^{-/-}$KRAS(G12V) cells was dissipated by a combination of cell death, detachment and re-circulation within the bloodstream or migration from the blood vessel into surrounding tissue. Monitoring the cells for longer periods enabled us to assess their ability to extravasate, invade the surrounding tissue, survive and proliferate. Each mouse injected with PTEN$^{-/-}$KRAS(G12V) cells developed a tumour in the lung leading to an experimental endpoint between 9 and 14 weeks (Fig. 4d,f). No evidence of tumour formation was observed in mice injected with parental MCF-10A, PTEN$^{-/-}$ nor KRAS(G12V) cells. Taken together, these results demonstrate that PTEN$^{-/-}$KRAS(G12V) cells have high metastatic potential while MCF-10A, PTEN$^{-/-}$ and KRAS(G12V) cells failed to form metastases in vivo. Furthermore, the MAqCI assay accurately predicted the metastatic potential of these cell lines.

MAqCI evaluation of cells from patient-derived xenografts. Patient-derived xenografts (PDXs) are used as a model for breast cancer tumour growth and metastasis because of their ability to recapitulate the hallmarks of the patient's disease[24,25]. We selected two well-characterized tumour specimens from patients with metastatic triple-negative breast cancer (HCI-001 and HCI-002)[24] and expanded them as xenografts in mice. As these specimens developed metastasis in the original patients and during their initial characterization as xenografts in mice, we expected that after quantifying the percentages of migratory and Ki-67-positive cells dissociated from the xenograft tumours using MAqCI, our logistic regression formula would calculate a high probability of metastasis. Indeed, dissociated PDX cells were able to migrate through the feeder channel and enter the branch channels (Fig. 5a and Supplementary Video 3). After each experiment, cells were fixed in the microfluidic device and immunostained for human mitochondria to ensure that the migratory cells strictly originated from the cancer patient and were not mouse stromal cells (Fig. 5c). Both HCI-001 and HCI-002 samples exceeded the threshold of migratory cells for each of the four optimal combinations of time and percentage of migratory cells (Fig. 5b,d). These samples also had high levels of Ki-67 cells (78% and 67%, respectively—Fig. 5e). Use of the migratory and proliferation indices in the logistic regression formula developed from the breast cancer and breast epithelial cell lines identifies both HCI-001 and HCI-002 as metastatic (probability = 1). These experiments establish proof-of-principle that MAqCI can be used to predict the metastatic potential of clinically relevant specimens.

MAqCI testing of therapeutics from active clinical trials. Accumulating evidence suggests that patients with breast cancer with the same molecular subtype may have differing responses to treatment[25–27]. To date, no reliable indicator of response to a specific therapeutic regimen exists. As MAqCI is capable of identifying highly migratory and proliferative cells that have enhanced metastatic potential, we hypothesized that it would be well suited for screening the effectiveness of potential therapeutics to inhibit the

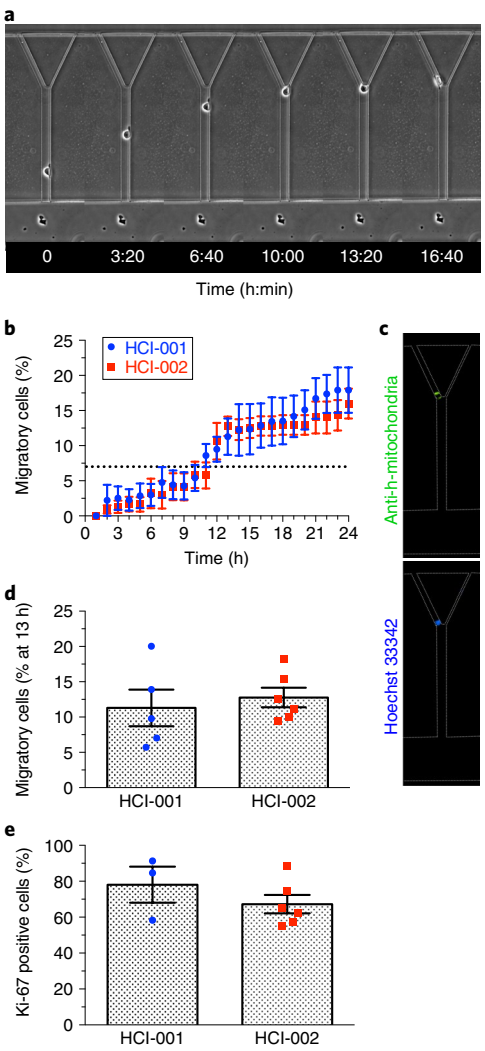

Fig. 5 | MAqCI accurately predicts the metastatic potential of cells obtained from PDXs. a, Migratory HCI-002 cell entering left branch channel. **b**, Percentage of migratory cells for PDX cell lines. Data represent mean ± s.e.m. from $n \geq 5$ independent experiments. Dotted line designates 7% migratory cells. **c**, Immunostaining for human mitochondria and nucleus performed after migration experiment. **d**, Percentage of migratory cells in HCI-001 and HCI-002 after 13 h. Each data point represents the percentage from one experiment for $n \geq 5$ independent experiments. Column and error bars represent mean ± s.e.m. **e**, Percentage of Ki-67-positive cells in MAqCI. Each data point represents the percentage from one experiment for $n \geq 3$ independent experiments. Column and error bars represent mean ± s.e.m.

motility of these cells. As migratory cells have altered expression of numerous genes from the RAS/MAPK and PI3K pathways, we chose to evaluate inhibitors targeting these pathways that have potential for use in the clinic. The MEK1/2 inhibitor, trametinib, is Food and Drug Administration approved for use in melanomas harbouring *BRAF* V600E or V600K gene mutations and is under evaluation for effectiveness in breast cancer in several active clinical trials. The PI3K inhibitor, BKM120, is also under evaluation in several clinical trials for use in breast cancer. A clinically relevant concentration of each drug was selected based on published pharmacokinetic data of each drug's volume of distribution and patient dosage/target concentration[28,29]. For trametinib, 70 nM was selected, and for BKM120, 1 μM. Before testing in MAqCI, we confirmed that 24 h treatment

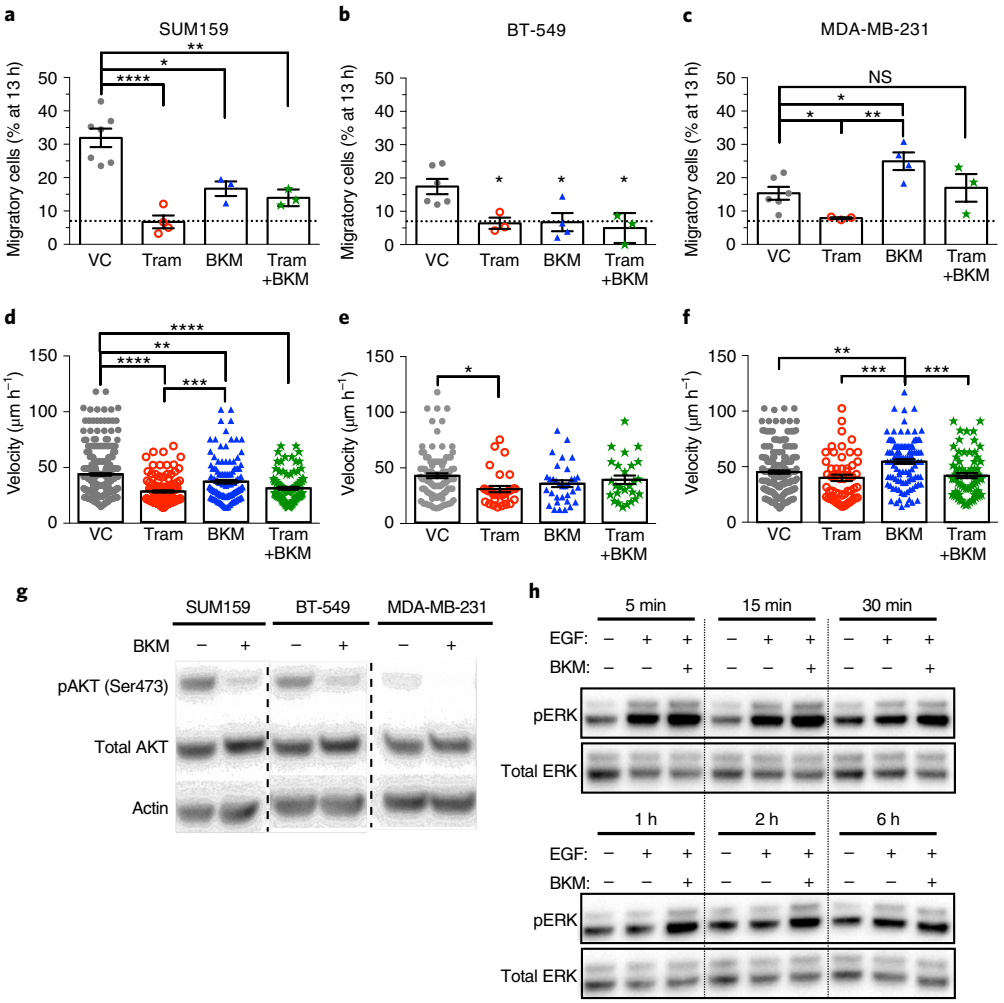

Fig. 6 | MAqCI testing of therapeutic agents from ongoing clinical trials. a–c, Percentage of migratory SUM159 (**a**), BT-549 (**b**) and MDA-MB-231 (**c**) cells after 13 h during treatment with vehicle control (VC), BKM120 (BKM; 1 μM), trametinib (TRAM; 70 nM) or BKM120 and trametinib. Dotted line designates 7% of migratory cells. Each data point represents the percentage from one experiment. Column and error bars represent mean ± s.e.m. of $n \geq 3$ independent experiments. **d–f**, Velocity of SUM159 (**d**), BT-549 (**e**) and MDA-MB-231 (**f**) cells migrating in the feeder channel. *P* values were calculated using one-way ANOVA followed by Tukey's multiple comparison test. NS, $P \geq 0.05$. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. Exact *P* values are listed in Supplementary Dataset 3. **g**, Representative western blot for pAKT in SUM159, BT-549 and MDA-MB-231 cells treated with vehicle control or BKM120 (1 μM) for 24 h. **h**, Representative western blot for pERK in MDA-MB-231 cells serum-starved for 24 h, treated with vehicle control or BKM120 (1 μM) for 30 min, then in select samples stimulated with EGF (100 nM) for the indicated time. Western blot experiments were repeated $n = 3$ times. Full scans of the cropped blots appear in Supplementary Fig. 7.

of either or both therapeutics did not significantly affect the percentage of viable cells compared to treatment with a vehicle control (Supplementary Fig. 6a). Thus, any effect of the inhibitors on the percentage of migratory cells cannot be attributed to cell toxicity.

Trametinib treatment of three triple-negative breast cancer cell lines with high metastatic potential, SUM159, BT-549 and MDA-MB-231, was effective in reducing the percentage of migratory cells down to the threshold levels exhibited by cells with low metastatic potential (Fig. 6a–c). This pharmacological intervention also decreases the migration velocity of all three cell lines (Fig. 6d–f). The marked reduction in the percentage of migratory cells in response to trametinib treatment suggests its potential to reduce the risk of metastasis in patients bearing tumours with the characteristics of these breast cancer cell lines. Interestingly, pharmacological inhibition of PI3K yields different outcomes in these triple-negative breast cancer cell lines. Although BKM120 treatment reduces the percentage of migratory SUM159 and BT-549 cells compared with vehicle control (Fig. 6a,b), it significantly augments the migratory potential of MDA-MB-231 cells (Fig. 6c). In a clinical setting, an increase in the percentage of migratory cells is predicted to exacerbate the risk of metastasis. Similarly, BKM120 exerts divergent effects on the migration velocity of the aforementioned breast cancer cell lines (Fig. 6d–f). The combination of trametinib and BKM120 treatment failed to enhance the inhibitory effects of trametinib alone in all cell lines tested (Fig. 6a–f).

Further examination of the genotype of these three breast cancer cell lines[30] provided insights into their divergent responses to PI3K inhibition (Supplementary Fig. 6b). SUM159 and BT-549 both harbour activating mutations in the PI3K pathway (H1047L mutation in the PI3K catalytic subunit or loss of PTEN, respectively). Inhibition of the overactive PI3K pathway in both of these cell lines decreased their percentage of migratory cells. In contrast, MDA-MB-231 cells have no activating mutations in this pathway. PI3K activity leads to the phosphorylation of the effector protein, AKT on serine 473 (pAKT). Western blotting indicates that while these three cell lines express similar levels of AKT, SUM159 and

BT-549 both have significantly higher pAKT than MDA-MB-231 (Fig. 6g and Supplementary Fig. 7a–c), in line with their activating mutations in this pathway. Addition of BKM120 decreases the level of pAKT in all three cell lines. It is established that AKT can inhibit RAF, reducing downstream signalling to MEK and ERKs[31]. We thus hypothesized that inhibition of PI3K in MDA-MB-231 alleviates AKT inhibition of RAF, increasing downstream signalling in the RAS/MAPK pathway (Supplementary Fig. 6c).

To test this hypothesis, we used western blotting to assess the phosphorylation levels of the RAS/MAPK pathway's downstream effectors, ERK1/2 (pERK) in response to epidermal growth factor (EGF) stimulation in the presence or absence of BKM120. EGF increases the levels of pERK in vehicle control-treated cells from 5–30 min, while at longer times (≥1 h), the pERK levels are similar to that of unstimulated cells (Fig. 6h and Supplementary Fig. 7d–f). In line with our hypothesis, cells pre-treated with BKM120 and then stimulated with EGF displayed higher pERK levels than those of EGF-stimulated vehicle control cells from 5 min to 1 h (Fig. 6h and Supplementary Fig. 7d–f); thus, BKM120 treatment not only increased the levels of pERK at early time points but also prolonged the duration of this signal beyond the time frame that the vehicle control-treated cells showed activation. Taken together, these results demonstrate that inhibition of PI3K can increase RAS-mediated cell motility in MDA-MB-231, but not in SUM159 or BT-549 cells. Using the MAqCI assay, we herein demonstrate the potential to use a phenotypic test to rapidly screen the efficacy of therapeutics for reducing metastatic potential without the need for genetic testing and analysis to attempt to predict the response.

Discussion

To date, the diagnosis of metastatic potential in breast cancer involves the consideration of multiple independent parameters such as the grade, stage and molecular subtype, which do not provide complete information on the expected outcome of a patient's cancer. A second limitation to breast cancer treatment is the lack of a functional assay to evaluate the efficacy of novel therapeutic drugs on a patient-specific basis for personalized medicine. We herein developed a companion assay that provides information on the patient's risk of metastasis by measuring both the motility and proliferative potentials of their cancerous cells, which are key properties for establishing metastatic colonies. The advantages of MAqCI include the physical isolation of highly migratory/highly proliferative cells for detailed molecular and genetic characterization and the potential for high-throughput evaluation of potential antimetastatic therapeutics for precision medicine. MAqCI has been standardized by examining a large panel of normal-like breast epithelial and cancer cell lines and validated using engineered cell lines with demonstrated metastatic potential and PDX specimens. This assay can yield results in hours, compared with the current gold standard of 4–12 months required for PDXs to grow in mice. Moreover, PDXs are only successful in 28–37% of cases, leading to a significant loss of patient representation[24,25]. The ability to complete the assay within 24 h enables prognostic clinical applications that would not be possible for a system that models the slow pace and complexity of in vivo metastasis. Furthermore, the assay evaluates the ability of potential therapeutics to inhibit the migration and/or proliferation of highly motile metastasis-initiating cells, rather than the growth of the unsorted tumour population.

We and others have demonstrated that metastatic cells are more migratory than non-metastatic ones[32–34]. Our results indicate that the percentage of migratory cells correlates with the metastatic potential of breast cancer cell lines. Interestingly, we found that the predictive power of the MAqCI motility assay was greater than that of the widely used transwell-migration assay. In view of these findings, we conclude that entry into a confining pore itself is not sufficient for the prediction of metastatic potential but must be combined with a measure of directional persistence, and potentially with higher geometric complexity (decision-making). Other conventional migration assays, such as wound healing, were not compared with MAqCI due to their inability to isolate migratory cells for further study. We also extended our studies to clinically relevant PDX specimens and found that our logistic regression formula accurately predicts their metastatic propensity. Assessing two metastatic PDX specimens provides a proof-of-principle for the use of MAqCI with clinically relevant cell preparations but lacks comparison with non-metastatic specimens. A limitation of the PDX model is that the tumour specimens that do grow tend to be the most aggressive[24], thereby limiting the availability of non-aggressive samples for comparison. MAqCI avoids this pitfall—as well as the selective pressures introduced by long-term growth—by directly assessing tumour cells isolated from specimens without a requirement for growth either in vitro or as mouse xenografts. More comprehensive comparison of PDX specimens with low versus high metastatic potential will require the invention of new technologies that permit the reliable growth of a broader spectrum of patient tumours as PDXs. Although some established prognostic markers are limited to specific breast cancer subtypes[4], MAqCI analyses cell motility and proliferation, which are universal phenotypic features required by all metastatic cells independent of subtype and perhaps of tissue.

Combining our motility index with a clinically established proliferation index improves the predictive power of MAqCI by eliminating any false-negative readings assessed by motility alone. Using normal breast epithelial and breast cancer cell lines, our accuracy is 100%. Ki-67 levels can be reproducibly assessed clinically by immunohistochemistry, which has already been shown to have prognostic value in breast cancer[16], and is recommended by the St. Gallen guidelines to help determine the molecular subtype. Furthermore, RNA-seq reveals that Ki-67 (*MKI67*) is upregulated by migratory cells, affirming its role as a marker for poor prognosis. The percentage of Ki-67-positive cells was determined for tumour xenografts of MDA-MB-231, BT-549, HCI-001[35] and HCI-002[36] using immunohistochemistry (50%, 50%, 55% and 31%, respectively). These reported values are lower than our measurements in MAqCI. Percentages of Ki-67-positive cells are likely reduced in tumour samples due to the increased competition for limited nutrients and oxygen within the tumour microenvironment, as compared with the controlled media, pH and oxygen available to isolated cancer cells in MAqCI. By using well-established breast cancer cell lines, we did not detect any differences in the proliferative index of migratory versus non-migratory cells. However, in patient's samples, there may be a difference caused by competition in the tumour microenvironment. Thus, it will be important to specifically measure the proliferative index of migratory cells, which can only be done in situ using MAqCI. Future studies with patient specimens will use the methods presented here to further validate the logistic regression coefficients developed for MAqCI.

MAqCI correctly predicts that the simultaneous activation of the PI3K and RAS/MAPK pathways enables breast epithelial cells to form metastases in vivo. The tail-vein injection model captures the latter stages of the metastatic cascade, including survival in the bloodstream, reattachment in lung capillaries, migration out of the blood vessel, migration/invasion in the lung tissue, persistent survival (dormancy) and recurrent growth. The MAqCI assay models the migration steps of this cascade, and assesses proliferation using Ki-67, which should predict recurrent growth. PTEN$^{-/-}$KRAS(G12V) cells have high motility and proliferation indices, and only this cell line formed lung tumours following tail-vein injection in mice, in line with results from a murine model of prostate cancer[37]. PTEN$^{-/-}$KRAS(G12V) cells were retained in the lung more efficiently at early time points, which was followed by the formation of large tumours by 9–14 weeks. Both PTEN$^{-/-}$ and parental MCF-10A cells, which have low indices for migration and proliferation, fail to generate any tumours. Interestingly,

KRAS(G12V) cells, although they have a high motility index, could not generate metastatic colonies, presumably due to their low proliferative index; this highlights the importance of evaluating the combined indices. Mice injected with parental, PTEN$^{-/-}$ or KRAS(G12V) cells were monitored for one year without signs of tumour formation. The observations we have made in the tail-vein model are in line with results from the subcutaneous injection of these cell lines[23], where PTEN$^{-/-}$KRAS(G12V) cells form rapidly growing tumours. In that study, PTEN$^{-/-}$KRAS(G12V) cells grown as xenografts in mice had increased Ki-67 levels compared with PTEN$^{-/-}$ or KRAS(G12V) cells, which is in line with our results in MAqCI. Collectively, these data demonstrate the ability of the assessment of migration and proliferative potentials in MAqCI to predict the formation of metastatic colonies.

RNA-seq revealed that key pathways that are deregulated in breast cancer, including RAS/MAPK and PI3K, contribute to enhanced migratory and metastatic potentials. Signalling pathways and biological processes related to cell migration, proliferation, survival and metabolism influence the phenotype of migratory cells. Since the PI3K and RAS/MAPK pathways play a role in breast cancer progression and metastasis, they have both been targeted by numerous therapeutics; however, cross-talk between these pathways has limited the success of targeting either pathway individually, leading to the theory that simultaneous inhibition of these pathways is required for maximal inhibition of tumour growth and metastatic activity[38]. This approach has had clinical success for antitumour activity in *RAS* or *BRAF* mutated non-small cell lung, ovarian and pancreatic cancer[39], and is also under examination in advanced solid tumours (NCT01449058). It follows logically that for maximal metastatic potential, the activity of both pathways is required, which is confirmed by our functional and sequencing results.

Drugs are typically tested in patients based on the molecular subtype of breast cancer. However, the presence of different mutations among patients leads to varied responses to the same treatment. By using MAqCI to screen the efficacy of BKM120 at reducing metastatic potential, we discovered divergent responses in three triple-negative cell lines, highlighting the importance of our phenotypic assay to measure the potential efficacy of therapeutic drugs in a rapid, reliable and reproducible manner. We also deciphered the genetic factors responsible for the distinct responses of these breast cancer cell lines to the same treatment. PI3K, the target of BKM120, can regulate cell migration by altering the localization and activity of Rho GTPases[40]. Cell lines with PI3K pathway activation (through *PIK3CA* mutation or PTEN loss) exhibited reduced motility after treatment with BKM120. Interestingly, in the cell line without PI3K pathway activation, MDA-MB-231, BKM120 enhanced cell motility. We revealed that this occurs due to a cross-talk mechanism between AKT and RAF, which resulted in increased RAS/MAPK signalling when BKM120 treatment reduced AKT inhibition of RAF. This result is in line with the fact that knockdown of AKT1, a downstream effector of PI3K, potentiates invasion in breast epithelial and breast cancer cells[41]. We also examined the effect of trametinib, a MEK inhibitor, in these three triple-negative cell lines. This intervention consistently reduced the migratory potential of all three cell lines, in accordance with the well-studied roles of the RAS/MAPK pathway in supporting cell migration[42]. Trametinib is currently being evaluated in clinical trials for breast cancer and advanced solid tumours, either alone or in combination with other inhibitors, including ones targeting the PI3K pathway. Previous work has suggested that cell lines can develop dominant signalling pathways that inhibit the activity of other pathways through negative feedback loop mechanisms[43]. Under this regime, MDA-MB-231 cells, which do not bear constitutive PI3K activation, would be resistant to PI3K inhibition, while sensitive to inhibition of the RAS/MAPK pathway. Indeed, it has been demonstrated that MDA-MB-231 and other basal-like breast cancer cell lines that depend on RAS/MAPK signalling (similar to that of RAS-transformed cell lines), are insensitive to PI3K inhibition while remaining sensitive to MEK inhibition[44]; this study also found that PTEN loss reduced sensitivity to MEK inhibition. Further, in prostate cancer cell lines that retain PTEN, it has been shown that AKT inhibition enhances RAS/MAPK signalling[45]. Interestingly, although SUM159 and BT-549 cells have constitutive PI3K activity, they were still sensitive to MEK inhibition. This is likely due to sustained dependence on C-RAF/MEK/ERK activity, which can be activated as a result of the PI3K pathway in cell lines with constitutive activation of this pathway[46]. Collectively, these results highlight the advantage of a phenotypic assay to evaluate the efficacy of therapeutics in the complex and varied genetic landscapes present in tumour populations.

MAqCI has the potential to be used in the clinical setting. When an area of abnormal tissue is detected in the body of a patient, a biopsy is performed to determine whether the lesion is benign or malignant. If the lesion is malignant, traditional pathology cannot predict the metastatic propensity of the primary tumour. Here MAqCI can be applied to rapidly distinguish between aggressive and non-aggressive cancers to provide information about a patient's risk of metastasis and to potentially help generate individualized treatments. A small specimen of the cancerous lesion can be obtained via biopsy or as part of a resection surgery. Mass production of microfluidic devices with high-quality assurance using thermoplastics or other materials has been demonstrated by companies such as Optotrack and μFluidix. Lens-free imaging technology can be employed as an inexpensive way to monitor cell migration[33]. This technology provides high cell-to-background contrast that is amenable to automated analysis of the time-lapse images. Preclinical validation using a cohort of prospective patients is required to confirm the logistic regression coefficients and probability threshold used to determine patient prognosis. Given the promising performance of MAqCI in breast cancer, this functional assay might be relevant to other solid cancer types.

Methods

Microfluidic Assay for quantification of Cell Invasion (MAqCI). Microfluidic devices were fabricated and seeded with 50,000 cells as described previously[32,33,47]. Briefly, microchannels were patterned in polydimethylsiloxane (PDMS, Sylgard 184) by replica moulding using master moulds created on silicon wafers (Wafer World) by photolithography. Patterned PDMS was cleaned, activated by oxygen plasma in a Harrick PDC-32G plasma cleaner (Harrick Plasma) and bonded to glass slides (Electron Microscopy Sciences). Assembled microfluidic devices were immediately coated with 20 μg ml^{-1} collagen type I (BD Biosciences). Migration experiments were performed in DMEM plus 10% (v/v) fetal bovine serum (FBS, Life Technologies/Gibco) and 1% (v/v) penicillin/streptomycin (P/S, Gibco). Time-lapse images were recorded in 20 min intervals for 24 h using a Digital Sight Qi1Mc camera mounted on an inverted Eclipse Ti Microscope (Nikon) equipped with a 10×/0.30 numerical aperture lens.

Transwell-migration assay. Cell migration was monitored using the xCELLigence RTCA DP device (Acea Biosciences) using the manufacturer's protocol. Briefly, 40,000 cells per well were added to the upper chamber of a CIM-Plate 16. These plates have chambers that are similar to a Boyden chamber; they consist of an upper chamber where the cells are seeded in serum-free DMEM with 1% P/S, a microporous polyethylene terephthalate membrane (average pore diameter of 8 μm), electrodes directly below this membrane to detect cell migration, and a lower chamber that was filled with DMEM containing 10% FBS as a chemoattractant and 1% P/S. Migration progress was recorded in real time as cell index (a change in the electrical impedance between the electrodes caused by migrating cells). Measurements were taken every 15 min for 48 h in an incubator maintained at 37 °C and 5% CO_2. Each cell type was run in triplicate for each experiment.

Cell lines used in this study and viability measurements. Immortalized human mammary epithelial cells (HMLE) were transduced with a retrovirus carrying pMSCV-luciferase PGK-hygro expression vector (Addgene 18782, used for all luciferase-expressing cell lines). Single HMLE clones were isolated and the clones with the brightest bioluminescent signal were selected. To ensure that results were not due to clonal variance, two of the brightest clones, HMLE Luc 26 and 8, were selected at random and used in this project. These cells were grown in mammary epithelial cell growth medium (Lonza). 184B5 cells were grown in mammary epithelial cell growth medium plus 1 ng ml^{-1} cholera toxin

(Sigma-Aldrich). 184A1 cells were grown in mammary epithelial cell growth medium plus 1 ng ml^{-1} cholera toxin and 5 µg ml^{-1} transferrin (Sigma-Aldrich). MCF-10A cells and variants were created and cultured as described previously[23]. MCF-12F cells were a generous gift from D. Wirtz (Johns Hopkins University) and were grown in the same media as the MCF-10A cells. HCC1428 and ZR75-1 cells were grown in RPMI 1640 plus 10% and 1% P/S. MDA-MB-468, MCF7, T47D, MDA-MB-436, Hs578t, BT-549 and MDA-MB-231 cells were grown in DMEM (Life Technologies/Gibco) plus 10% FBS and 1% P/S. SkBr3 cells were grown in McCoy's 5A (modified) medium (Gibco) plus 10% FBS and 1% P/S. BT20 cells were grown in minimum essential medium Eagle (Sigma-Aldrich) plus 10% FBS and 1% P/S. SUM159 cells were a generous gift from D. Sharma (Johns Hopkins University); they and SUM149 cells were grown in Ham's F-12 medium (Corning Cellgro) plus 5% FBS, 1% P/S, 1 µg ml^{-1} hydrocortisone (Sigma-Aldrich) and 5 µg ml^{-1} insulin (Sigma-Aldrich). MCF7 targeted wild type and MCF7 Her2 were a generous gift from B. Ho Park (Johns Hopkins University) and were cultured as described previously[48]. The MCF7 cell genome contains two copies of an activating E545K mutation in the PIK3CA gene, and one wild-type copy of this gene. The MCF7 targeted wild-type cells have had these alleles corrected to wild type[48]. MCF7 Her2 cells have one wild-type allele of HER2 and one activating mutant V777L allele[48]. MCF7-luciferase cells were transduced with lentivirus for luciferase expression and have been shown to be non-metastatic. These cells were grown in DMEM plus 10% FBS, 1% P/S, and 5 µg ml^{-1} puromycin (Gibco). MDA-MB-231 tumour, lung met and CTC variants were grown in DMEM plus 10% FBS, 1% P/S and 125 µg ml^{-1} hygromycin (ThermoFisher). To produce these cell lines, MDA-MB-231 cells were transduced with lentivirus for luciferase expression and injected into the mammary fat pads of mice. Cells from the primary tumour were dissociated and human cells were selected with hygromycin (MDA-MB-231 tumour). Single cells were dissociated from secondary tumours formed in the lung and human cells were selected with hygromycin (MDA-MB-231 lung met). Blood samples were taken and a CTC cell was isolated (MDA-MB-231 CTC). Cells were maintained in a humidified incubator at 37 °C, 95% air/5% CO_2. Cells were routinely checked for mycoplasma contamination via PCR using the primers: F-(5′-GGGAGCAAACAGGATTAGATACCCT-3′) and R-(5′-TGCACCATCTGTCACTCTGTTAACCTC-3′). Unless otherwise specified, cell lines were purchased from the American Type Culture Collection. Cell viability measurements were performed as previously described[23].

Classification of cell lines. Cell lines were classified as having low or high metastatic potential based on reports of their ability to consistently metastasize in mouse models (spontaneously or following injection into the circulatory system)[49–52].

Identification of threshold. The threshold percentage of migratory cells and time were systematically varied over a range of 1–20% and 1–24 h. For each combination of values, the percentage of migratory cells from each cell line was compared with the threshold percentage to see whether MAqCI predicts that the cell line would have low metastatic potential (<threshold) or high metastatic potential (≥threshold). MAqCI's predictions were compared with our assessment of the cell lines (Supplementary Table 1) and used to classify each prediction as true positive, true negative, false positive or false negative (where true denotes a match between MAqCI's prediction and cell line classification and positive/negative denotes high/low metastatic potential, respectively). Using these classifications, the sensitivity (percentage true positive predictions out of number of cell lines with high metastatic potential), specificity (percentage true negative predictions out of number of cell lines with low metastatic potential) and accuracy (percentage of true predictions out of the total number of cell lines tested) of MAqCI's predictions were calculated for each combination of threshold and time.

Immunostaining. Cells were fixed, permeabilized, blocked against nonspecific adhesion, immunostained for target proteins and then imaged on an inverted Eclipse Ti epifluorescence microscope (Nikon). Primary antibodies were administered at manufacturer recommended concentration: anti-human CD45 (clone HI30, Becton Dickinson), anti-mouse CD45 (clone 30-F11, Becton Dickinson), anti-human mitochondria (clone 113-1, MilliporeSigma), anti-Ki-67 (clone 8D5, Cell Signaling Technology), and anti-phospho-histone H2A.X (Ser139, clone 20E3, Cell Signaling Technology). The secondary antibodies used were Alexa Fluor 488 goat anti-mouse immunoglobulin-G (IgG) (H+L) (Invitrogen) and Alexa Fluor 568 goat anti-rabbit IgG (H+L) (Invitrogen).

Logistic regression. Logistic regression was used to calculate the probability of a cell line having high metastatic potential based on the predictors (X_i) percentage migratory cells (X_1) and percentage Ki-67-positive cells (X_2) (equation (1)). Logistic regression coefficients (b_i) were trained in MATLAB using the glmfit function for the panel of 25 breast epithelial and breast cancer cell lines (Supplementary Table 2). Probability values were calculated in MATLAB using the glmval function.

$$p = \frac{e^{(b_0+b_1X_1+b_2X_2)}}{1+e^{(b_0+b_1X_1+b_2X_2)}} \times 100 \quad (1)$$

CellTracker labelling. MDA-MB-231 and MCF7 cells were labelled with either Red CMTPX dye or Green CMFDA dye (Invitrogen) according to the manufacturer's protocol. The colour assigned to each of the cell types was randomized for each experiment. Briefly, the medium was aspirated from 60–80% confluent cells in a T25. The cells were washed with PBS, and then incubated with 2 ml of DMEM containing 1 µM CellTracker dye at 37 °C for 30 min. After incubation, cells were passaged and used for the experiment.

Isolation of migratory cells. Fifty thousand MDA-MB-231 cells were seeded in MAqCI and incubated for 24 h. Migratory cells that exited the branch channels and entered the collection channel were washed with PBS and then detached from the device with 0.05% trypsin-EDTA (or TrypLE Express for RNA isolation experiments). Microbore tubing was attached to a collection inlet and connected to a sterile 0.2 µm filter (Corning) and a 10 ml syringe (Beckton Dickinson) containing DMEM plus 10% FBS and 1% P/S. Observing under a microscope, the medium was pushed from the syringe through the collection channel to force the migratory cells into the collection outlet. These cells were collected with a micropipette and transferred to a 15 ml Falcon tube that contained 5 ml of medium. Each microfluidic device typically yielded 50–100 migratory cells. Cells from multiple microfluidic devices were pooled to generate greater numbers of migratory cells. To collect a control population, cells from the seeding channel were dissociated and collected in a similar manner to the migratory cells. Since 50,000 cells were seeded in the MAqCI assay, thousands of cells adhere to the device, of which only hundreds are positioned close enough to the feeder channels to migrate, and of which an even smaller fraction are migratory, collection of cells from the seeding channel yields essentially the unsorted population, which contains both non-migratory (~80%) and migratory (~20%) cells. Thus, for this experiment, comparison was made between the migratory cells and the unsorted population.

Subcutaneous injection and bioluminescent imaging. All animal studies were performed following Institutional Animal Care and Use Committee procedures and guidelines at University of Maryland, Baltimore under an approved protocol. Eight-to-twelve-week-old female NOD/SCID mice weighing 19–25 g were obtained from Charles River and fed food and water ad libitum. Animals were randomly assigned to groups. For subcutaneous injections, equal numbers (500) of migratory or unsorted luciferase-tagged MDA-MB-231 cells were suspended in 100 µl PBS and mixed with an equal volume of Matrigel (Corning). Cell number was quantified via hemocytometer and confirmed by bioluminescent imaging.

Bioluminescence was only detected in viable cells expressing the firefly luciferase gene, indicative of an active metabolism. At the indicated time points following injection, mice were injected intraperitoneally with luciferin (150 mg kg^{-1}, Perkin Elmer) and returned to their cages for 5 min to allow for biodistribution. Mice were anaesthetised with 2% isoflurane gas and imaged at 5 min intervals for the maximum photon emission. Total 60 photon flux (photons s^{-1}) was calculated and corrected for tissue depth by spectral imaging using Living Image 3.0 software (IVIS). To quantify the bioluminescent signal from each organ, total bioluminescent signal was measured using regions of interest of equal areas for comparison between migratory and unsorted cell samples and the background signal was subtracted. To avoid false-positive detection of bioluminescent signal, the background subtracted value was normalized to the background reading and only readings that were ≥50× the background reading were considered to be positive for metastasis. To control for differences in tumour size, bioluminescent values were normalized to the volume of the primary tumour at the time of necropsy.

Quantitative PCR. Quantitative PCR for human long interspersed nuclear elements was conducted as described previously[19]. Dilutions of human DNA isolated from MDA-MB-231 cells were included in each plate for comparison. Primers used for verification of DEG were: *FGF5* F-(5′-CTCTTCCCCTCTCCCCTTCT-3′) and R-(5′-GGCTGATTCTGGGCTCTGTA-3′), *CCND1* F-(5′-GGATGCTGGAGGTCTGCGA-3′) and R-(5′-AGAGGCCACGAACATGCAAG-3′), *TAB2* F-(5′-GACCTGCCCTGGAAAAGAAGT-3′) and R-(5′-CTCTTCTGTCCAAGCATTTTCTGG-3′), *AKT2* F-(5′-GCTCCACAAGCGTGGTGAAT-3′) and R-(5′-GGCCTCTCGGTCTTCATCAG-3′), *KRAS* F-(5′-CATTGGTGAGGGAGATCCGA-3′) and R-(5′-AGGCATCATCAACACCCAGAT-3′) and *FOS* F-(5′-TAGTTAGTAGCATGTTGAGCCAGG-3′) and R-(5′-ACCACCTCAACAATGCATGA-3′).

Immunohistochemistry. Animals with primary tumour formation that exceeded the designated endpoint, including saturation exceeding 1,000-fold over the initial bioluminescence signal, were killed. Tissue samples were removed, fixed in formalin for 24 h, embedded in paraffin wax and serially sectioned (4 µm thick). All immunohistochemistry and hematoxylin and eosin staining was performed by Mass Histology Services.

Cell tracking. Cells were tracked manually every 20 min using ImageJ software (National Institutes of Health) using the Manual Tracking plugin[47]. In select experiments, polygonal regions of interests (ROIs) were manually drawn around the cell periphery in 40 min intervals using ImageJ and saved to the ROIs Manager.

The Measure ROI function was then used to calculate the cell area, aspect ratio and solidity[47]. Values of each metric were averaged over the time the cell spent in the feeder channel.

RNA-seq and analysis. RNA was isolated from samples of 1,000 migratory or unsorted cells in triplicate using the Nucleospin RNA XS kit (Macherey-Nagel). Complementary DNA libraries were amplified using the SMART-seq ultralow input RNA kit (Takara) and then tagmented and barcoded by indexing primers using Nextera XT DNA library prep kit (Illumina). Samples were pooled and paired-end sequenced on an Illumina NextSeq 500 using the NextSeq 500/550 Mid Output v2 kit with 150 cycles and an output of up to 130 million reads. After quality control of raw data using Illumina pipeline, RNA-seq reads were mapped to hg38 reference genome using HISAT2[53] aligner. Htseq-count command from the HTSeq framework[54] was used to quantify read counts per gene from aligned reads using human ENSEMBL 86 (GRCh38.p7) gene models. The Bioconductor/R package DESeq2[55] was used for normalization and differential gene expression analysis. Pathway analysis and gene ontology clustering were performed using the Database for Annotation, Visualization, and Integrated Discover (DAVID)[56,57].

DNA damage quantification. Cells were imaged on a Nikon A1 confocal microscope using a Plan Apo 60× objective (numerical aperture, 1.4). ImageJ was used to convert images to 8 bit, set a binary threshold, and quantify the number and area of foci.

Tail-vein injection. Eight-to-twelve-week-old female athymic nude-Foxn1nu mice weighing 19–25 g were obtained and fed as described above. MCF-10A, 10A-KRAS(G12V), PTEN and PTEN$^{-/-}$KRAS(G12V) cells (1×10^6 cells 100 µl^{-1}) were injected intravenously in cohorts of 5 animals. Animals were monitored by bioluminescence signal and carried out through 20 weeks or until recurrent growth (observed only in mice injected with PTEN$^{-/-}$KRAS cells) required the mouse to be sacrificed. After 20 weeks, mice injected with MCF-10A, PTEN$^{-/-}$ or KRAS(G12V) cells still survived and were monitored for recurrence and clinical distress by gross examination.

PDXs. Cryogenically preserved tumour specimens were implanted into the cleared mammary fat pads of NOD/SCID mice. Both samples formed tumours, which were excised, dissociated to single cells as described previously[58], and allowed to recover on collagen-I coated tissue culture dishes in DMEM plus 10% FBS and 1% P/S for 48 h. Cells were then seeded in MAqCI and monitored via time-lapse microscopy for 24 h. After imaging, cells were fixed and immunostained for human mitochondria and other markers. To exclude the possibility of the dissociated cells being human leukocytes preserved during resection from the patient, select samples were immunostained for CD45 and were negative in all cases (data not shown). To assess Ki-67-positive cells, previously stained samples were quenched with 10 mg ml^{-1} sodium borohydride (Sigma-Aldrich) in PBS for 15 min. Samples were then rinsed thoroughly with PBS, blocked and immunostained for Ki-67 and human mitochondria (AB3598, MilliporeSigma).

Western blotting. Western blots were performed as described previously[59]. Primary antibodies were administered at the manufacturer recommended concentration: anti-actin (clone Ab-5, Becton Dickinson), anti-AKT (clone C67E7, Cell Signaling Technologies), anti-pAKT (Ser473) (clone D9E, Cell Signaling Technologies), anti-ERK (clone L34F12, Cell Signaling Technologies), anti-pERK (clone D13.14.4E, Cell Signaling Technologies). Secondary antibodies: anti-mouse IgG horseradish peroxidase-linked antibody (Cell Signaling Technologies), anti-rabbit IgG horseradish peroxidase-linked antibody (Cell Signaling Technologies).

Statistical methods. Data means ± s.e.m were calculated and plotted using GraphPad Prism 7 (GraphPad Software). The D'Agostino–Pearson omnibus normality test was used to determine whether data were normally distributed. Datasets with Gaussian distributions were compared using an unpaired Student's *t*-test (two-tailed) or analysis of variance (ANOVA) followed by Tukey's multiple comparisons test. Datasets with non-Gaussian distributions were compared using an unpaired Mann–Whitney test (two-tailed). Test statistics and *P* values for each test are reported in Supplementary Dataset 3.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The main data supporting the results of this study are available within the paper and its Supplementary Information files. Source data for the figures in this study are available from the corresponding author upon reasonable request. RNA sequencing data are available at the National Center for Biotechnology Information Gene Expression Omnibus, under accession number GSE128313.

Received: 15 August 2018; Accepted: 5 April 2019;
Published online: 6 May 2019

References

1. Steeg, P. S. Targeting metastasis. *Nat. Rev. Cancer* **16**, 201–218 (2016).
2. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018. *CA Cancer J. Clin.* **68**, 7–30 (2018).
3. Harms, W. et al. DEGRO practical guidelines for radiotherapy of breast cancer VI: therapy of locoregional breast cancer recurrences. *Strahl. Onkol.* **192**, 199–208 (2016).
4. Paik, S. et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **351**, 2817–2826 (2004).
5. Nagrath, S. et al. Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* **450**, 1235–1239 (2007).
6. Lippman, M. & Osborne, C. K. Circulating tumor DNA—ready for prime time? *N. Engl. J. Med* **368**, 1249–1250 (2013).
7. Chandler, Y. et al. Cost effectiveness of gene expression profile testing in community practice. *J. Clin. Oncol.* **36**, 554–562 (2018).
8. Alix-Panabières, C. & Pantel, K. Clinical applications of circulating tumor cells and circulating tumor DNA as liquid biopsy. *Cancer Discov.* **6**, 479–491 (2016).
9. Garcia-Murillas, I. et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci. Transl. Med.* **7**, 302ra133 (2015).
10. Riggi, N., Aguet, M. & Stamenkovic, I. Cancer metastasis: a reappraisal of its underlying mechanisms and their relevance to treatment. *Annu. Rev. Pathol.* **13**, 117–140 (2018).
11. Paul, C. D., Mistriotis, P. & Konstantopoulos, K. Cancer cell motility: lessons from migration in confined spaces. *Nat. Rev. Cancer* **17**, 131–140 (2017).
12. Wolf, K. et al. Collagen-based cell migration models in vitro and in vivo. *Semin. Cell Dev. Biol.* **20**, 931–941 (2009).
13. Fidler, I. J. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat. Rev. Cancer* **3**, 453–458 (2003).
14. Irianto, J. et al. Nuclear constriction segregates mobile nuclear proteins away from chromatin. *Mol. Biol. Cell* **27**, 4011–4020 (2016).
15. Irianto, J. et al. DNA damage follows repair factor depletion and portends genome variation in cancer cells after pore migration. *Curr. Biol.* **27**, 210–223 (2017).
16. Abubakar, M. et al. Prognostic value of automated KI67 scoring in breast cancer: a centralised evaluation of 8088 patients from 10 study groups. *Breast Cancer Res.* **18**, 104 (2016).
17. Cidado, J. et al. Ki-67 is required for maintenance of cancer stem cells but not cell proliferation. *Oncotarget* **7**, 6281–6293 (2016).
18. Duval, K. et al. Modeling physiological events in 2D vs. 3D cell culture. *Physiology (Bethesda)* **32**, 266–277 (2017).
19. Dallas, M. R. et al. Divergent roles of CD44 and carcinoembryonic antigen in colon cancer metastasis. *FASEB J.* **26**, 2648–2656 (2012).
20. López-Knowles, E. et al. PI3K pathway activation in breast cancer is associated with the basal-like phenotype and cancer-specific mortality. *Int J. Cancer* **126**, 1121–1131 (2010).
21. McLaughlin, S. K. et al. The RasGAP gene, RASAL2, is a tumor and metastasis suppressor. *Cancer Cell* **24**, 365–378 (2013).
22. Giltnane, J. M. & Balko, J. M. Rationale for targeting the Ras/MAPK pathway in triple-negative breast cancer. *Discov. Med* **17**, 275–283 (2014).
23. Thompson, K. N. et al. The combinatorial activation of the PI3K and Ras/MAPK pathways is sufficient for aggressive tumor formation, while individual pathway activation supports cell persistence. *Oncotarget* **6**, 35231–35246 (2015).
24. DeRose, Y. S. et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat. Med.* **17**, 1514–1520 (2011).
25. Dobrolecki, L. E. et al. Patient-derived xenograft (PDX) models in basic and translational breast cancer research. *Cancer Metastasis Rev.* **35**, 547–573 (2016).
26. Rouzier, R. et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin. Cancer Res.* **11**, 5678–5685 (2005).
27. Prat, A. et al. Research-based PAM50 subtype predictor identifies higher responses and improved survival outcomes in HER2-positive breast cancer in the NOAH study. *Clin. Cancer Res.* **20**, 511–521 (2014).
28. Leonowens, C. et al. Concomitant oral and intravenous pharmacokinetics of trametinib, a MEK inhibitor, in subjects with solid tumours. *Br. J. Clin. Pharm.* **78**, 524–532 (2014).
29. Csonka, D. et al. A phase-1, open-label, single-dose study of the pharmacokinetics of buparlisib in subjects with mild to severe hepatic impairment. *J. Clin. Pharm.* **56**, 316–323 (2016).
30. Hollestelle, A., Elstrodt, F., Nagel, J. H., Kallemeijn, W. W. & Schutte, M. Phosphatidylinositol-3-OH kinase or RAS pathway mutations in human breast cancer cell lines. *Mol. Cancer Res.* **5**, 195–201 (2007).
31. Zimmermann, S. & Moelling, K. Phosphorylation and regulation of Raf by AKT (protein kinase B). *Science* **286**, 1741–1744 (1999).
32. Tong, Z. et al. Chemotaxis of cell populations through confined spaces at single-cell resolution. *PLoS ONE* **7**, e29211 (2012).

33. Mathieu, E. et al. Time-lapse lens-free imaging of cell migration in diverse physical microenvironments. *Lab Chip* **16**, 3304–3316 (2016).
34. Chen, Y. C. et al. Functional isolation of tumor-initiating cells using microfluidic-based migration identifies phosphatidylserine decarboxylase as a key regulator. *Sci. Rep.* **8**, 244 (2018).
35. Song, W. et al. Targeting EphA2 impairs cell cycle progression and growth of basal-like/triple-negative breast cancers. *Oncogene* **36**, 5620–5630 (2017).
36. Camarda, R. et al. Inhibition of fatty acid oxidation as a therapy for MYC-overexpressing triple-negative breast cancer. *Nat. Med.* **22**, 427–432 (2016).
37. Mulholland, D. J. et al. Pten loss and RAS/MAPK activation cooperate to promote EMT and metastasis initiated from prostate cancer stem/progenitor cells. *Cancer Res.* **72**, 1878–1889 (2012).
38. Mendoza, M. C., Er, E. E. & Blenis, J. The Ras-ERK and PI3K-mTOR pathways: cross-talk and compensation. *Trends Biochem. Sci.* **36**, 320–328 (2011).
39. Bedard, P. L. et al. A phase Ib dose-escalation study of the oral pan-PI3K inhibitor buparlisib (BKM120) in combination with the oral MEK1/2 inhibitor trametinib (GSK1120212) in patients with selected advanced solid tumors. *Clin. Cancer Res.* **21**, 730–738 (2015).
40. Ridley, A. J. et al. Cell migration: integrating signals from front to back. *Science* **302**, 1704–1709 (2003).
41. Toker, A. & Yoeli-Lerner, M. AKT signaling and cancer: surviving but not moving on. *Cancer Res.* **66**, 3963–3966 (2006).
42. Huang, C., Jacobson, K. & Schaller, M. D. MAP kinases and cell migration. *J. Cell Sci.* **117**, 4619–4628 (2004).
43. Cheng, H. et al. PIK3CA^{H1047R} and Her2 initiated mammary tumors escape PI3K dependency by compensatory activation of MEK-ERK signaling. *Oncogene* **35**, 2961–2970 (2016).
44. Hoeflich, K. P. et al. In vivo antitumor activity of MEK and phosphatidylinositol 3-kinase inhibitors in basal-like breast cancer models. *Clin. Cancer Res.* **15**, 4649–4664 (2009).
45. Butler, D. E. et al. Inhibition of the PI3K/AKT/mTOR pathway activates autophagy and compensatory Ras/Raf/MEK/ERK signalling in prostate cancer. *Oncotarget* **8**, 56698–56713 (2017).
46. Ebi, H. et al. PI3K regulates MEK/ERK signaling in breast cancer via the Rac-GEF, P-Rex1. *Proc. Natl Acad. Sci. USA* **110**, 21124–21129 (2013).
47. Paul, C. D. et al. Interplay of the physical microenvironment, contact guidance, and intracellular signaling in cell decision making. *FASEB J.* **30**, 2161–2170 (2016).
48. Zabransky, D. J. et al. HER2 missense mutations have distinct effects on oncogenic signaling and migration. *Proc. Natl Acad. Sci. USA* **112**, E6205–E6214 (2015).
49. Sflomos, G. et al. A preclinical model for ERα-positive breast cancer points to the epithelial microenvironment as determinant of luminal phenotype and hormone response. *Cancer Cell* **29**, 407–422 (2016).
50. Jiang, Y., Woosley, A. N., Sivalingam, N., Natarajan, S. & Howe, P. H. Cathepsin-B-mediated cleavage of Disabled-2 regulates TGF-β-induced autophagy. *Nat. Cell Biol.* **18**, 851–863 (2016).
51. Rizwan, A. et al. Breast cancer cell adhesome and degradome interact to drive metastasis. *NPJ Breast Cancer* **1**, 15017 (2015).
52. Wiegmans, A. P. et al. Rad51 supports triple negative breast cancer metastasis. *Oncotarget* **5**, 3261–3272 (2014).
53. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
54. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
55. 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).
56. Huang, dW., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1–13 (2009).
57. Huang, dW., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
58. DeRose, Y. S. et al. Patient-derived models of human breast cancer: protocols for in vitro and in vivo applications in tumor biology and translational medicine. *Curr. Protoc. Pharmacol.* **60**, 14.23.1–14.23.43 (2013).
59. Shea, D. J., Li, Y. W., Stebe, K. J. & Konstantopoulos, K. E-selectin-mediated rolling facilitates pancreatic cancer cell adhesion to hyaluronic acid. *FASEB J.* **31**, 5078–5086 (2017).

Acknowledgements

This line of research was supported by the National Cancer Institute through grants R01-CA183804 (K.K., A.K.-K., S.S.M.), R01-CA216855 (K.K.), R01-CA154624 (S.S.M.), R01-CA174385 (N.V.) and K01-CA166576 (M.I.V.), as well as by CPRIT RP180466 (N.V.), MRA Award 509800 (N.V.), CDMRP CA160591 (N.V.) and Department of Defense grant W81XWH-17-1-0246 (V.K.B.). M.I.V. was also supported by a Research Scholar Grant, RSG-18-028-01-CSM, from the American Cancer Society.

Author contributions

C.L.Y., C.D.P. and K.K. designed the study. C.L.Y. performed experiments, interpreted the data and wrote the manuscript. K.N.T., C.D.P., M.I.V. and P.M. contributed to design the study, performed experiments and interpreted the data. A.M. and V.K.B. helped to design, perform and analyse the RNA sequencing experiments. D.J.S. and K.M.M. performed select experiments. A.C.C. wrote code and used it to analyse data. N.V., A.K.-K. and S.S.M. interpreted data, provided critical insights and edited the manuscript. K.K. designed and supervised the study, and wrote the manuscript.

Competing interests

The PTEN$^{-/-}$ cells are licensed to Horizon Discovery Ltd (Cambridge, UK). M.I.V receives compensation for the sale of these cells. MAqCI is the subject of US Utility Patent applications 15/780,768 and 14/906,055.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41551-019-0400-9.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to K.K.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Konstantinos Konstantopoulos

Last updated by author(s): Apr 3, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Data were collected by using standard features of NIS Elements Viewer (version 4.11.0) and FIJI (Version 1.0).
Data analysis	Data were primarily collected and organized in Microsoft Excel (Version 16.14.1) and Graphpad Prism (Version 7.0a). Select analysis were performed in MATLAB (Version R2018a, 9.4.0.813654), as described in Methods.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The main data supporting the results of this study are available within the Article and its Supplementary Information. Source data for the figures in this study are available from the corresponding author upon reasonable request. RNA-seq data is available at NCBI GEO under accession number GSE128313.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was chosen based on the throughput of technique used. Sample sizes were sufficient to show the same trends between the three or more replicates performed for each experiment, and by statistical testing.
Data exclusions	Data were not excluded from analysis.
Replication	Each experiment was repeated 3 or more times, with similar results observed each time.
Randomization	Mice were assigned randomly to experimental groups.
Blinding	Blinding was not performed for this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For immunostaining: Primary antibodies were administered at manufacturer recommended concentration: anti-CD45 (clone HI30, Becton Dickinson, 555482, 1:100), anti-CD45 (clone 30-F11, Becton Dickinson, 553080), anti-human mitochondria (clone 113-1, MilliporeSigma, MAB1273, 1:100, lot#2722860), anti-Ki67 (clone 8D5, Cell Signaling Technology, 9449, 1:800, lot#4), and anti-phospho-histone H2A.X (Ser139, clone 20E3, Cell Signaling Technology, 9718, 1:400, lot#13). The secondary antibodies used were Alexa Fluor 488 goat anti-mouse IgG (H+L) (Invitrogen, A28175, 1:100) and Alexa Fluor 568 goat anti-rabbit IgG (H+L) (Invitrogen, A11011, 1:100). For western blotting: Primary antibodies were administered at the manufacturer recommended concentration: anti-actin (clone Ab-5, Becton Dickinson, 612656, 1:10000), anti-Akt (clone C67E7, Cell Signaling Technologies, 4691, 1:1000, lot#20), anti-p-Akt (Ser473) (clone D9E, Cell Signaling Technologies, 4060, 1:2000), anti-Erk (clone L34F12, Cell Signaling Technologies, 4696, 1:2000, lot#22), anti-p-Erk (clone D13.14.4E, Cell Signaling Technologies, 4370, 1:2000, lot#12). Secondary antibodies: anti-mouse IgG HRP-linked antibody (Cell Signaling Technologies, 7076, 1:1000), anti-rabbit IgG HRP-linked antibody (Cell Signaling Technologies, 7074, 1:1000).
Validation	Antibodies were validated by the manufacturer prior to purchasing. Additionally, antibodies were verified for western blotting based on the correct molecular weight identified and for immunofluorescence by comparison of their cellular distribution to that obtained by the manufacturer.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Parental cell lines were purchased from American Type Culture Collection (ATCC). Select cell lines were modified from the parental line, as described in Methods.

Authentication	Cell lines were originally authenticated by ATCC, and were not further authenticated as part of this study.
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination, as assessed by PCR.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	For mammary fat pad injection, Eight-to-twelve-week-old female NOD SCID mice weighing 19-25g were obtained from Charles River (Fredrick, MD). For tail vein injection, Eight-to-twelve-week-old female athymic nude-Foxn1nu mice weighing 19-25g were obtained Charles River (Fredrick, MD).
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve samples collected from the field.
Ethics oversight	All animal studies were performed following Institutional Animal Care and Use Committee procedures and guidelines at the University of Maryland, Baltimore, under an approved protocol.

Note that full information on the approval of the study protocol must also be provided in the manuscript.