E-cadherin promotes cell hyper-proliferation in breast cancer

Gabriella C. Russo[1,2], Michelle N. Karl[1,2], David Clark[3], Julie Cui[1], Ryan Carney[4], Boyang Su[5], Bartholomew Starich[1,2], Tung-Shing Lih[3], Qiming Zhang[1,2], Pei-Hsun Wu[1,2], Meng-Horng Lee[1], Hon S. Leong[5,6], Hui Zhang[3], and Denis Wirtz[1,2]*

[1]*Department of Chemical and Biomolecular Engineering, Johns Hopkins University, 3400 N Charles St, Baltimore, Maryland 21218, USA*

[2]*Johns Hopkins Physical Sciences– Oncology Center and Institute for NanoBioTechnology, Johns Hopkins University, 3400 N Charles St, Baltimore, Maryland 21218, USA*

[3]*Department of Pathology, The Johns Hopkins University School of Medicine, Baltimore 21231, Maryland, USA*

[4]*Department of Biophysics, Johns Hopkins University, 3400 N Charles St, Baltimore. MD 21218, USA*

[5]*Department of Medical Biophysics, University of Toronto, Toronto, ON*

[6]*Biological Sciences Platform, Sunnybrook Research Institute, Toronto, ON*

ABSTRACT

The loss of the intercellular adhesion molecule E-cadherin is a hallmark of the epithelial-mesenchymal transition (EMT), which promotes a transition of cancer cells to a migratory and invasive phenotype. E-cadherin is associated with a decrease in cell proliferation in normal cells. Here, using physiologically relevant 3D *in vitro* models, we find that E-cadherin induces hyper-proliferation in breast cancer cells through activation of the Raf/MEK/ERK signaling pathway. These results were validated and consistent across multiple *in vivo* models of primary tumor growth and metastatic outgrowth. E-cadherin expression dramatically increases tumor growth and, without affecting the ability of cells to extravasate and colonize the lung, significantly increases macrometastasis formation via cell proliferation at the distant site. Pharmacological inhibition of MEK1/2, blocking phosphorylation of ERK in E-cadherin-expressing cells, significantly depresses both tumor growth and macrometastasis. This work suggests a novel role of E-cadherin in tumor progression and identifies a potential new target to treat hyper-proliferative breast tumors.

SUMMARY

E-cadherin, an extensively studied transmembrane molecule ubiquitously expressed in normal epithelial tissues, promotes and maintains intercellular adhesion. In cancer, the loss of adhesion molecule E-cadherin is associated with onset of invasion via epithelial-to-mesenchymal transition (EMT) process.[1] EMT consists of a highly orchestrated cascade of molecular events where epithelial cells switch from a non-motile phenotype to an invasive, migratory phenotype accompanied by a change in cell morphology.[1,2] These processes are believed to then trigger metastasis in carcinomas (cancers of epithelial origin). Moreover, the expression of intercellular adhesion molecule E-cadherin (E-cad) is associated with a decrease in cell proliferation in normal cells. Classical experiments in fibroblasts and epithelial cells show that the expression of E-cad not only promotes cell-cell adhesion, but also reduces cell proliferation and onset of apoptosis.[3,4] Altogether these results have long supported that E-cad acts as a tumor suppressor gene.[1,2]

However, despite its role in cell-adhesion the requirement for loss of E-cad in metastasis has recently been re-assessed.[5,6,7,8]

These investigations focus on E-cad's role in EMT, even though the relationship between E-cad and proliferation is just as intriguing. While E-cad has been shown to have anit-proliferative effects in normal cells, E-cad also helps maintain a pluripotent and proliferative phenotype in stem cells, and notably is lost during differentiation, a non-proliferative step of stem cell progression.[9,10] Yet, despite potentially important implications in our understanding of tumor progression, whether E-cad expression affects growth in cancer cells remains mostly unexplored.

Here, utilizing a physiologically relevant 3D *in vitro* model and multiple *in vivo* models, we studied the impact of E-cad on cell proliferation at the primary tumor site and proliferation at a secondary site. Remarkably, E-cad upregulates multiple proliferation pathways, including hyper-activation of the ERK cascade within the greater MAPKinase pathway, resulting in a dramatic increase in cell proliferation *in vitro* and tumor growth *in vivo*. When the phosphorylation of ERK is blocked utilizing a MEK1/2 inhibitor, PD0325901[11], this effect is reversed *in vitro* and *in vivo*. Thus, E-cad plays an oncogenic role in tumorigenesis and merits evaluation as a potential new drug target.

MAIN

We began our exploration into whether and how E-cad may impact proliferation by determing the endogenous expression of E-cad in seven commonly used breast cell lines, across multiple breast-cancer subtypes, including triple negative, Her2-receptor-positive, and hormone-receptor-positive breast cancer (Fig S1c). After testing, we chose three cell lines that were suitable for *in vivo* modeling. As previously shown, MCF7 and MDA-MB-468 are E-cad positive (denoted E-cad^{+}). MDA-MB-231 is E-cad negative (E-cad^{-}) and is known for its highly proliferative and metastatic phenotypes[12,13]. Next, we generated E-cad shRNA lentiviral knockdown cell line pairs for MDA-MB-468 and MCF7 cells and an E-cad lentiviral knock-in for MDA-MB-231 cells. After confirmation of 80% or more depletion of E-cad in the knock-down cells and high E-cad expression in the knock-in cells via western blots (Fig. S1a), we also determined functionality changes in modified cell lines by assessing the impact of E-cad on the localization of β-catenin, which links the cytoplasmic domain of E-cad to the actin cytoskeleton.[14] Immunofluorescence microscopy showed that when E-cad was naturally absent or depleted, β-catenin accumulated in and around the nucleus, as anticipated (Fig. S1b).[14] *Vice versa*, when E-cad was naturally present or induced, we observed an accumulation of cytosolic and membrane-bound β-catenin (Fig. S1b).

Since MDA-MB-231 cancer cells are widely used due to their ability to grow rapidly *in vivo*, our expectation was that ectopic addition of E-cad into MDA-MB-231 cells would have minimal impact on tumor growth[15]. We orthotopically injected E-cad^{+}/E-cad^{-} MDA-MB-231 cells in the mammary fat pad of NOD-SCID Gamma (NSG) mice utilizing a bilateral approach to directly compare tumor size (Fig. 1a). Remarkably, we observed a

three-fold increase in the growth rate of tumors when E-cad expression was induced (Fig. 1b) as well as a six-fold increase in the final tumor weight (measured after tumor excision) (Fig. 1c, Fig. S2a). We observed the same trend in MDA-MB-468 and MCF7 tumors: E-cad^{+} tumor growth rate was significantly higher than E-cad^{-} tumor growth rate (Fig. 1e-h, Fig. S2 b,c). We also stained for proliferation marker Ki67 and observed more Ki67 positive cells in the E-cad^{+} tumors when compared to E-cad^{-} tumor, e.g. shown in (Fig. 1d). To ensure there was no crosstalk between the two implanted tumors, we also implanted a single bolus of cancer cells per mouse and observed the same trend for growth and final tumor size, analogous to the average size in the bilateral study (Fig. S1l).

To determine whether E-cad-induced proliferation *in vivo* was an endogeneous cell effect, we assessed proliferation in standard culture dishes. We observed a slight but significant increase in the proliferation rate of E-cad-bearing cells compared to cells lacking E-cad (Fig. S1d). To ensure that these observations were not an artifact of 2D cell culture, we further validated our *in vivo* results using a 3D double-layer spheroid system. Cancer cells were enclosed in a basement membrane material core (Matrigel) surrounded by a collagen corona to respectively mimic the basement membrane and the surrounding collagen-rich stroma of solid breast tumors (Fig. 1i). This assay allowed us to monitor the effect of E-cad on proliferation, as well as cell invasion and their interaction with the outer collagen layer. Here, we observed a greater and more significant impact in cell growth when E-cad was expressed in all tested cell pairs (Fig. 1j). Time-lapsed phase-contrast microscopy also showed that E-cad prevented single-cell dissemination from the core of the spheroid into the collagen layer but did not necessarily prevent invasion completely (Fig. S1h). This suggests our manipulation of E-cad affects the migration as expected.

Together these results suggest that E-cad is pro-tumorigenic, promoting cell proliferation and tumor growth.

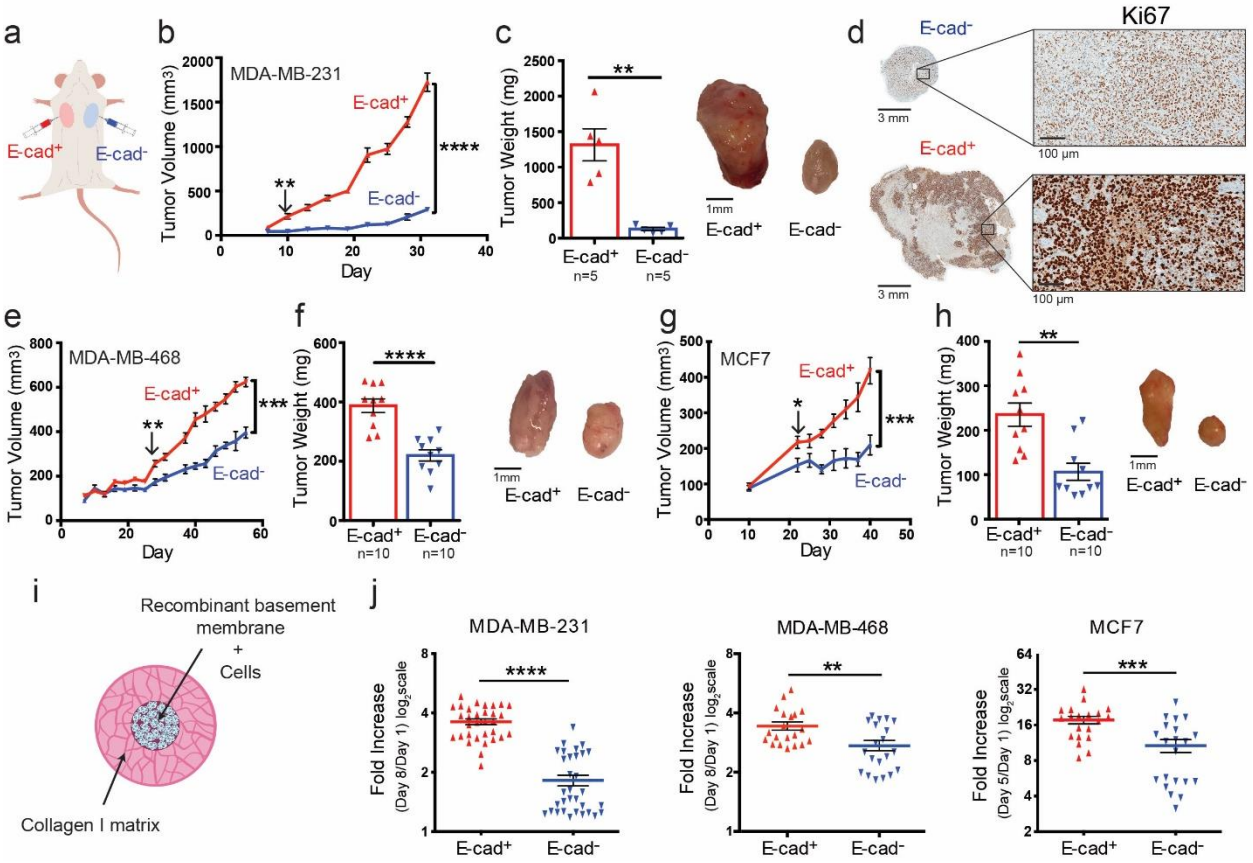

Figure 1: E-cadherin increases tumor progression *in vivo* and *in vitro*

(a) Schematic of bilateral mammary fat pad injection for *in vivo* modeling of primary tumor growth. (b, e and g) Time-dependent volume of control MDA-MB-231 scramble and E-cad^{+} MDA-MB-231 tumors (****P< 0.0001) (b), control MDA-MB-468 scramble and E-cad^{-} shRNA MDA-MB-468 tumors (***P= 0.0009) (e), and control MCF7 scramble and E-cad^{-} shRNA MCF7 tumors (***P= 0.0003) (g). Apparent tumor volume was assessed using calipers to measure x and y dimensions. (c, f, and h) Tumor weight at the end of the study

with pictures of representative excised tumors. Scale bars, 1 mm. **P=0.0079, ****P<0.0001, **P=0.0015 respectively. (d) IHC staining of proliferative marker Ki67 in corresponding tumors. (i) Schematic of the 3D spheroid system composed of Matrigel in the inner core and collagen I in the outer layer. (j) PrestoBlue fold increase of MDA-MB-231 (n= 60) ****P=<0.0001, MDA-MB-468 (n= 60) **P=0.0073, and MCF7 (n= 50) ***P = 0.0008 comparing spheroids on day 1 to day 8 or day 1 to day 5. All data are mean ± SEM. P values determined from paired T.Test for mice, and Mann-Whitney test, two sided for spheroids.

E-cadherin promotes growth at the distant metastatic site

Next, we sought to explore the role of E-cad in metastatic outgrowth, as we wanted to see if this phenomenon is preserved after cells arrive at a distant site. To assess how E-cad impacts proliferation at the secondary site, i.e. the lung for breast cancer in mice, we designed an experimental metastasis assay, in which cancer cells are “forced” to the secondary site via tail vein injection. (Fig. 2a). We injected either MDA-MB-231 E-cad^{+} cells or E-cad^{-} cells into the tail veins of NSG mice and assessed metastatic burden in the lung 4 weeks after injection. We chose an extended timeline to minimize effects of survival or colonization ability and focus on metastatic growth.

After four weeks, we observed many more macrometastases in the lungs of mice injected with E-cad^{+} cells (Fig. S3b). These observations were confirmed via RT-qPCR: we assessed the expression of a human genomic marker *HK2*, which is often used to quantify the metastastic burden in human cancer cells in the lungs of NSG mice.[16] We observed a remarkable 109-fold increase in *HK2* expression in the lungs of mice injected

with E-cad^{+} cells (Fig. 2b). These results were further confirmed via H&E staining of lung sections, where dense foci of cancer cells were present in the lungs of mice injected with E-cad^{+} cells (Fig. 2c). We also observed that a majority of the cells in the lungs of mice injected with E-cad^{+} cells were positive for Ki67 (Fig. 2d).This was not true when the mice were injected with E-cad^{-} cells. This data suggests that E-cad may also play a role in promoting proliferation at a secondary site.

While the tail vein model is a well-established experimental assay to assess metastatic burden, it has shortcomings; predominantly that cells cannot be tracked in real-time from their injection site (the tail vein) to the lungs. To confirm that our observations from the tail vein model were not due to enhanced colonization or survival ability of E-cad^{+} cancer cells, we utilized the chicken chorioallantonic membrane (CAM) assay, in which intravital microscopy can be used to track cells as they extravasate, colonize, and proliferate (Fig. 2e).[17,18,19] To fully control the impact E-cad expression has on the cellular phenotype, we generated a MDA-MB-231 cell line that would only express E-cad after receiving an engineered stimulus. This tunable E-cad-zsGreen cell line gains E-cad expression when activated by the cell-permeant high-affinity ligand Shield-1.[20]

First, we sought to determine if E-cad influenced extravasation rates, which would be an indicator of differential survival and colonization ability. In all cell lines (MDA-MB-231 E-cad^{+}/E-cad^{-} cell line pair and the MDA-MB-231 E-cad-zsGreen cell line treated with Shield-1), we did not observe a statistically significant difference in the percentage of cells that successfully extravasated into the stroma of the CAM (Fig. 2f). Next, we assessed the ability of these cells to form metastases. We set up three experimental conditions: (1) cells that did not express E-cad (E-cad^{-} cells and cells expressing tunable E-cad treated

with vehicle control-DMSO), (2) cells that expressed E-cad all the time (E-cad$^+$ cells), and (3) cells that did not express E-cad until after extravasation into the stroma (tunable cell line treated with Shield-1). Fluorescent images of CAM showed larger colonies in embryos injected with E-cad$^+$ cells and the tunable cell line treated with Shield-1 to induce E-cad expression when compared to E-cad$^-$ cells and embryos treated with DMSO (Fig. 2g). When quantified, the E-cad$^+$ conditions had a 1.5 to 2-fold increase in the number of metastases present throughout the CAM when compared to those lacking E-cad expression (Fig. 2h). This confirms that E-cad expression promotes growth at a secondary site, as observed in our mouse models, but does not significantly affect the ability of cells to extravasate or colonize the secondary site (i.e. lungs).

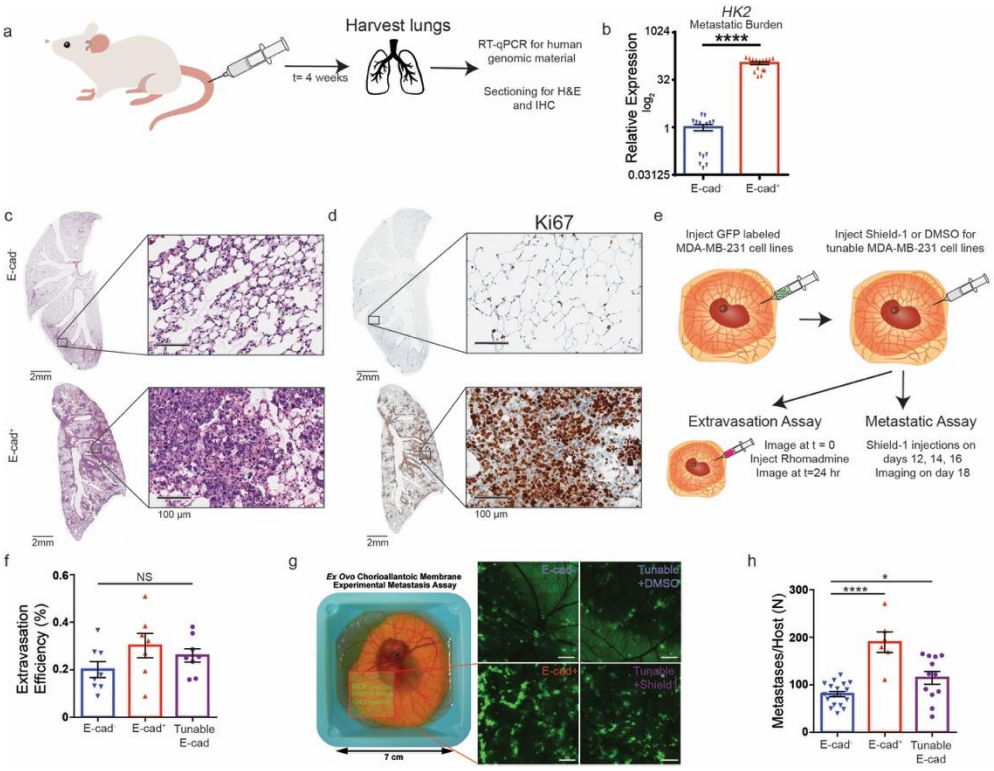

Figure 2: E-cadherin promotes growth at a secondary site. (a) Schematic of tail vein injection of breast cancer cells. (b) Relative expression level of *HK2* a marker of human genomic material comparing lungs of mice injected with E-cad$^+$ to E-cad$^-$ cells (****P<0.0001), Mann-Whitney. (c) H&E staining of lungs from E-cad$^+$ and E-cad$^-$ mice. (d) IHC staining for Ki67 in the lungs of from E-cad$^+$ and E-cad$^-$ mice. (e) Schematic of intravital microscopy experiment. (f) Extravasation rate of E-cad$^-$, E-cad$^+$, and tunable E-cad MDA-MB-231 cells. (g) Confocal images of embryos on day 18 (post injection) showing greater accumulation of GFP tagged E-cad$^+$ or Shield-1 treated tunable cells when compared to the scramble control (E-cad$^-$) or DMSO treated tunable cells. (h) Quantified from fluorescent images, number of metastases in embryos with E-cad$^-$, E-cad$^+$, and tunable E-cad MDA-MB-231 cells (****P <0.0001, *P = 0.02, One Way Anova).

E-cadherin promotes proliferation via the ERK cascade

To understand how E-cad promotes cancer cell proliferation and tumor growth, we conducted global proteomics and phosphoproteomic characterization on the tumors from the MDA-MB-231 bilateral tumor study (Fig. 1b,c). A total of 7766 global proteins and 7258 phosphosites were identified in our report. Hierarchical clustering revealed that distinct groups of proteins were differentially expressed when E-cad expression was manipulated (Fig. 3a). We designed the proteomics analysis to directly compare tumor pairs (i.e. tumors grown in the same mouse). When comparing E-cad$^+$ to E-cad$^-$ tumors, we observed that ~15% of proteins were differentially expressed at the global level and ~8% of phosphosites were differentially occupied by more than 2-fold (Fig. 3b). Utilizing pathway analytical tools[21-25], we consistently found that the MAPKinase and WNT

signaling pathways were the most highly upregulated among cancer-related proliferation pathways (Fig 3c). We also observed that pathways related to protein translation (ribosomal biogenesis) and metabolism were also among the upregulated pathways when comparing E-cad^{+} tumors to E-cad^{-} tumors (FDR ≤ 0.05)[22].

We sought to further explore these proliferation pathways utilizing our 3D double layered spheroid system to conduct a mRNA-level relative expression analysis on E-cad^{+}/E-cad^{-} cells. We observed significant upregulation of ERK1/2 and downstream proteins FOS and ELK1 at the mRNA level in spheroids, which correlated with their upregulation at the protein level in tumors (Fig. S4a,b). This spheroid system also allowed us to conduct a high-throughput drug screening with pharmacological inhibitors targeting various proteins in the MAPK and WNT pathways (Fig. S5c). The inhibitor PD0325901[11], which blocks MEK1/2 from phosphorylating ERK1/2, was the most effective inhibitor in reverting the hyper-proliferative phenotype induced by E-cad (Fig. S4 d-g,j). Indeed, the IC_{50} for PD0325901 was 3.5-fold higher in E-cad^{+} cancer cells when compared to E-cad^{-} cancer cells (Fig. S4h). Our results suggest that E-cad increases the phosphorylation of ERK1/2 by increasing the expression of its kinase MEK1/2 - as increased phosphorylation of ERK1/2 by MEK1/2 leads to the upregulation of downstream transcription factors, including c-FOS and ELK1, which have been directly linked to cellular growth, proliferation, and differentiation.[11,26]

After this validation, we delved further into the global proteomics to determine the connection between E-cad and ERK1/2. First we focused on cell surface receptor expression, observing a 1.4-fold increase in the receptor tyrosine kinase (RTK) encoded by *EGFR* in E-cad^{+} tumors compared to E-cad^{-} tumors (Fig. 3f). While we did not observe

large differences in the expression of Ras and Raf proteins (downstream of EGFR, and upstream of ERK1/2) in the canonical Ras/Raf/MEK/ERK pathway, we observed an increase in the expression of oncogenic Ras protein, RAB38 (Fig. 3f). We also observed a 1.5 fold increase in the expression of c-Myc encoded by *MYC* and a 1.7 fold increase in FOS-related genes *FOSL1* and *FOSL2*, downstream of the ERK1/2 complex (Fig. 3f).

Next, we assessed the activation of these molecules via phosphorylation utilizing phosphosite proteomic analysis. At the cell surface, we observed a 1.8-fold increase in the abundance of EGFR autophosphorylation at site Y1172 in E-cad^{+} tumors when compared to the E-cad^{-} tumors (Fig. 3g). EGFR autophosphorylation can be induced by interactions with E-cad, which we suspect initiates the upregulation of the ERK cascade, leading to the hyper-proliferative phenotype induced by high E-cad expression.[27] We also observed significant increases in the phosphorylation abundance of key molecules in the MAPKinase pathway: Y187 and Y204 on ERK1/2 and S62 on c-MYC (Fig. 3d,e). We further confirmed that the upregulation of ERK1/2 was due increased occupancy of MEK1/2 on sites Y187 and Y204, and the upregulation of downstream molecule c-Myc was due to increased phosphorylation of S62 by ERK1/2 (Fig. 3g).

With this proposed mechanism, we returned to the preserved tissues from the mouse model and performed immunohistochemistry (IHC) staining on the tumors from the bilateral study and lungs from the tail vein model. We observed that phospho-ERK was much more prevalent in E-cad^{+} tumors (Fig. 3i). This correlation was maintained in the lungs from the tail vein experimental metastasis model (Fig. 3j). It is an important observation that the relationship between E-cad and phospho-ERK does not depend on the microenvironment.

We also conducted western blot analysis on the tumors and observed the same trend: E-cad^{+} tumors had more phospho-ERK compared to E-cad^{-} tumors, without significantly impacting the total ERK content (Fig. 3h). This suggested that the increased cell proliferation induced by E-cad expression is not due a change in cell cycle. We further verified by performing a cell cycle distribution analysis (Fig. S4k). In sum, we propose that E-cad induces a hyper-proliferative phenotype via hyper-activation of ERK1/2.

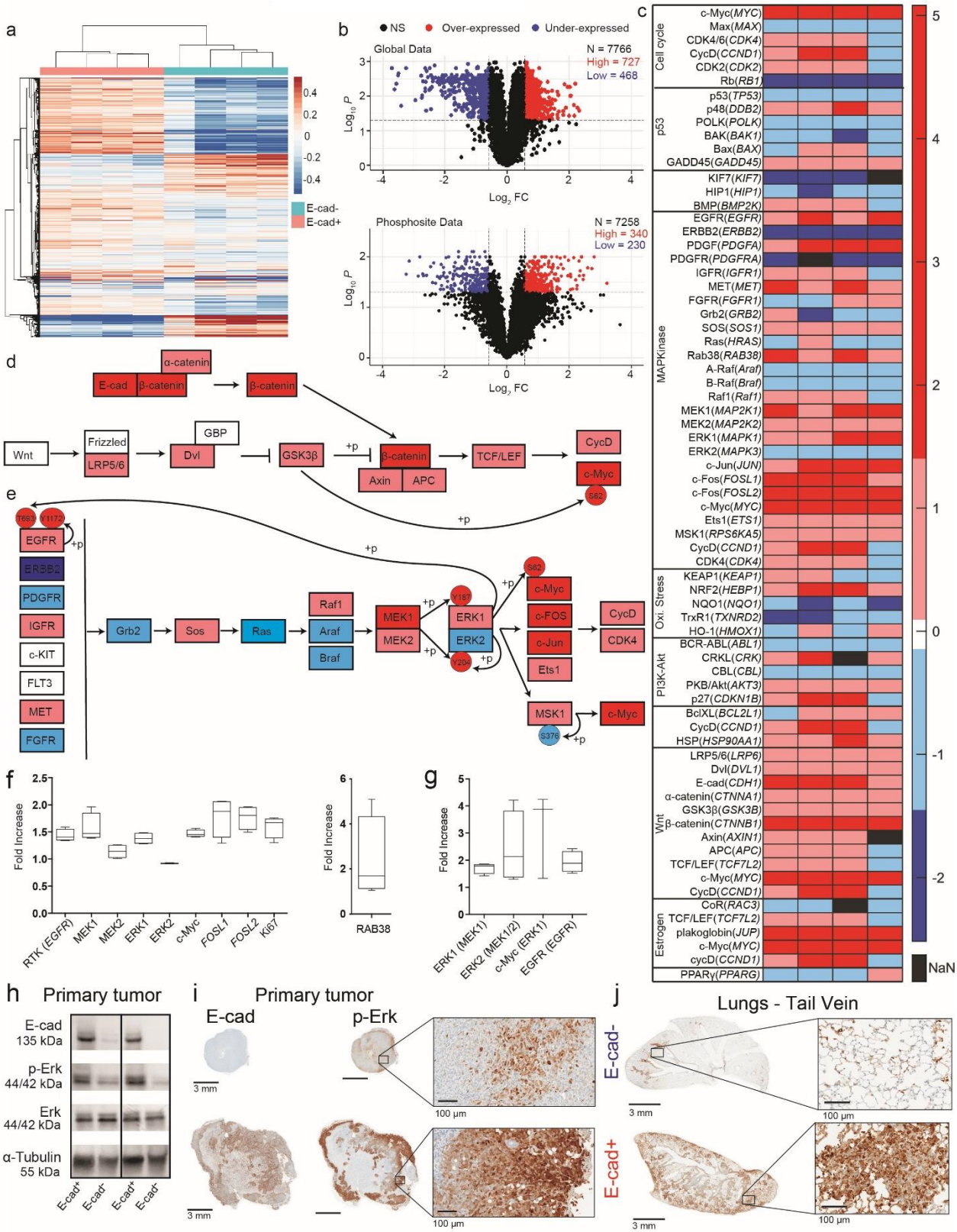
a
E-cad-
E-cad+
b
NS
Over-expressed
Under-expressed
Global Data
N = 7766
High = 727
Low = 468
Phosphosite Data
N = 7258
High = 340
Low = 230
Log10 P
Log2 FC
c
Cell cycle
p53
MAPKinase
Oxi. Stress
PI3K-Akt
Wnt
Estrogen
NaN
d
α-catenin
E-cad
β-catenin
Wnt
Frizzled
LRP5/6
GBP
Dvl
GSK3β
+p
Axin
APC
TCF/LEF
CycD
c-Myc
S62
e
EGFR
ERBB2
PDGFR
IGFR
c-KIT
FLT3
MET
FGFR
Grb2
Sos
Ras
Raf1
Araf
Braf
MEK1
MEK2
ERK1
ERK2
c-FOS
c-Jun
Ets1
CDK4
MSK1
S376
f
Fold Increase
RTK (EGFR)
MEK1
MEK2
ERK1
ERK2
c-Myc
FOSL1
FOSL2
Ki67
RAB38
g
ERK1 (MEK1)
ERK2 (MEK1/2)
c-Myc (ERK1)
EGFR (EGFR)
h
Primary tumor
E-cad
135 kDa
p-Erk
44/42 kDa
Erk
44/42 kDa
α-Tubulin
55 kDa
E-cad+
E-cad-
i
Primary tumor
E-cad
p-Erk
3 mm
100 µm
j
Lungs - Tail Vein
E-cad-
E-cad+
3 mm
100 µm

Figure 3: Proteomics assessment of E-cadherin's effect on key pathways

(a) Clustering heatmap of global proteomics from MDA-MB-231 bilateral tumors. (b) Volcano plot of differential expression (fold change) of E-cad^{+} compared to E-cad^{-} tumor samples. Fold change cut off = 1.5. p-value cut off = 0.05. (c) Heatmap of pathways in cancer (KEGG)[23,24] utilizing the differential expression (fold change) of E-cad^{+} tumor compared to E-cad^{-} tumor pair from MDA-MB-231 bilateral. Color key: dark red = FC ≥ 1.5, light red = FC < 1.5, >0, light blue = FC <0, > -1.5, dark blue = FC ≤ -1.5, and black = NaN (not detected in proteomics data acquisition). (d) WNT pathway schematic of fold change, demonstrating fold changes of proteins and activation (phosphorylation). White boxes represent proteins that were not detected by proteomics. Color key: white boxes = NaN (not detected in proteomics data acquisition), red and blue variations match heatmap in panel (c). (e) MAPKinase pathway schematic of fold change. (f) Box plots showing the fold change of key proteins in the MAPKinase pathway at the global proteomics level. (g) Box plots showing activation differences of key proteins in the MAPKinase pathway (phosphosite proteomics data). (h) Western blot of two sets of primary tumors from the MDA-MB-231 bilateral mouse model, showing more phospho-ERK in E-cad^{+} tumor. (i) IHC of bilateral primary tumors, showing stronger phospho-ERK signal in E-cad^{+} tumors. (j) IHC of lungs from the tail vein model using MDA-MB-231 E-cad^{+} and E-cad^{-}, showing stronger phospho-ERK signal in lungs of mice injected with MDA-MB-231 E-cad^{+} cells when compared to lungs injected with E-cad- cells.

Blocking phosphorylation of ERK reverses E-cad-induced cancer cell hyperproliferation

To validate the above mechanism of action of E-cad, we subjected tumor-bearing mice to a MEK1/2 treatment. Compared to vehicle control, we observed a significant decrease in tumor progression and final tumor weight and size when E-cad^{+} tumors were treated with MEK1/2 inhibitor PD0325901[11] (Fig. 4b,c, Fig. S5a). We also assessed metastatic burden in the lungs using RT-qPCR and observed a 90-fold decrease in relative expression of *HK2* when mice were treated with the MEK1/2 inhibitor. This suggests that the MEK1/2 inhibitor blocked E-cad induced growth at both the primary and secondary sites. To directly observe how the inhibitor affected the phosphorylation of ERK, tumors were harvested on day 13 (one day after the treatment was completed) for IHC staining. Both the phospho-ERK and Ki67 IHC stains confirmed our previous results: when the phosphorylation of ERK is blocked, there are fewer proliferative cells in the tumors (Fig 4f,g).

To ensure that the inhibitor PD0325901 worked specifically and did not significantly impact other pathways, we conducted global and phosphosite proteomics on the treated and untreated tumors. We observed that only a small percentage of proteins and phosphosites were impacted by the MEK1/2 treatment (Fig 4h). We also observed a statistically significant decrease in FOS-related gene *FOSL1* and Ki67 in the treated tumors when compared to non-treated tumors (Fig 4i). The phosphosite data confirmed that the inhibitor prevented phosphorylation of ERK by MEK1/2 at sites Y186 and Y204 (Fig 4j). After assessing the impact of treatment on pathway dynamics, we observed that the inhibitor worked as intended (Fig. 4k). This observation confirms that E-cad induces

hyper-phosphorylation of ERK1/2 by MEK1/2 and blocking this phosphorylation event can reverse this phenotype.

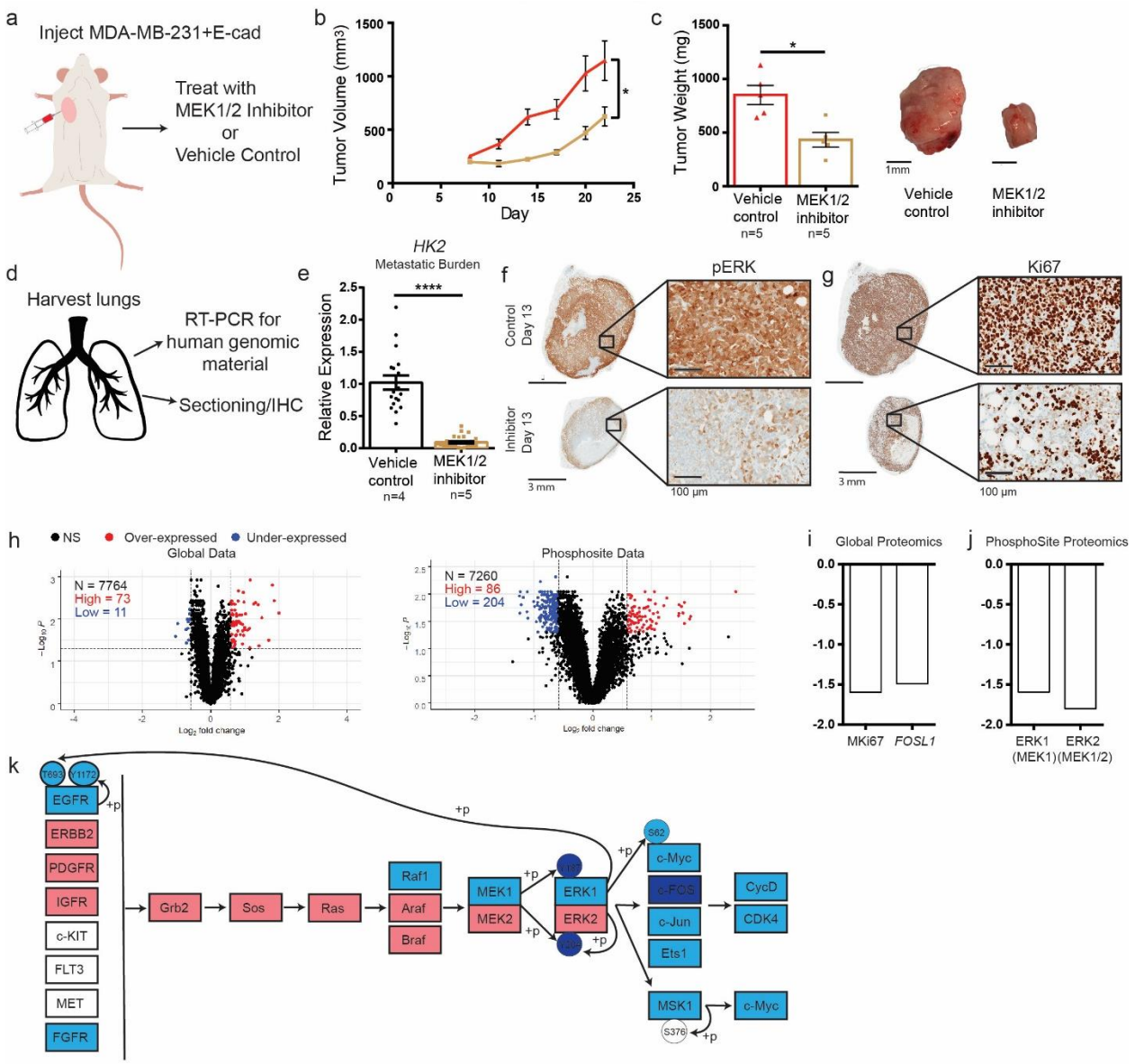

Figure 4: Blocking ERK phosphorylation stops hyperproliferative effect of E-cad. (a) Schematic of orthotopic injection into mammary fat pad of NSG mice to assess impact of MEK1/2 inhibitor on primary tumor progression (N=5 per group). (b) Time-dependent

volume of MDA-MB-231 E-cad^{+} untreated (red line) and treated (gold line). Apparent tumor volume was assessed using calipers to measure x and y dimensions. *P =0.0345 (c) Tumor weight at the end of the study with representative pictures of excised tumors. *P=0.0159. (d) Schematic of metastatic burden assessment utilizing both RT-qPCR and IHC. (e) Relative expression level of *HK2* (marker for human genomic material) in the vehicle control and MEK1/2 inhibitor treated groups. ****P<0.0001. N=2, n=3 per mouse for qPCR. (f) IHC staining of tumors harvested on day 13, 1 day after treatment was completed, showing downregulation of phospho-ERK with treatment of MEK1/2 inhibitor PD0325901. (g) IHC staining of tumors harvested on day 13, 1 day after treatment was completed for Ki67. There is a significant decrease in Ki67 positive cells when mice were treated with MEK1/2 inhibitor PD0325901. (h) Volcano plots showing statistically significant fold increase/decrease during treatment in the global proteomics and phospho-site proteomics. (i) Fold decrease in global protein levels of proliferation marker Ki67 (P= 0.0001) and FOS related protein encoded by FOSL1 (P= 0.0015), (j) Fold decrease in phospho-sites Y187 (P= 0.254) and Y204 (P= 0.053) on ERK1/2 when treated with MEK1/2 inhibitor P0D0325901. (k) Pathway schematic comparing relative expression of treated tumors to control tumors.

Discussion

Our results indicate that E-cad can induce a hyper-proliferative phenotype in breast cancer cells. E-cad expression dramatically increases tumor growth at the primary site and greatly promotes the formation of macrometastasis via cell proliferation *in situ*.

E-cad has been long classified as a tumor suppressor gene due in part to its role in the epithelial-to-mesenchymal transition of cells. These studies, however, have been typically conducted in fibroblasts or epithelial cell lines that in some cases were engineered to induce a tumorigenic phenotype.[1,28] Recent studies have shown that in breast cancer, this classification may not be correct: E-cad may not only be required for metastasis, but also have an impact on cancer cells' survival in the circulatory system.[5,6] Here, we intentionally designed our experiments to focus on tumor and distal site growth. We offer insight into the role of E-cad in the primary tumor and the secondary metastatic site as a promoter of proliferation via the ERK cascade within the greater MAPKinase signaling pathway.

The Ras/Raf/MEK/ERK cascade has been linked to the regulation of many cellular processes, including cell adhesion, survival, and proliferation.[26] This pathway is upregulated in a variety of cancers, even in the absence of oncogenic Ras and Raf mutations, which have a high occurrence in multiple cancer types, including breast, pancreatic, and colorectal cancers.[26] We show in a MDA-MB-231 mouse model the the E-cad induced increase in ERK1/2 phosphorylation resulted in a 7-fold increase in the expression of c-FOS. This transcription factor has previously been linked to the ability of MCF7 breast cancer cells to proliferate.[29] Clinical data corroborates our results: Patients with high E-cad expression have a worse overall survival. Moreover, the vast majority of invasive ductal carcinoma (IDC) are E-cad positive; we show that pateints with IDC tumors displaying relatively high E-cad expression have a worse overall survival (Fig. 5a). There are also implications for other cancer types, as in the ovarian cancer cell line

SKOV3, E-cad mediated MEK/ERK activation was observed *in vitro*, which suggests this phenomenon may not be exclusive to breast cancer.[30]

From our global and phospho-proteomics pathway schematics, we propose that E-cad promotes EGFR activity, due to increases in both EGFR global protein expression levels and EGFR autophosphorylation. This is supported by recent studies which show that ligated E-cad can recruit EGFR, induce its ligand-independent activation, and result in activation of the MAPK pathway.[27,31] Thus explaining the direct link between E-cad and the Ras/Raf/MEK/ERK cascade (Fig 5b). This pathway has implications for patients with high E-cad expressing tumors, affecting both primary and secondary sites, regardless of the potential impacts E-cad has on metastasis (Fig 5c).

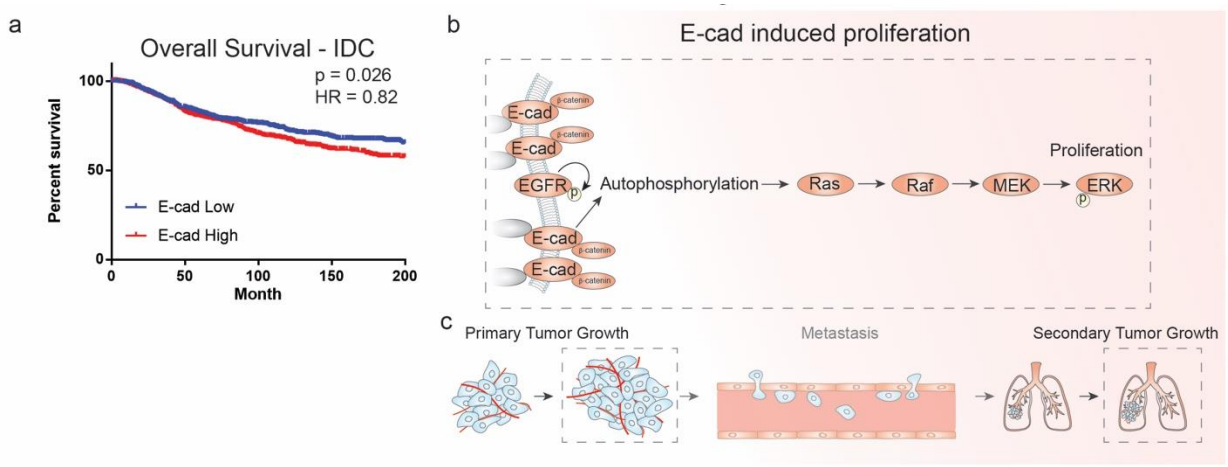

Figure 5: Summary figure. (a) Overall survival plot based on Metabric data set[32] for IDC breast cancer. (b) Schematic of proposed mechanism. (c) Cartoon of metastatic cascade, highlighting the areas of interest.

References

1. Polyak, K. & Weinberg, R. A. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nature Reviews Cancer* **9**, 265 (2009).

2. Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* **139**, 871-890.

3. Stockinger, A. et al. E-cadherin regulates cell growth by modulating proliferation-dependent b-catenin transcriptional activity. *J. Cell Biol.* **154**, 1185-1196 (2001).

4. Soler, C. et al. Dermal fibroblast proliferation is improved by beta-catenin over-expression and inhibited by E-cadherin expression. *FEBS Lett.* **442**, 178-182 (1999).

5. Padmanaban, V., Krol, I., Suhail, Y. *et al.* E-cadherin is required for metastasis in multiple models of breast cancer. *Nature* **573,** 439–444 (2019).

6. Na TY, Schecterson L, Mendonsa AM, Gumbiner BM. The functional activity of E-cadherin controls tumor cell metastasis at multiple steps. *Proc. Natl. Acad. Sci. U S A*. **117**, 5931–5937 (2020).

7. Rodriguez, F.J., Lewis-Tuffin, L.J., and Anastasiadis, P.Z. E-cadherin's dark side: possible role in tumor progression. *Biochim. Biophys. Acta.* **1826**, 23-31 (2012).

8. Chu, K., Boley, K. M., Moraes, R., Barsky, S. H. & Robertson, F. M. The Paradox of E-Cadherin: Role in response to hypoxia in the tumor microenvironment and regulation of energy metabolism. *Oncotarget* **4**, 446-462 (2013).

9. Soncin, F. and Ward, C.M. The function of E-cadherin in stem cell pluripotency and self-renewal. *Genes*. **2**, 229-259 (2011).

10. Redmer, T. et al. E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *EMBO Rep.* **12**, 720-726 (2011).

11. Barrett, S.D., et al. The discovery of benzhydroxamate MEK inhibitors CI-1040 and PD032590. *Bioorganic & Medicinal Chemistry Letters*. **18**, 6501-6504(2008).

12. Liu, K., Newbury, P.A., Glicksberg, B.S. *et al.* Evaluating cell lines as models for metastatic breast cancer through integrative analysis of genomic data. *Nat Commun* **10,** 2138 (2019).

13. Wu. P.H., et al. Single cell morphology encodes metastatic potential. *Science Advances*. **12**, EAAW6938 (2020).

14. Van Roy, F. & Berx, G. The cell-cell adhesion molecule E-cadherin. Cellular and Molecular *Life Sciences* **65**, 3756-3788 (2008).

15. Iorns, E., et al. A new mouse model for the study of human breast cancer metastasis. *PLoS One*. **7**, e47995 (2012).

16. Chaturvedi P, Gilkes DM, Wong CCL, et al. Hypoxia-inducible factor–dependent breast cancer–mesenchymal stem cell bidirectional signaling promotes metastasis. J Clin Invest. 2013;123(3):1402. doi:10.1172/JCI69244

17. Kim Y, Williams KC, Gavin CT, Jardine E, Chambers AF, Leong HS. Quantification of cancer cell extravasation in vivo. Nat Protoc. 2016;11(5):937-948. doi:10.1038/nprot.2016.050

18. Williams KC, Cepeda MA, Javed S, et al. Invadopodia are chemosensing protrusions that guide cancer cell extravasation to promote brain tropism in metastasis. Oncogene. 2019;38(19):3598-3615. doi:10.1038/s41388-018-0667-4

19. Leong HS, Robertson AE, Stoletov K, et al. Invadopodia are required for cancer cell extravasation and are a therapeutic target for metastasis. Cell Rep. 2014;8(5):1558-1570. doi:10.1016/j.celrep.2014.07.050

20. Leong HS, Lizardo MM, Ablack A, et al. Imaging the impact of chemically inducible proteins on cellular dynamics in vivo. PLoS One. 2012;7(1):e30177. doi:10.1371/journal.pone.0030177

21. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources. Nature Protoc. 2009;4(1):44-57.

22. Jing Wang, Suhas Vasaikar, Zhiao Shi, Michael Greer, Bing Zhang, WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit, Nucleic Acids Research, Volume 45, Issue W1, 3 July 2017, Pages W130–W137, https://doi.org/10.1093/nar/gkx356

23. Kanehisa, M. and Goto, S.; KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27-30 (2000).

24. Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., and Tanabe, M.; New approach for understanding genome variations in KEGG. Nucleic Acids Res. 47, D590-D595 (2019).

25. Kanehisa, M; Toward understanding the origin and evolution of cellular organisms. Protein Sci. 28, 1947-1951 (2019)

26. Dhillon, A. S., Hagan, S., Rath, O. & Kolch, W. MAP kinase signalling pathways in cancer. *Oncogene.* **26**, 3279 (2007).

27. Fedor-Chaiken, Mary & Hein, Patrick & Stewart, Jane & Brackenbury, Robert & Kinch, Michael. (2003). E-Cadherin Binding Modulates EGF Receptor Activation. Cell communication & adhesion. 10. 105-18. 10.1080/714040416.

28. Onder, T. T. et al. Loss of E-Cadherin Promotes Metastasis via Multiple Downstream Transcriptional Pathways. *Cancer Research.* **68**, 3645-3654(2008).

29. Lu, C. et al. cFos is critical for MCF7 breast cancer cell growth. *Oncogene*. **24**, 6516–6524(2005).

30. Dong, L. et al. E-cadherin promotes proliferation of human ovarian cancer cells *in vitro* via activating MEK/ERK pathway. *Acta Pharma. Sinica*. **33**, 817-822(2012).

31. Cavallaro, U., Christofori, G. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. Nat Rev Cancer 4, 118–132 (2004). https://doi.org/10.1038/nrc1276

32. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, Pugh M, Jones L, Russell R, Sammut SJ, Tsui DW, Liu B, Dawson SJ, Abraham J, Northen H, Peden JF, Mukherjee A, Turashvili G, Green AR, McKinney S, Oloumi A, Shah S, Rosenfeld N, Murphy L, Bentley DR, Ellis IO, Purushotham A, Pinder SE, Børresen-Dale AL, Earl HM, Pharoah PD, Ross MT, Aparicio S, Caldas C. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun. 2016 May 10;7:11479. doi: 10.1038/ncomms11479. PMID: 27161491; PMCID: PMC4866047.

Tables

Table 1: List of primers used to assess MAPK and WNT signaling pathways in 3D double layered spheroid system.

AKT1 FWD	AGCGACGTGGCTATTGTGAAG
AKT1 RVS	GCCATCATTCTTGAGGAGGAAGT
ATF2 FWD	AATTGAGGAGCCTTCTGTTGTAG
ATF2 RVS	CATCACTGGTAGTAGACTCTGGG
DAXX FWD	GATACCTTCCCTGACTATGGGG
DAXX RVS	GTAACCTGATGCCCACATCTC
ELK1 FWD	CTGCCTCCTAGCATTCACTTC
ELK1 RVS	GCTGCCACTGGATGGAAACT
FOS FWD	CCGGGGATAGCCTCTCTTACT
FOS RVS	CCAGGTCCGTGCAGAAGTC
GREB1 FWD	ATGGGAAATTCTTACGCTGGAC
GREB1 RVS	CACTCGGCTACCACCTTCT
JUN FWD	TCCAAGTGCCGAAAAAGGAAG
JUN RVS	CGAGTTCTGAGCTTTCAAGGT
JUND FWD	TCATCATCCAGTCCAACGGG
JUND RVS	TTCTGCTTGTGTAAATCCTCCAG
MAP2K1 FWD	CAATGGCGGTGTGGTGTTC
MAP2K1 RVS	GATTGCGGGTTTGATCTCCAG
MAP2K2 FWD	AGGTCCTGCACGAATGCAA
MAP2K2 RVS	CGTCCATGTGTTCCATGCAA
MAP2K3 FWD	GAGGGAGACGTGTGGATCTG
MAP2K3 RVS	CCGCACGATAGACACAGCAAT
MAP2K4 FWD	TGCAGGGTAAACGCAAAGCA
MAP2K4 RVS	CTCCTGTAGGATTGGGATTCAGA
MAP2K5 FWD	CCGGACCCTCTCAACACAG
MAP2K5 RVS	ATGACCAAGAGTGTCCCGATA
MAP2K6 FWD	AAACGGCTACTGATGGATTTGG
MAP2K6 RVS	CAGTGCGCCATAAAAGGTGAC
MAPK1 FWD	TCACACAGGGTTCCTGACAGA
MAPK1 RVS	ATGCAGCCTACAGACCAAATATC
MAPK10 FWD	CAGATGGAATTAGACCATGAGCG
MAPK10 RVS	TCAATGTGCAATCAGACTTGACT
MAPK12 FWD	AGTGGCTTTTACCGCCAGG
MAPK12 RVS	GACTGGAAGGGCCGATACAG
MAPK13 FWD	TGAGCCGACCCTTTCAGTC
MAPK13 RVS	AGCCCAATGACGTTCTCATGC

MAPK14 FWD	TCAGTCCATCATTCATGCGAAA
MAPK14 RVS	AACGTCCAACAGACCAATCAC
MAPK3 FWD	CTACACGCAGTTGCAGTACAT
MAPK3 RVS	CAGCAGGATCTGGATCTCCC
MAPK8 FWD	TCTGGTATGATCCTTCTGAAGCA
MAPK8 RVS	TCCTCCAAGTCCATAACTTCCTT
MAPK9 FWD	GAAACTAAGCCGTCCTTTTCAGA
MAPK9 RVS	TCCAGCTCCATGTGAATAACCT
MAX FWD	CAATCTGCGGCTGACAAACG
MAX RVS	GCACTTGACCTCGCCTTCT
TRAF2 FWD	CCTTCCCAGATAATGCTGCCC
TRAF2 RVS	GCTCTCGTATTCTTTCAGGGTC
APC FWD	AAGCATGAAACCGGCTCACAT
APC RVS	CATTCGTGTAGTTGAACCCTGA
AXIN1 FWD	GGTTTCCCCTTGGACCTCG
AXIN1 RVS	CCGTCGAAGTCTCACCTTTAATG
BTRC FWD	TGCCCAAGCAACGGAAACT
BTRC RVS	GCCCATGTTGGTAATGACACA
CCND1 FWD	GCTGCGAAGTGGAAACCATC
CCND1 RVS	CCTCCTTCTGCACACATTTGAA
CSNK1A1 FWD	AGTGGCAGTGAAGCTAGAATCT
CSNK1A1 RVS	CGCCCAATACCCATTAGGAAGTT
CSNK2A1 FWD	CGAGTTGCTTCCCGATACTTC
CSNK2A1 RVS	ACTTGCCAGCATACAACCCAA
CTBP1 FWD	CGACCCTTACTTGTCGGATGG
CTBP1 RVS	TTGACGGTGAAGTCGTTGATG
CTNNB1 FWD	AGCTTCCAGACACGCTATCAT
CTNNB1 RVS	CGGTACAACGAGCTGTTTCTAC
DVL1 FWD	GAGGGTGCTCACTCGGATG
DVL1 RVS	GTGCCTGTCTCGTTGTCCA
FZD1 FWD	ATCTTCTTGTCCGGCTGTTACA
FZD1 RVS	GTCCTCGGCGAACTTGTCATT
GSK3B FWD	GGCAGCATGAAAGTTAGCAGA
GSK3B RVS	GGCGACCAGTTCTCCTGAATC
HDAC1 FWD	CTACTACGACGGGGATGTTGG
HDAC1 RVS	GAGTCATGCGGATTCGGTGAG
LEF1 FWD	AGAACACCCCGATGACGGA
LEF1 RVS	GGCATCATTATGTACCCGGAAT
NLK FWD	CCAACCTCCACACATTGACTATT
NLK RVS	ACTTTGACATGATCTGAGCTGAG
PPARD FWD	GCCTCTATCGTCAACAAGGAC

PPARD RVS	GCAATGAATAGGGCCAGGTC
PPP2CA FWD	GTTCGTTACCGTGAACGCATC
PPP2CA RVS	TGGCGAGAGACCACCATGT
WIF1 FWD	TCTCCAAACACCTCAAAATGCT
WIF1 RVS	GACACTCGCAGATGCGTCT
WNT1 FWD	CGATGGTGGGGTATTGTGAAC
WNT1 RVS	CCGGATTTTGGCGTATCAGAC
WNT11 FWD	GGAGTCGGCCTTCGTGTATG
WNT11 RVS	GCCCGTAGCTGAGGTTGTC
WNT5A FWD	ATTCTTGGTGGTCGCTAGGTA
WNT5A RVS	CGCCTTCTCCGATGTACTGC

Table 2: List of primers used to assess metastatic burden (*HK2*).

18s Human FWD	AGAAGTGACGCAGCCCTCTA
18s Human RVS	GAGGATGAGGTGGAACGTGT
18s Mouse FWD	CGGCGACGACCCATTCGAAC
18s Mouse RVS	GAATCGAACCCTGATTCCCCGT
aTubulin Human FWD	AGGAGTCCAGATCGGCAATG
aTubulin Human RVS	GTCCCCACCACCAATGGTTT
aTubulin Mouse FWD	CACACAAGCTCACTCACCCT
aTubulin Mouse RVS	CTGTTATTAGGGATGTGACTCCA
HK2 (Human) FWD	CCAGTTCATTCACATCATCAG
HK2 (Human) RVS	CTTACACGAGGTCACATAGC
GAPDH Human FWD	GGAGCGAGATCCCTCCAAAAT
GAPDH Human RVS	GGCTGTTGTCATACTTCTCATGG
GAPDH Mouse FWD	TCACCACCATGGAGAAGGC

Materials and Methods

Cell culture

Human breast carcinoma cells MDA-MB-231, MCF7, MDA-MB-468, T47D, HCC1954, ZR751 (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM, Corning, 10-013-CV) supplemented with 10% (v/v) fetal bovine serum (FBS, Corning, 35-010-CV) and 1% Penicillin-Streptomycin (Gibco, 15140-122). Cells were maintained at 37°C and 5% CO_2 in a humidified incubator during cell culture.

Lentivirus production and transduction

HEK-293T cells (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM, Sigma Aldrich) supplemented with 10% (v/v) fetal bovine serum (FBS, Corning) and 1% Penicillin-Streptomycin (Corning). Mission shRNA glycerol stock for CDH1 (NM004360, TRCN0000237841) was obtained from Sigma (Bacterial culture was started and incubated for 16 hrs (overnight) on a shaker at 37 C and a speed of 200 RPM. Plasmid DNA was extracted using Midi-Prep Kit from (Machery-Nagel, 740412.50). Calcium chloride and HEPES-buffered-saline were used for transfection/lentivirus production. Plasmid DNA and reagents were added to HEK-293T cells and virus was harvested every 24 hours for 3 days. E-cadherin positive cells (MCF7, MDA-MB-468, T47D, HCC1954, HCC1937, and ZR751) received two rounds of virus for a total period of 48 h. After 48 h, cells were given fresh DMEM. After 24-48 h, cells were given DMEM with 5µg/mL puromycin for selection. Knockdown cells were cultured in puromycin conditioned medium for three days before checking protein expression. Cells were kept in selection medium for the duration of the experiments.

Spheroid modeling

A double-layered spheroid model was generated with matrigel (Corning) and rat tail type I (Corning). Cells were trypsinized, pelleted, and re-suspended in matrigel at the desired cell density and then into a collagen I gel was mixed as previously described by Fraley *et al*.[1] The spheroids are plated in a 96-well round-bottom dish and suspended in 200 µL medium.

Quantifying double-layered spheroid volume

Inner and outer layers of the spheroids were manually traced using NIS Elements software to generate a radius measurement. Volume of inner and total spheroid was calculated using the volume for a sphere: $V = \frac{4}{3}\pi r^3$. To calculate the volume of the collagen layer, the inner volume was subtracted from the total volume.

2D and 3D Proliferation using PrestoBlue viability assay

Cells were seeded in 96 well plate at 2,000 cells/well for MDA-MB-231 and MCF7 and at 5,000 cells/well for MDA-MB-468 (day 0). On day 1 and day 7, cells were incubated in 1x PrestoBlue (ThermoFisher) at 37ºC for 1 hr. To account for background signal, 5-10 wells of 1x PrestoBlue were prepared and measured to subtract off background medium signal from the experimental wells. After reading the signal, cells were washed with DPBS 3x, given 100 µL of fresh medium.

Spheroids were incubated at 37°C in a 1x PrestoBlue for 3 hours on day 1 (day of generation). Standard curves were generated relating cell number to RFU to ensure

linear relationship (Fig. S1 e-g). After reading, spheroids were wash in DBPS 5x and given 200 µL of fresh medium. Experiment was carried out for either 5 or 8 days. To account for background signal, 5-10 wells of 1x PrestoBlue were prepared and measured to subtract off background signal from the experimental wells. SpectraMax plate reader (Molecular Devices) was used to read the RFU for all experiments at (EX 540, EM 600, cutoff 590).

mRNA extraction and cDNA synthesis

Spheroids were generated and then incubated in fresh media. After 24 hrs, spheroids were harvested and washed 3x in DPBS (Corning). To isolate mRNA, TriZol reagent (Invitrogen) was added to the spheroids and then vortexed until completely dissolved. Direct-zol Mini Prep kit (Zymo research) reagents and protocol were used to complete RNA extraction. cDNA was synthesized using the Bio-Rad cDNA iScript kit following their recommended protocol to make 1000 ng of cDNA for a total reaction volume of 20 µL.

RT-qPCR

qPCR was conducted with iTaq-SYBR Green (Bio-Rad) using the Bio-Rad CFX CFX384 Touch Real-Time PCR detection system. Primers were obtained from Integrated DNA Technologies.

Inhibitor studies

Inhibitor studies were conducted using spheroids described above. Experiments were carried out for a week and imaged daily using phase contrast microscopy. Inhibitors were

obtained from Sellechem and fresh or conditioned medium was added on the day the spheroids were generated and not changed during the week-long experiment.

Protein extraction, concentration, and western blot assays

A clear sample buffer containing SDS with protease inhibitor (Roche, 05892970001) was used to lyse 10-15 spheroids. Samples were sonicated to manually break up the gel layers and dissolve the collagen and Matrigel of the spheroids. Samples were then heated at 100°C for 5 min to completely denature the protein. Protein samples were stored in -80°C freezer until ready for use. Protein concentration was measured using MicroBCA Kit (ThermoFisher, 23235) following the manufacturer's protocol. A loading buffer made from 2x Laemlli (Biorad) was prepared with the instructed amount of beta-mercaptoethanol and added to the protein samples for a final concentration of 2 µg/µL. Bio-rad mini-protean gels (4-15% gradient, 4561086) were used for gel electrophoresis. Transblot Turbo packs (Bio-rad) were used to transfer the protein from the gel to PVDF membrane. Antibodies were purchased from Cell Signaling Technology (E-cadherin 14472 and 3195S, phospho-ERK 9101S, total-ERK 4696S, α-tubulin 2125S, and GAPDH 97166S) and recommended dilutions for the primary and secondary antibodies were used in incubations. Blots were imaged using a Bio-Rad imager and bands were quantified using ImageLab software (BioRad).

Animal Studies

All mouse experiments were carried out according to protocols approved by the Johns Hopkins University Animal Care and Use Committee in accordance with the NIH Guide

for the Care and Use of Laboratory Animals. All mice were housed at a temperature of 25°C under a 12 h dark/light cycle. Primary tumors were established with cell lines obtained from ATCC. Cells were cultured in 5% CO_2 and 37°C. Cells were rinsed with PBS 3 times, pelleted, and resuspended in 1:1 DPBS/Matrigel (Corning). 5-week-old NSG (NOD SCID Gamma) female mice were obtained. Hair was removed from mice to ensure accurate orthotopic injection into the second mammary fat pad. Mice were briefly anesthetized using isoflurane to immobilize the animal for injection of tumor cells. For the bilateral studies, E-cadherin positive cell lines were implanted on the right side of the mouse and E-cadherin negative cell lines were implanted on the left side of the mouse.

For inhibitor studies, MDA-MB-231+E-cadherin primary tumors were established in the second mammary fat pad on the left side. The MEK1/2 inhibitor PD0325901 (Selleckchem) was orally administered via peanut butter pellets daily for 5 days starting one week post tumor cell injection at a dosage of 20 mg/kg and ending on day 12. Control group was given vehicle control (DMSO).

Tumors were measured in two dimensions (x and y) with calipers every three days and weights of animals were recorded. Tumors were categorized as spheres or ellipses by calculating the difference between x and y. If $x-y < 1$mm, the tumor volume was calculated using the volume formula for a sphere. If $x-y > 1$mm, the tumor volume was calculated using the volume formula for an ellipse.

For tail vein “forced metastasis” model, mice were placed under heat lamp for 5-7 min to dilate the tail veins. 100 µL of either E-cad$^+$ or E-cad$^-$ cell solution (5,000 cells/µL) was injected into the lateral tail vein using a 27-gauge needle. Mice were monitored

every three days and sacrificed one week or four weeks post injection using isoflurane and cervical dislocation. Lungs were harvested and then inflated with 1.5% agarose.

The mice were sacrificed using isoflurane and cervical dislocation. For primary tumor models, tumors were excised, weighed and saved. For non-bilateral models, lungs and livers were excised and saved. All IHC staining was performed by an internal core facility.

Processing and analysis of harvested mouse tissues

Portions of primary tumors, lungs and liver were snap frozen in liquid nitrogen in addition to sections that were fixed in formalin were sent for sectioning and staining. H&E slides were obtained for visualization of the tumor microenvironment and regular slides were obtained for immunohistochemistry staining. Protein was extracted from the primary tumor samples using 2X SDS sample buffer. Tumor tissue was broken down using tissue grinder, samples were then heated at 100°C and sonicated. DNA was extracted from tumor tissue using DNA all prep kit (Qiagen).

2D fix and stain

Cells were seeded at desired density in 96 well flat-bottom plate. After 24 h growth, 4% paraformaldehyde (Sigma) was added to each well to fix cells and removed after 20 min incubation. After washes with DPBS, 0.1% Triton (Sigma) was added to cells to permeabilize the cell membrane and incubated for 10 min at room temperature before removal. Wells were then blocked with 1% Bovine Serum Albumin (Sigma) for 30 min.

Primary antibodies were added and samples were incubated overnight in a 2-8°C fridge. After removal of the primary antibodies, samples were stained with H33342 (nuclear DNA), phalloidin (actin), and secondary antibodies, and incubated for 1 h at room temperature. Samples were then washed 3x with DBPS (Corning) before imaging. Images were analyzed using NIS Elements software and ImageJ.

Generation of Tunable E-cadherin-zsGreen Expressing Cell Lines

A DNA plasmid encoding the chemically tunable form of E-cadherin-zsGreen was stably transfected into MDA-MB-231 triple negative breast cancer cells as previously described in.[2]

***Ex ovo* chick embryo chorioallantoic membrane (CAM) platform**

Fertilized white leghorn eggs were used to generate *ex ovo* chick embryos for extravasation efficiency assays[3,4] and experimental metastasis assays[5]. In summary, freshly fertilized eggs (Day 0) were incubated in a Sportsman Rocker Incubator at 37°C and a relative humidity >45% for 4 days. Egg contents were removed from the egg shell as described in[3,6] and placed into individual weigh boats on Day 4. Each *ex ovo* embryo was returned to the incubator until Day 9 of embryonic development.

Experimental Metastasis Assay in the CAM of Chick Embryos

Each cell line was harvested and diluted into D-PBS (Gibco Inc.) at a concentration of 2×10^5 cells/mL. Using Day 9 chick embryos, 2×10^4 cells (100 µL delivery volume) were intravenously (I.V.) injected into a small microvessel (20 µm diameter) to minimize

bleeding post-IV injection. Equal distribution of cancer cells arrested throughout the CAM was confirmed using a fluorescence stereoscope (Nikon Inc.). Injected CAMs were returned to the incubator. At Day 18, each CAM was enumerated for the number of metastases (single metastasis = >5 cells/colony) using a fluorescence stereoscope. For the MDA-MB-231 cell line expressing the E-cadherin-zsGreen-DD protein (chemically tunable E-cadherin with Shield-1 ligand), 50 µL of Shield-1 (5 µM, 5% EtOH in D-PBS, Takeda Inc.) was I.V. injected into each embryo on Days 12, 14, and 16 as described by Leong *et al*.[2]

Extravasation Efficiency Assay in the CAM of Chick Embryos

Each cell line was harvested and diluted into D-PBS (Gibco Inc.) at a concentration of 1×10^6 cells/mL. Using Day 13 chick embryos, 1×10^5 cells (100 µL delivery volume) were intravenously (I.V.) injected into a small microvessel (20 µm diameter) to minimize bleeding post-IV injection. At t=0 after I.V. injection of cancer cells, individual cancer cells within a 1.5×1.5 cm field of view (as described in Leong *et al*[3]) were enumerated with a fluorescence stereomicroscope and then returned to an incubator. At t=24 hours post-I.V. injection, l*ens culinaris* agglutinin-Rhodamine (Vector Laboratories Inc.) was I.V. injected (20 µg/embryo) to visualize the CAM microvasculature and vessel lumen. After agglutinin injection, the same field of view on each embryo was enumerated for extravasated cancer cells using an upright confocal microscope (Fast A1R, Nikon Inc.). Only cells that had extravasated past the CAM, i.e., not overlapping with rhodamine signal, were enumerated.

Proteomics Sample Preparation

Sample preparation for proteomic and phosphoproteomic characterization were carried out as previously described.[7,8] In brief, tumors were excised and immediately snapfrozen in liquid nitrogen to ensure maximum preservation of phosphorylated proteins. Frozen tumor tissue was cryopulverized and resuspended in lysis buffer (8M urea, 75 mM NaCl, 50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 2 µg/mL aprotinin, 10 µg/mL leupeptin, 1 mM PMSF, 10 mM NaF, Phosphatase Inhibitor Cocktail 2 and Phosphatase Inhibitor Cocktail 3 (1:100 dilution), and 20 µM PUGNAc). To remove any particulates, lysates were centrifuged at 20,000 x g for 10 min at 4°C. Protein lysates were subjected to reduction with 5mM, 1,4-Dithiothreitol for 30 min at RT. Then were subjected to alkylation with 10mM iodoacetamide for 45 min at RT, away from light. 50 mM Tris-HCl (pH 8.0) was used to reduce urea concentration to < 2M. Samples were subjected to digestion of Lys-C (Wako Chemicals) (1:50) for 2 h at RT and then to trypsin (Promega) (1:50) overnight at RT. Generated peptides were then acidified to final concentration of 1% formic acid, subjected to clean-up using C-18 SepPak columns (Waters), and then dried utilizing a SpeedVac (Thermo Scientific). Desalted peptides were then labeled with 10-plex Tandem-mass-tag (TMT) reagents following manufacturer's instructions (Thermo Scientific). The samples were then reconstituted in volume of 20 mM ammonium formate (pH 10) and 2% acetonitrile (ACN) and subjected to basic reverse-phase chromatography with Solvent A (2% ACN, 5 mM ammonium formate, pH 10) and non-linear gradient of Solvent B (90% ACN, 5mN ammonium formatem pH 10) at 1 mL/min in the following sequence: 0% Solvent B for 9 min, 6% solvent B for 4 min, 6% increase to 28.5% solvent B over 50 min, 28% increase to 34% solvent B over 5.5 min,

34% increase to 60% solvent B over 13 min, ending with hold at 60% solvent B for 8.5 min – conducted on Agilent 4.6 mm x 250 mm RP Zorbax 300 A Extend-C18 column with 3.5 µm beads (Agilent). Collected fractions were concatenated and 5% of fractions were alqiuopted for global proteomic analysis. Samples for global proteomic analysis were then dried down and resuspended in 3% ACN and 0.1% formic acid before ESI-LC-MS/MS analysis. The remaining 95% of sample was utilized for phosphopeptide enrichment and further concatenated using immobilized metal affinity chromatrography (IMAC). Ni-NTA agarose beads were used to prepare Fe^{3+}-NTA agarose beads, and then 300 µg of peptides reconstituted in 80% ACN/0.1% trifluoroacetic acid were incubated with 10 µL of Fe^{3+}-IMAC beads for 30 min. Samples were then centrifuged and supernatant was removed. Beads were washed 2x and then loaded onto equilibrated C-18 Stage Tips with 80% ACN, 0.1% trifluoroacetic acid. Tips were rinsed 2x with 1% formic acid. Sample was then eluted (3x) off the Fe^{3+}-IMAC beads and onto C-18 Stage tips with 70 µL of 500 mM dibasic potassium phosphate (pH 7.0). C-18 Stage Tips were washed 2x with 1% formic acid, followed by elution (2x) of phosphopeptides from C-18 stage tips with 50% ACN and 0.1% formic acid.

Following elution, samples were dried and resuspended in 3% ACN, 0.1% formic acid prior to ESI-LC-MS/MS analysis. 1 µg of peptide was separated using Easy nLC 1200 UHPLC system on an in-house packed 20 cm x 75 µm diameter C18 column (1.9 µm Reprosil-Pur C18-AQ beads (Dr. Maisch GmbH); Picofrit 10 µm opening (New Objective)). Column parameters: 50°C, 0.300 µL/min with 0.1% formic acid, 2% acetonitrile in water and 0.1% formic acid, 90% acetonitrile. Using a 6 to 30% gradient

over 84 mins peptides were separated and analyzed using Thermo Fusion Lumos mass spectrometer (Thermo Scientific). Following parameters were used- MS1: 60,000 for resolution, 350 to 1800 m/z for mass range, 30% for RF Lens, $4.0e^5$ for AGC Target, 50 ms for Max IT, 2 to 6 for charge state include, 45 s for dynamic exclusion, top 20 ions selected for MS2. For MS2: 50,000 for resolution, 37 for high-energy collision dissociation activation energy (HCD), 0.7 for isolation width (m/z), $2.0e^5$ for AGC Target, and 105 ms for Max-IT.

Proteomics Data Analysis

All LC-MS/MS files were analyzed by MS-PyCloud, a cloud-based proteomic pipeline developed at Johns Hopkins University. MS-PyCloud was used to perform database search for spectrum assignments[9], using MS-GF+ against a combined human and mouse UniprotKB Swiss-Prot database[10,11]. We focused on human-derived proteins for downstream analyses. False discovery rate at PSM, peptide, and protein levels were obtained using a decoy database[12]. Peptides were searched with two trypic ends, allowing for up to two missed cleavages. The following search parameters were used: 20 ppm precursor tolerance and 0.06 Da fragment ion tolerance, static modification of carbamidomethylation at cysteine (+57.02146), TMT-labeled modification of N-terminus and lysine (+229.1629) and variable modifications of oxidation of methionine (+15.99491), phosphorylation at serine, threonine, tyrosine (+79.96633). The following filters were used for global data analysis: one PSM per peptide, two peptides per protein, 1% FDR at protein level. The following filters were used for phosphoproteome data: one PSM per peptide, one peptide per protein, 1% FDR threshold at peptide level.

Phosphopeptide-level data was used to examine overall relationship between substrates and kinases. Kinase-substrate associations wre extracted from PhosphoSitePlus (Hornbeck 2015) to eliminate phosphopeptides containing phosphosites that were not reported as well as those without associated kinases identified in our global dataset. We then calculated the fold changed (log_2 scale) and then ranked each tumor (greater than 1.5 fold increase).[7,8]

Generation of heatmaps and volcano plots for proteomics data

Normalized log_2 transformed protein expression values were used to generate heatmaps. Unsupervised hierarchical clustering was used to group samples (top dendrogram) and proteins (side dendrogram) using Euclidean distances. Heatmaps were generated in R, using the pheatmap package[13]. Normalized log_2 transformed protein expression values were used to generate volcano plots. Log_2 fold changes were calculated between groups of samples. A two-tailed t-test was used to generate P-values. P-values were adjusted via a Benjamini-Hochberg procedure to correct for multiple comparisons. Statistically significant differential expression was determined to be greater than 1.5-fold change and an adjusted P-value < 0.05. Volcano plots were generated in R using the EnhancedVolcano package[14].

Statistical Analysis

Figure legends describe statistical tests used for individual experiments and also include p values.

References for Methods

1. Fraley, S. I. et al. Three-dimensional matrix fiber alignment modulates cell migration and MT1-MMP utility by spatially and temporally directing protrusions. Scientific Reports 5, 14580, doi:10.1038/srep14580 (2015).
2. Leong HS, Lizardo MM, Ablack A, et al. Imaging the impact of chemically inducible proteins on cellular dynamics in vivo. PLoS One. 2012;7(1):e30177. doi:10.1371/journal.pone.0030177
3. Kim Y, Williams KC, Gavin CT, Jardine E, Chambers AF, Leong HS. Quantification of cancer cell extravasation in vivo. Nat Protoc. 2016;11(5):937-948. doi:10.1038/nprot.2016.050
4. Williams, K.C., Cepeda, M.A., Javed, S. et al. Invadopodia are chemosensing protrusions that guide cancer cell extravasation to promote brain tropism in metastasis. Oncogene 38, 3598–3615 (2019). https://doi.org/10.1038/s41388-018-0667-4
5. Leong H.S., Chambers A.F., Lewis J.D. (2012) Assessing Cancer Cell Migration and Metastatic Growth In Vivo in the Chick Embryo Using Fluorescence Intravital Imaging. In: Hoffman R. (eds) In Vivo Cellular Imaging Using Fluorescent Proteins. Methods in Molecular Biology (Methods and Protocols), vol 872. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-61779-797-2_1
6. DeBord LC, Pathak RR, Villaneuva M, et al. The chick chorioallantoic membrane (CAM) as a versatile patient-derived xenograft (PDX) platform for

precision medicine and preclinical research. Am J Cancer Res. 2018;8(8):1642-1660. Published 2018 Aug 1.

7. Clark DJ, Dhanasekaran SM, Petralia F, et al. Integrated Proteogenomic Characterization of Clear Cell Renal Cell Carcinoma [published correction appears in Cell. 2020 Jan 9;180(1):207]. Cell. 2019;179(4):964-983.e31. doi:10.1016/j.cell.2019.10.007.

8. Mertins P, Tang LC, Krug K, Clark DJ, Gritsenko MA, Chen L, Clauser KR, Clauss TR, Shah P, Gillette MA, Petyuk VA, Thomas SN, Mani DR, Mundt F, Moore RJ, Hu Y, Zhao R, Schnaubelt M, Keshishian H, Monroe ME, Zhang Z, Udeshi ND, Mani D, Davies SR, Townsend RR, Chan DW, Smith RD, Zhang H, Liu T, Carr SA. Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. Nat Protoc. 2018 Jul;13(7):1632-1661. doi: 10.1038/s41596-018-0006-9. PMID: 29988108; PMCID: PMC6211289.

9. Chen, L.; Zhang, B.; Schnaubelt, M.; Shah, P.; Aiyetan, P.; Chan, D.; Zhang, H.; Zhang, Z. MS-PyCloud: An Open-Source, Cloud Computing-Based Pipeline for LC-MS/MS Data Analysis. bioRxiv 2018.

10. Kim, S.; Pevzner, P. A. MS-GF+ Makes Progress towards a Universal Database Search Tool for Proteomics. Nat. Commun. 2014, 5, 5277. https://doi.org/10.1038/ncomms6277.

11. Kim, S.; Gupta, N.; Pevzner, P. A. Spectral Probabilities and Generating Functions of Tandem Mass Spectra a Strike Against the Decoy Database-Supplement-Supplementary_Info. J. Proteome Res. 2008, 7 (8), 1–6.

12. Elias, J. E.; Gygi, S. P. Target-Decoy Search Strategy for Increased Confidence in Large-Scale Protein Identifications by Mass Spectrometry. Nat. Methods 2007, 4 (3), 207–214. https://doi.org/10.1038/nmeth1019.

13. Raivo Kolde pheatmap: Pretty Heatmaps. R package version 1.0.10, https://CRAN.R-project.org/package=pheatmap (2018).

14. Blighe, K, S Rana, and M Lewis. 2018. “EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling.”

Proteomics Data Availability: The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [1] partner repository with the dataset identifier PXD021545 and 10.6019/PXD021545. Please use the following username and password to access the data:

Username: reviewer_pxd021545@ebi.ac.uk
Password: MYmOhJvc

Acknowledgements

We thank all members of the Wirtz Lab for discussions and feedback on this project. We also thank Alan Meeker and Sujayita Roy from the Oncology Tissue Services IHC Core at Johns Hopkins Medical Campus for their assistance. This work was supported through grants from the National Cancer Institute (U54CA143868) and the National Institite on Aging (U01AG060903) to D.W. and P.H.W.

Affiliations

[1]*Department of Chemical and Biomolecular Engineering, Johns Hopkins University, 3400 N Charles St, Baltimore, Maryland 21218, USA*

[2]*Johns Hopkins Physical Sciences– Oncology Center and Institute for NanoBioTechnology, Johns Hopkins University, 3400 N Charles St, Baltimore, Maryland 21218, USA*

[3]*Department of BioPhysics, Johns Hopkins University, 3400 N Charles St, Baltimore. MD 21218, USA*

[4]*Department of Pathology, The Johns Hopkins University School of Medicine, Baltimore 21231, Maryland, USA*

[5]*Department of Medical Biophysics, University of Toronto, Toronto, ON*

[6]*Biological Sciences Platform, Sunnybrook Research Institute, Toronto, ON*

Contributions

G.C.R. and D.W. developed the hypothesis and designed experiments. G.C.R. performed most experiments and data analysis. M.N.K., J.C., R.C., B.S. assisted with the experiments. D.C., T.L., H.Z. performed proteomics experiments, assisted in the analysis, and edited the manuscript. D.J.D. and H.S.L. performed CAM experiments, assisted in the analysis, and edited the manuscript. P.W. and Q.Z. assisted with the analysis of the clinical data. G.C.R. and D.W. wrote the manuscript with input from M.N.K. and B.S.

Corresponding Author

Correspondence to Denis Wirtz (wirtz@jhu.edu)

Competing interests

No declaration of competing interests from any authors.