The Orphan G Protein-Coupled Receptor GPR52 is a Novel Regulator of Breast Cancer Multicellular Organization

Sarah Z. Hanif[1,2], CheukMan Cherie Au[2], Ingrid Torregroza[3], Syeda Y. Jannath[2,4], Tabassum Fabiha[5], Bhavneet Bhinder[6], Michael Washburn[7,8], Dominic Devost[9], Shuchen Liu[2], Priya Bhardwaj[2], Todd Evans[3,] Pradeep Kumar Anand[10], Robert Tarran[10], Sailesh Palikhe[14], Olivier Elemento[6], Lukas Dow[11], John Blenis[11,12,13], Terence E. Hébert[9], Kristy A. Brown[2,8,14]

1. Weill Cornell Medicine/Rockefeller University/Memorial Sloan-Kettering Cancer Center Tri-Institutional MD-PhD Program, New York, NY, USA.
2. Department of Medicine, Weill Cornell Medicine, New York, NY, USA.
3. Department of Surgery, Weill Cornell Medicine, New York, NY, USA.
4. Hunter College, City University of New York, New York, NY, USA.
5. Columbia University Computational & Systems Biology Program, Sloan-Kettering Institute of Memorial Sloan-Kettering Cancer Center, New York, NY, USA.
6. Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, New York, USA.
7. Department of Cancer Biology, University of Kansas Medical Center Kansas City KS USA.
8. University of Kansas Cancer Center, Kansas City, KS, USA.
9. Department of Pharmacology and Therapeutics, McGill University, Montréal, QC, Canada.
10. Division of Genetic, Environmental and Inhalational Disease, Department of Internal Medicine, University of Kansas Medical Center, Kansas City, KS, USA
11. Sandra and Edward Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA.
12. Department of Pharmacology, Weill Cornell Medicine, New York, NY, USA.
13. Department of Biochemistry, Weill Cornell Medicine, New York, NY, USA.
14. Department of Cell Biology and Physiology, University of Kansas Medical Center Kansas City, KS, USA.

The authors declare no potential conflicts of interest.

Running Title
GPR52 Regulates Cancer Cell Clustering and Organization

Statement of Significance

We show that loss of the orphan G protein-coupled receptor GPR52 in human breast cell lines leads to increased cell clustering, hybrid/partial EMT, and increased tumor burden in zebrafish.

Abstract

G protein-coupled receptors (GPCRs) are the largest class of membrane-bound receptors and transmit critical signals from the extracellular to the intracellular spaces. Transcriptomic data of resected breast tumors shows that low mRNA expression of the orphan GPCR GPR52 correlates with reduced overall survival in breast cancer patients, leading to the hypothesis that loss of GPR52 supports breast cancer progression. CRISPR-Cas9 was used to knockout GPR52 in human triple-negative breast cancer (TNBC) cell lines MDA-MB-468 and MDA-MB-231, and in the non-cancerous breast epithelial cell line, MCF10A. Loss of GPR52 was found to be associated with increased cell-cell interaction in 2D cultures, altered 3D spheroid morphology, and increased propensity to organize and invade collectively in Matrigel. Furthermore, GPR52 loss was associated with features of EMT in MDA-MB-468 cells. To determine the *in vivo* impact of GPR52 loss, MDA-MB-468 cells were injected into zebrafish and loss of GPR52 was associated with a greater total cancer area compared to control cells. RNA-sequencing and proteomic analyses of GPR52-null breast cancer cells reveal an increased cAMP signaling signature. Consistently, we found that treatment of wild-type (WT) cells with forskolin, which stimulates production of cAMP, induces some phenotypic changes associated with GPR52 loss, and inhibition of cAMP production rescued some of the GPR52 KO phenotypes. Overall, our results reveal GPR52 loss as a potential mechanism by which breast cancer progression may occur and support the investigation of GPR52 agonism as a therapeutic option in breast cancer.

Introduction

Metastasis is the primary cause of death in breast cancer patients (1). The process required for cancer cells of solid tumors to metastasize is intensive, and the cells undergo several adaptive processes to enhance their metastatic potential. These include changes in cell-cell and cell-matrix adhesion, transitions between epithelial and mesenchymal cell states, and the ability to degrade and invade through tissue (2, 3). However, the upstream regulators of these processes are not well-characterized, which limits mechanistic understanding and therapeutic intervention.

G protein-coupled receptors (GPCRs) represent the largest protein family encoded by the human genome (4). These receptors consist of an extracellular N-terminus followed by seven transmembrane α-helices, which are connected by three intracellular and three extracellular loops, and a cytoplasmic C-terminal tail (5). Their transmembrane structure enables transmission of critical signals between the extracellular and intracellular spaces. The dissociation of the heterotrimeric G protein upon GPCR activation can regulate a diverse array of downstream molecules, allowing for regulation of various cell processes including proliferation, migration, adhesion, metabolism (6, 7).

GPR52 is a structurally unique GPCR that is enriched in the basal ganglia and whose endogenous ligand remains unknown, rendering it an orphan receptor (8). It has

garnered increased attention in recent years due to its emerging potential as a neurotherapeutic target for schizophrenia and Huntington's disease (9, 10). We examined GPR52 mRNA levels in nineteen solid tumor types and determined that GPR52 is significantly downregulated in tumor samples in fifteen of these (11). However, the role of GPR52 in cancer progression has not yet been reported. In patients with breast cancer, we found that GPR52 expression was further reduced in metastases compared to the primary tumor (11). Low GPR52 mRNA expression in resected breast tumors is also associated with a reduction in overall survival (12).

We generated GPR52 KO cancerous and non-cancerous breast epithelial cells in the widely used MDA-MB-468, MDA-MB-231, and MCF10A lines. Loss of GPR52 led to an increase in cell-cell interaction in 2D cultures, with the formation of cell clusters in each cell line. Transmission electron microscopy (TEM) of WT and GPR52 KO cells revealed differences in cell-cell adhesion properties, including length of the cell-cell interface and proximity of cells along this interface. In 3D Matrigel cultures, GPR52 KO was associated with changes in organization and morphology of MDA-MB-468 and MDA-MB-231 spheroids. Furthermore, GPR52 loss increased the propensity of the breast cancer cells to organize and invade collectively when cultured in Matrigel. Lastly, we found that culture of GPR52 KO cells on poly-D-lysine led to a partial EMT.

RNA-sequencing and proteomic studies of GPR52-null cells demonstrated an upregulation of several pathways implicated in breast cancer, including cAMP signaling (13)**.** Furthermore, we found that phosphorylation of CREB was increased in GPR52 KO cells and that treatment of WT cells with forskolin, which stimulates production of cAMP, promotes features associated with GPR52 loss while inhibition of cAMP production rescued some of the GPR52 KO phenotypes.

Overall, our results reveal GPR52 loss as a potential mechanism by which important processes in breast cancer progression, such as EMT and changes in multicellular organization, may occur. These processes have long been implicated in many solid tumors, but critical and targetable upstream regulators have not been identified. As GPCRs are the targets of more than 1/3 of FDA-approved small-molecule drugs, this data supports the investigation of GPR52 agonism as a viable therapeutic approach in breast cancer (14).

Results

Low GPR52 expression is associated with increased breast cancer progression and reduced survival probability of breast cancer patients

To date, there are no reports of the physiological role of GPR52 in any cancer type. However, GPR52 mRNA expression level has been reported in many transcriptome profiles of resected cancerous and non-cancerous tissues. We compared GPR52 mRNA expression level in normal and tumor samples in tissues from which solid tumors arise (Fig. 1A) using the TNMplot webtool (11). We found that in the majority of these tissue types, GPR52 mRNA expression level was lower in tumor compared to non-

cancerous samples (Fig. 1A, $p<0.05$ indicated with an asterisk). Importantly, low GPR52 mRNA expression in resected breast cancer correlates with a reduction in overall survival for patients with all subtypes of breast cancer (Fig. 1B), and notably in triple negative breast cancer (Fig. 1C) (12). Furthermore, GPR52 mRNA expression was lower in metastatic nodes compared to the primary resected tumor (Fig. 1D, $p=4.45e-17$) (11). Importantly, we found that GPR52 was detectable and differentially expressed in several human breast cell lines (Fig. 1E).

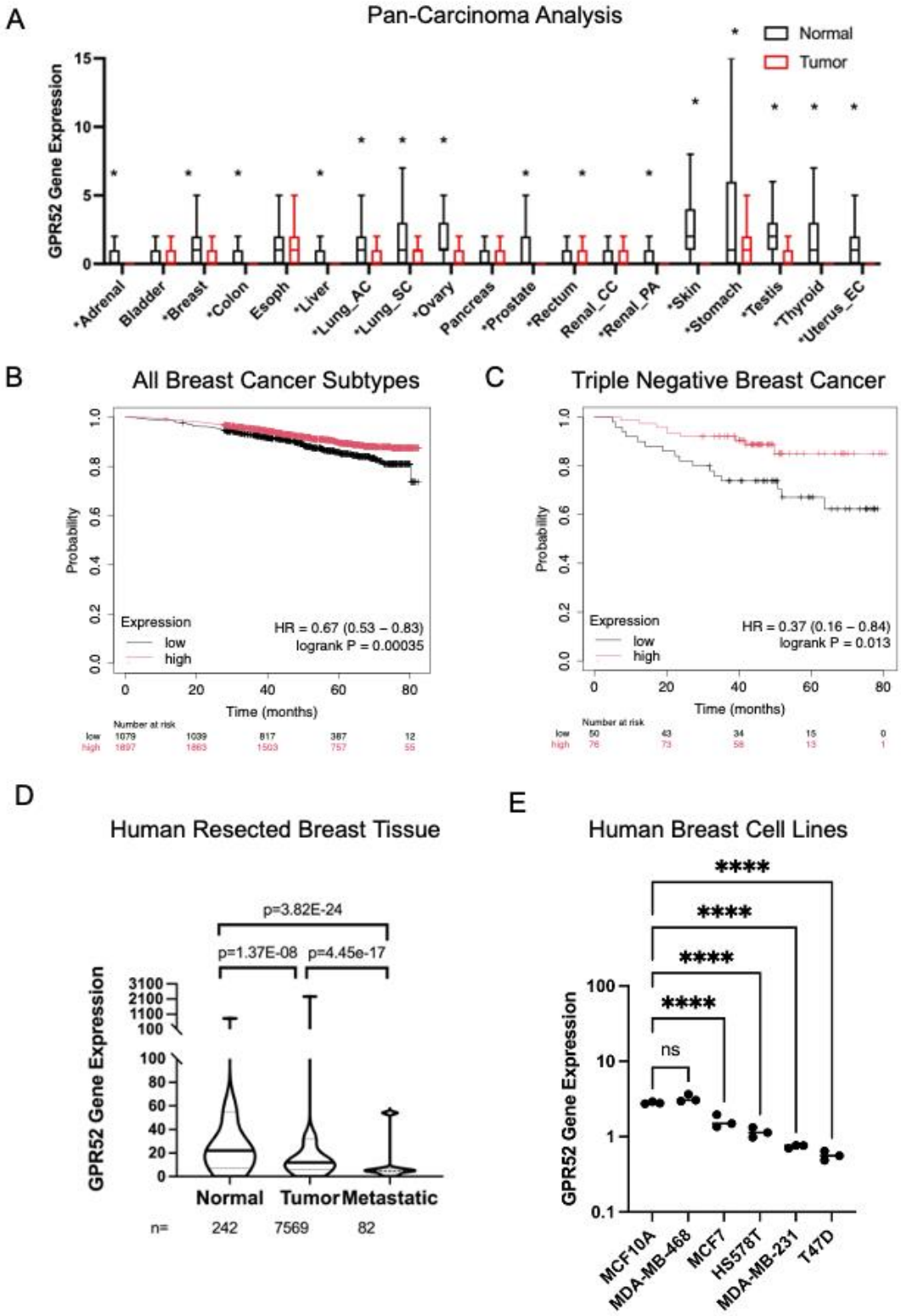

A
Pan-Carcinoma Analysis
Normal
Tumor
GPR52 Gene Expression
*Adrenal
Bladder
*Breast
*Colon
Esoph
*Liver
*Lung_AC
*Lung_SC
*Ovary
Pancreas
*Prostate
*Rectum
Renal_CC
*Renal_PA
*Skin
*Stomach
*Testis
*Thyroid
*Uterus_EC
B
All Breast Cancer Subtypes
Probability
Expression
low
high
HR = 0.67 (0.53 – 0.83)
logrank P = 0.00035
Time (months)
Number at risk
low 1079 1039 817 387 12
high 1897 1863 1503 757 55
C
Triple Negative Breast Cancer
Probability
Expression
low
high
HR = 0.37 (0.16 – 0.84)
logrank P = 0.013
Time (months)
Number at risk
low 50 43 34 15 0
high 76 73 58 13 1
D
Human Resected Breast Tissue
p=3.82E-24
p=1.37E-08
p=4.45e-17
GPR52 Gene Expression
Normal
Tumor
Metastatic
n= 242 7569 82
E
Human Breast Cell Lines
****
****
****
****
ns
GPR52 Gene Expression
MCF10A
MDA-MB-468
MCF7
HS578T
MDA-MB-231
T47D

Figure 1. Low GPR52 expression is associated with increased breast cancer progression and reduced survival probability of breast cancer patients. A) The median GPR52 mRNA expression, interquartile range, minimum value, and upper whisker are plotted for normal (black) and tumor (red) tissues. Mann-Whitney U tests were conducted to determine statistical significance. *P <0.05. AC=adenocarcinoma, SC=squamous cell, CC=clear cell, PA=papillary cell, EC=endometrial carcinoma. B) KMplot breast cancer overall survival curves for patients with low versus high GPR52 mRNA expression in resected tumors for (B) all breast cancer subtypes and (C) triple negative breast cancer. D) GPR52 mRNA expression collected from a gene chip dataset of non-cancerous breast tissue, primary tumor, and metastases of individuals with breast cancer (un-paired). Data are presented as median with upper and lower quartiles and minimum and maximum values. One-way ANOVA, p<0.05. E) GPR52 mRNA expression was determined by QPCR and normalized to the housekeeping gene RPL32. The normalized GPR52 expression was then divided by the average expression of GPR52 across the cell lines. n=3, line=median. One-way ANOVA, P-value<0.05; ****P <0.00005, ns=not significant.

GPR52 regulates cell-cell adhesion properties of cancerous and non-cancerous breast epithelial cells

We utilized CRISPR-Cas9 to generate indels in *GPR52* in two human breast cancer cell lines (MDA-MB-231 and MDA-MB-468) and one non-cancerous breast epithelial cell line (MCF10A) using two different guides RNAs (Supplementary Fig. S2). We observed that the GPR52 KO MDA-MB-231, MDA-MB-468, and MCF10A cell lines generated from both guide RNAs formed clusters of tightly packed cells with clearly defined borders in monolayer culture (Fig. 2A). To further investigate the effect of GPR52 loss on the interactions between cells, we performed TEM of vector control and GPR52 KO MDA-MB-468 cells that were cultured in suspension (Fig. 2B). We observed that the length of the interface between wild-type (WT) cells was longer than that between GPR52 KO cells (Fig. 2C). However, GPR52 KO cells interacted more closely with one another along the length of their interface than WT cells, as there was most intercellular free space along the WT cell-cell interface (Fig. 2D). The diameter of cells was not significantly different between the groups, suggesting that the reduction in cell interface length in the GPR52 KO group was not due to a reduction in cell size (Fig. 2E).

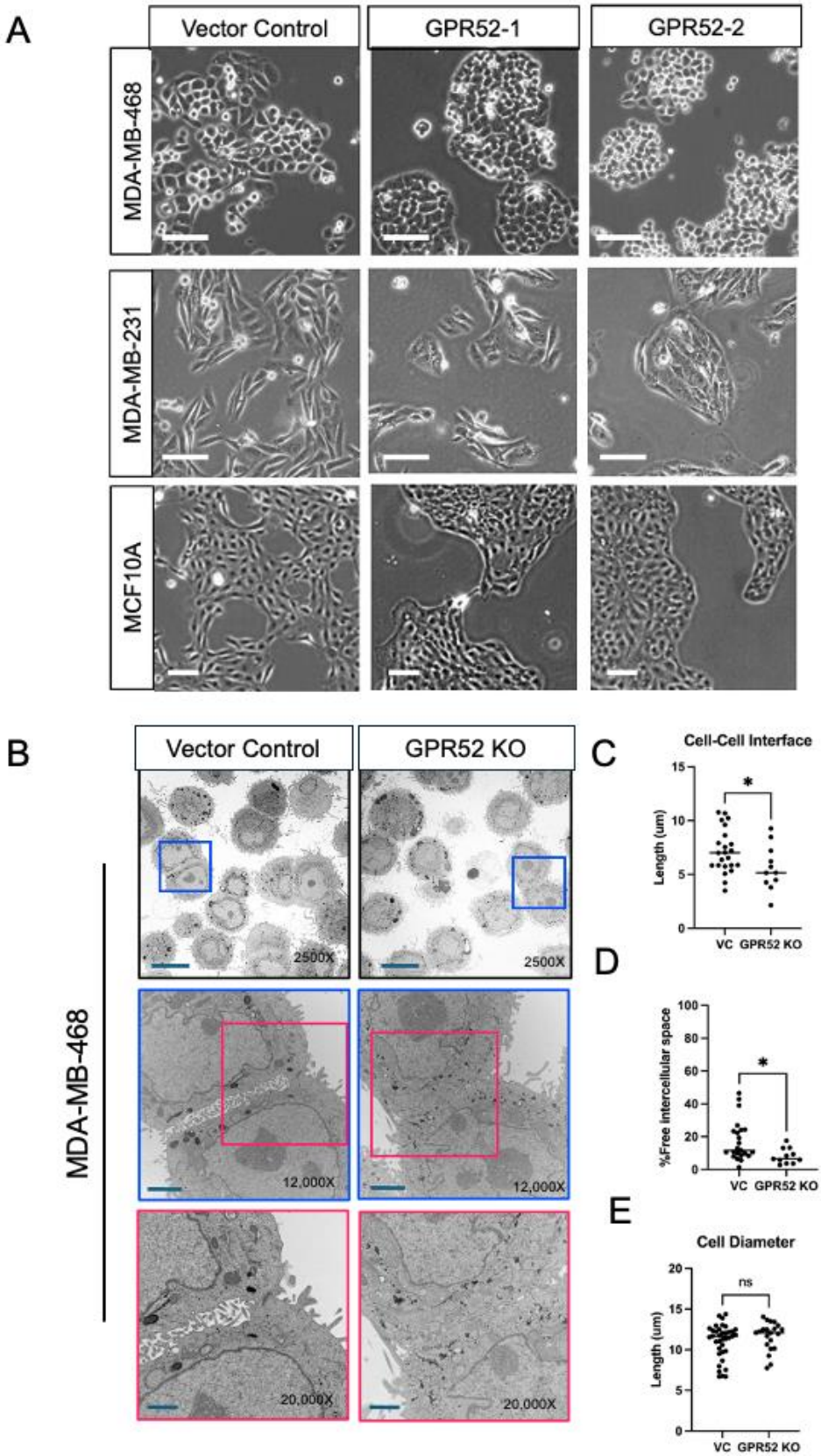

A
Vector Control
GPR52-1
GPR52-2
MDA-MB-468
MDA-MB-231
MCF10A
B
Vector Control
GPR52 KO
MDA-MB-468
2500X
2500X
12,000X
12,000X
20,000X
20,000X
C
Cell-Cell Interface
Length (um)
VC
GPR52 KO
D
%Free intercellular space
VC
GPR52 KO
E
Cell Diameter
ns
Length (um)
VC
GPR52 KO

Figure 2. GPR52 regulates cell-cell adhesion properties of cancerous and non-cancerous breast epithelial cells. A) Human breast cell lines were grown on tissue culture-treated plastic under standard cell culture conditions and imaged at 50-60% confluency with light microscopy. Scalebar=100 µm. B) MDA-MB-468 cells cultured in suspension were visualized with TEM from 2500-20,000x magnification. Scalebars: 2500x=10 μm; 12,000x=2 μm, 20,000x=1 μm. C) The length of the interface between cells, D) fraction of the cell-cell interface that was occupied by free space, and E) cell diameter were determined using 2500x images. Student's t-test, P-value<0.05; *P <0.05, ns=not significant. Line=median. VC=vector control.

Loss of GPR52 leads to altered breast cancer cell organization in Matrigel

Cell culture within a basement membrane extract such as Matrigel recapitulates some of the extracellular conditions that cells experience as they organize *in vivo* (15). We observed that WT MDA-MB-468 cells formed rounded spheroids with well-defined borders in Matrigel culture, while GPR52 KO MDA-MB-468 cells formed disorganized spheroids with irregular borders (Fig. 3A-B, top rows; 3C). This disorganization was further observed as the cells within the GPR52 KO spheroids were not as tightly packed as the WT cells. Furthermore, more GPR52 KO MDA-MB-468 spheroids were observed than WT spheroids (Fig. 3D). GPR52 loss in MDA-MB-231 resulted in a reduction in the average size of spheroids (Fig. 3A-B, bottom rows; 3E), but no change in the number or shape of spheroids (Supplementary Fig. S3A-B).

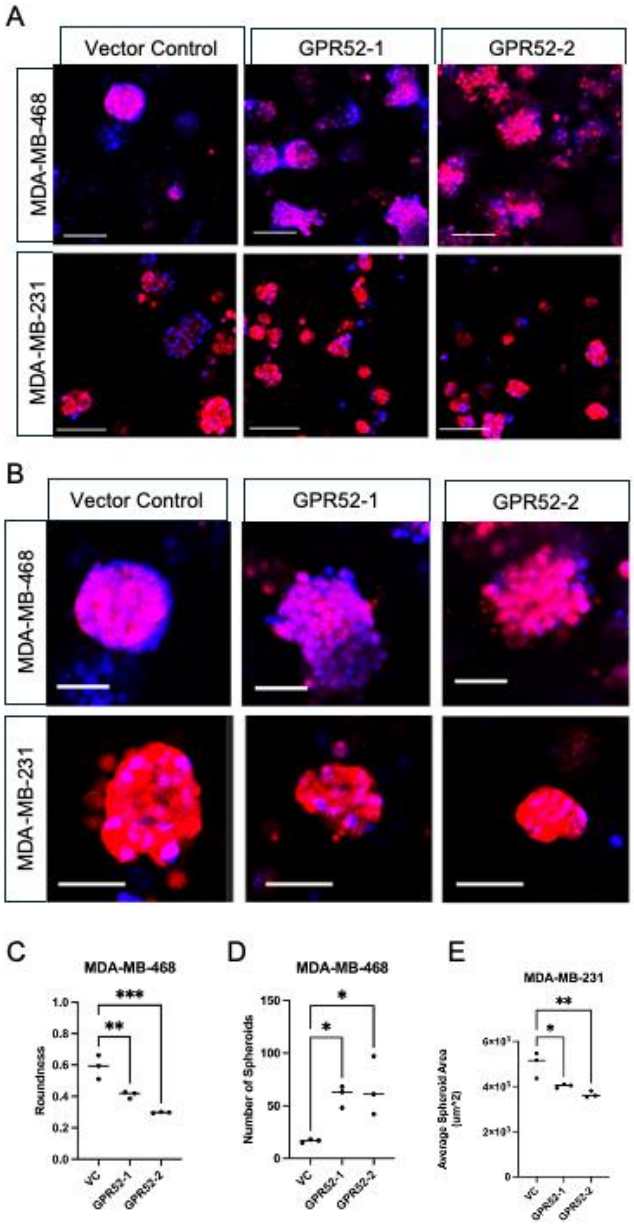


Figure 3. Loss of GPR52 leads to altered breast cancer cell organization in Matrigel. Cells were cultured in Matrigel for 10 days under complete media conditions. (A-B) cytoplasmic tdTomato (red) and 1:1000 nuclear Hoechst 33342 stain (blue) allow visualization of MDA-MB-468 (top row) and MDA-MB-231 (bottom row) WT and GPR52 KO spheroids. (A) Scalebar=100 μm, (B) Scalebar=50 μm. C-E) The roundness, number, and average area of spheroids was determined based on the tdTomato signal. VC=vector control. n=3, One-way ANOVA, P-value<0.05; **P* <0.05, ***P* <0.005, ****P* <0.0005, ns=not significant. Line=median.

GPR52 loss is associated with EMT, collective invasion, and an altered pattern of ECM digestion in breast cancer

The coating of plastic with poly-D-lysine (PDL) results in cell spreading and attachment but reduced phosphorylation and activation of cytoplasmic focal adhesion proteins by integrins, the major cell adhesion molecule that links cells to the extracellular matrix (16, 17). Therefore, PDL coating has been used as a method to distinguish cellular features associated with integrin activity (16, 17). We cultured MDA-MB-468 cells on plastic coated with poly-D-lysine (PDL) and found that this led to elongation of GPR52 KO MDA-MB-468 cells (Fig. 4A). Western blotting of GPR52 KO MDA-MB-468 cells revealed an upregulation of mesenchymal cell markers, such as snai1 and vimentin, but the cells notably continued to express E-cadherin (Fig. 4B).

As changes in cell adhesion properties and EMT can impact cancer invasiveness, we next determined whether loss of GPR52 affected breast cancer cell invasion through Matrigel (Fig. 4C). Interestingly, we observed at t=24 hours that GPR52 KO cells tended to organize and invade collectively in large clusters (MDA-MB-468) or sheets (MDA-MB-231) (Fig. 4D). We categorized all cancer foci with an area ≥ 3140 μm^2 at z=10 μm as Class B structures (Fig. 4C, yellow outline). This threshold was selected because the diameter of MDA-MB-468 and MDA-MB-231 cells was found to range from 10-20 μm, and therefore the cross-sectional area of approximately 10 cells was $10*\pi*r^2 = 10*\pi*10^2$=~3140). We then calculated the sum of the area of all Class B structures and divided this by the sum of the area of all cancer foci at z=10 μm. We found that the fraction of the area occupied by Class B structures was in fact increased in MDA-MB-468 and MDA-MB-231 GPR52 KO groups, suggesting an increased propensity to organize collectively (Fig. 4E-F).

Although collective invasion has been studied by many groups, the effect of this behavior on the ECM that the cells invade through has not been well-characterized. We incorporated dye-quenched (DQ)-collagen IV, which contains sequestered fluorophores that are released following proteolysis of collagen IV, a major component of the basement membrane that breast cancer cells invade through, to assess the pattern of ECM degradation (18). We found at the 24-hour timepoint that DQ signal co-localized strongly with the GPR52 KO cells, while it was more diffuse and less co-localized with the vector control cells for both MDA-MB-468 and MDA-MB-231 (Fig. 4H-I). Furthermore, we normalized the area of DQ signal to the area occupied by cancer cells and found that it tended to be lower for the GPR52 KO groups compared to the vector control, suggesting a focal area of matrix degradation (Fig. 4J-K).

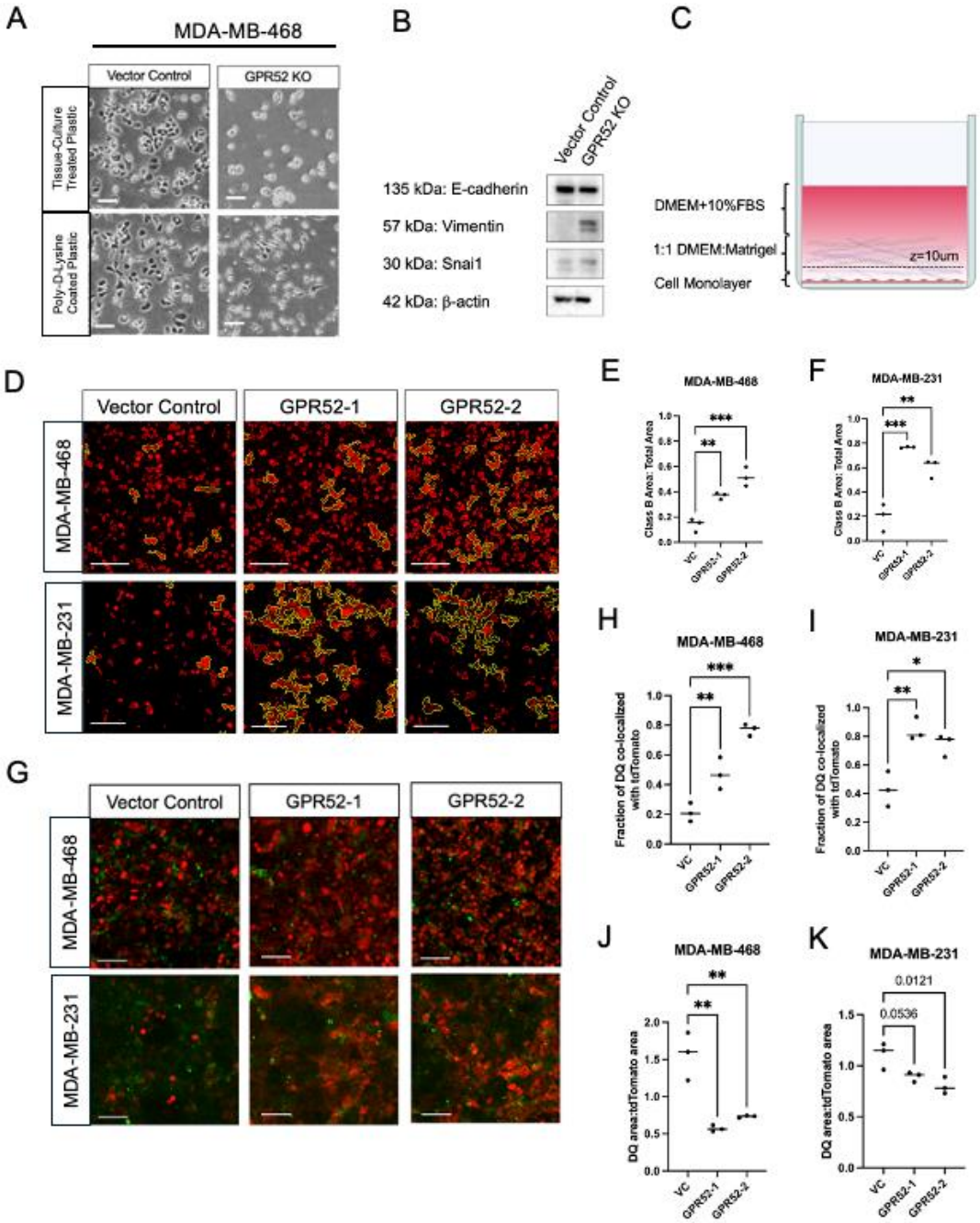


Figure 4. GPR52 loss is associated with EMT, collective invasion, and an altered pattern of ECM digestion in breast cancer. A) Vector control and GPR52 KO (sgRNA2) MDA-MB-468 cells were cultured on tissue-culture treated plates that were untreated (top row) or treated with 1μg/mL poly-D-lysine coated plastic (bottom row). Scalebar=100 μm. B) Western blotting of MDA-MB-468 cells cultured on poly-D-lysine. C) Schematic of invasion assay. FBS=fetal bovine serum, DMEM= Dulbecco's Modified Eagle Medium. D) Z-stacks obtained at t=24 hours were assessed at z=10 μm. All Class

B structures are outlined in yellow. Scalebar=150 µm. The fraction of the area occupied by Class B structures for (E) MDA-MB-468 and (F) MDA-MB-231. VC=vector control. n=3, One-way ANOVA, P-value<0.05. *P <0.05, **P <0.005, ns=not significant. Line=median. G) Representative images are shown of tdTomato-tagged cancer cells (red) and DQ-collagen (green) at t=24 hours at z=10 µm. Scalebar=100 µm. The fraction of DQ co-localized with tdTomato at z=10 µm (H-I) and area of DQ normalized to the area of tdTomato (J-K) were determined for MDA-MB-468 and MDA-MB-231, respectively. n=3, One-way ANOVA, P-value<0.05. *P <0.05, **P <0.005, ***P <0.0005, ns=not significant. Line=median.

Proteomic and transcriptomic analyses of GPR52 KO breast cell lines and resected breast tumors

We next conducted proteomic analyses of WT and GPR52 KO MDA-MB-468, MDA-MB-231, and MCF10A cell lines. Using Ingenuity Pathway Analysis (IPA, Qiagen), we observed upregulation of signatures associated with cellular homeostasis, viability, survival, proliferation, and migration in all three cell lines, while the signatures associated with organismal death and cell death of tumor cell lines were both significantly reduced (Fig. 5A). There were several upstream regulators that were predicted to be similarly altered across the GPR52 KO lines (Fig. 5B). We were particularly intrigued by the signature associated with the cAMP analogue 8-bromo-cAMP, as GPR52 activation and function has been associated with its regulation of intracellular cAMP levels (9, 10)

The TCGA-BRCA dataset reports RNA-sequencing data of resected breast tumors from a cohort of patients. The normalized GPR52 mRNA expression in TCGA-BRCA tumors was visualized (Fig. 5C) and it was determined that 45.7% did not express GPR52 while the remainder expressed non-zero levels of GPR52 (Fig. 5D). The RNA-sequencing datasets were compared for the GPR52 non-expressing and -expressing cohorts and the differentially expressed genes (DEGs) were imported into IPA. We also conducted RNA-sequencing of GPR52 KO and WT MDA-MB-468 cells that were cultured in monolayer and imported the DEGs into IPA. We identified several common signatures in both datasets. Notably, an upregulation of the signature associated with cAMP response element binding protein (CREB1), a transcription factor that is activated downstream of cAMP signaling, was increased in both datasets (Fig. 5E) (19). Based on phosphoproteomic analyses of MDA-MB-468, MDA-MB-231, and MCF10A cells, we identified kinases that were predicted to be active in the vector control (Fig. 5F) and GPR52 KO (Fig. 5G) datasets (20). The rank of each of the top kinases in the GPR52 KO cell lines was then compared to its rank in the vector control group (Fig. 5H). This revealed the greatest increase in rank of the kinase VRK1, which regulates cell cycle progression via phosphorylation and activation of CREB (21).

A

Top Diseases and Functions	MDA-MB-468	MDA-MB-231	MCF10A
Organismal death	-4.044	-2.366	-5.462
Cellular homeostasis	1.523	1.453	3.127
Cell survival	2.153	0.957	2.91
Cell viability	2.07	1.038	2.89
Cell viability of tumor cell lines	1.642	1.303	2.547
Connective or soft tissue tumor	-1.607	-2.369	-1.316
Cell proliferation of tumor cell lines	2.485	-0.816	1.927
Growth of tumor	2.408	-0.385	2.368
Cell movement	0.301	1.173	3.479
Cell spreading	-1.419	2.43	0.988
Leukopoesis	N/A	2.008	2.732
Cell movement of endothelial cells	-0.8	N/A	3.873
Cardiopathy	-1.654	-0.252	-2.743
Non-malignant disorder	-2.472	1.577	-0.58
Migration of endothelial cells	-0.61	N/A	3.966
Pancreaticobiliary cancer	-2.204	-1.964	0.329
Pancreaticobiliary carcinoma	-2.204	-1.964	0.329
Hepato-pancreato-biliary cancer	-2.103	-1.492	0.892
Migration of cells	0.444	0.657	3.358
Migration of tumor cell lines	-0.608	-0.643	3.145
Cell proliferation of breast cancer cel	1.719	0.797	1.865
Cell death of tumor cell lines	-1.691	-1.021	-1.644
Necrosis of epithelial tissue	0.993	1.573	1.743
Infection of cells	1.859	1.466	0.966
Upper abdominal cancer	-2.085	-1.105	1.034
Abdominal cancer	-2.409	-1.093	0.699

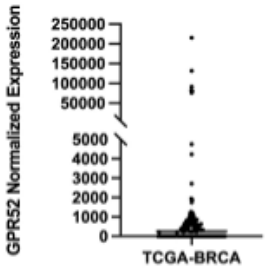
negative z-score positive z-score

B

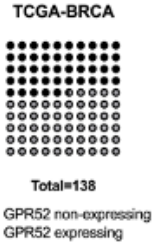
Upstream Regulators	MDA-MB-468	MDA-MB-231	MCF10A
ST1926	-3.86	-1.195	-1.726
beta-estradiol	2.087	2.078	2.373
CD437	-3.559	-1.134	-1.667
medroxyprogesterone acetate	1.511	2.034	2.183
Z-LLL-CHO	2.647	N/A	2.77
HRAS	-2.538	-1.026	1.693
HGF	0.441	-2.209	2.583
MEK	1.203	-1.521	2.169
topotecan	-2.429	-1.5	-0.894
INFG	0.758	0.287	3.673
TNF	0.843	0.075	3.775
tetradecanoylphorbol acetate	0.368	0.857	3.393
NUPR1	-1.414	1.069	-2.117
LEP	N/A	2.556	1.938
8-bromo-cAMP	1.604	N/A	2.846
cisplatin	1.787	0.878	1.718
miR-30a-3p (and others with seed UUUCACU	-2.423	-1.131	-0.49
miR-16-5p (and others with seed AGCAGCA)	-2.496	-1.221	0.295
CNTF	1.667	1.066	1.238
IL4	2.054	0.297	1.602
XBP1	0.664	2.041	1.211
STAT3	1.25	N/A	2.644
miR-1-3p (and others with seed GGAAUGU)	-2.177	-0.711	-0.99
ASPSCR1-TFE3	1.89	1.982	N/A
SYVN1	1.732	0	2.111

negative z-score positive z-score

C



D

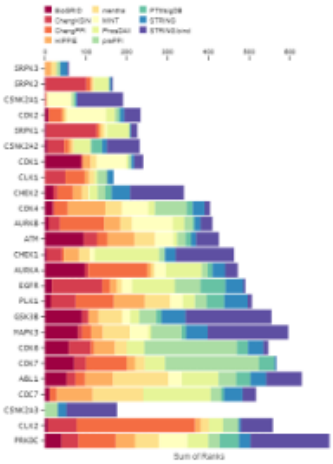


E

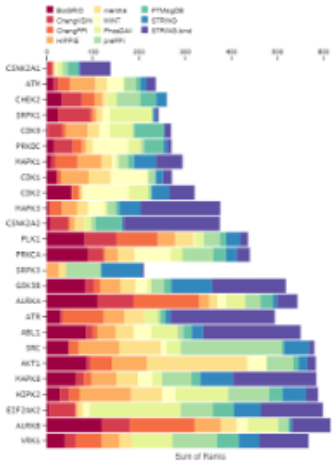
Upstream Regulators	MDA-MB-468	TCGA-BRCA
epicatechin	2.07E+00	2.701
FOXA2	2.685	2.008
N-CoR	1.698	-2.646
TNC	1.941	-2.219
CDC20	-1.89	-2.236
propylthiouracil	-1.769	-2.331
10E-12Z-octadecadienoic acid	1.681	-2.275
TNF	-1.014	-2.937
REST	-1.42	-2.487
PDPK1	-1.924	-1.944
NR2F2	2.592	1.236
beta-estradiol	2.192	1.615
CFI-402257	2.36	-1.432
SOX11	1.809	-1.943
cycloheximide	1.103	-2.576
ciprofibrate	1.421	2.241
bexarotene	1.621	2.007
insulin	1.444	2.179
PAF1	2.2	-1.4
PPARG	-0.649	2.913
NGF	2.324	1.193
BDNF	1.946	1.518
rosiglitazone	1.13	2.331
EFNA3	1.414	2
EFNA4	1.414	2
MALAT1	1.026	-2.387
CREB1	2.939	0.468
CTNNB1	2.493	-0.894
ritonavir	-1.254	-2.111
TEAD4	-0.967	-2.376

negative z-score positive z-score

F



G



H

Phosphoprotein Upstream Regulator	ΔRank(GPR52 KO-VC)
VRK1	240
HIPK2	79
PRKCA	31
MAPK8	29
SRC	24
CDK9	23
MAPK1	23
AKT1	21
PRKDC	19
EIF2AK2	13
ATM	10
ATR	9
MAPK3	8
CHEK2	6
PLK1	4
ABL1	3
CSNK2A1	2
GSK3B	2
SRPK1	1
CDK1	-1
AURKA	-2
CDK2	-5
CSNK2A2	-5
SRPK3	-13
AURKB	-13

Increase in rank Decrease in rank

Figure 5. Proteomic and transcriptomic analyses of GPR52-null breast cell lines and resected breast tumors. (A-B) Differentially expressed proteins in the GPR52 KO

cell lines with P-value <0.05, FDR<0.05, and fold-change greater than two were imported into IPA for each of the three cell lines. The common predicted upstream regulators (A) and top diseases and functions (B) with a P-value <0.05 are depicted. Pathways are marked N/A (not-applicable) if the z-score is not determined. (C) Normalized expression of GPR52 in TCGA-BRCA cohort resected breast tumors D) Graphical representation of proportion of patients with undetected (black) and detected (grey) GPR52 mRNA in resected breast tumors. E) DEGs in GPR52 KO MDA-MB-468 cells and the GPR52-null TCGA-BRCA cohort with P-value <0.05, FDR<0.05, and fold-change greater than two were imported into IPA for pathway analysis. Upstream regulators predicted to be responsible for the DEGs observed in both datasets with a P-value <0.05 are depicted. (F-H) Phosphoproteomic analysis was conducted on the groups outlined in Fig 5A. KEA3 software was used to rank kinases that were predicted to be active in the vector control (F) and GPR52 KO (G) cell lines. H) The difference in rank of kinases which were reported for both the GPR52 KO and vector control groups.

cAMP production is regulated by GPR52 and mediates phenotypes associated with GPR52 loss

To further explore the relationship between GPR52 and cAMP signaling, we next determined whether GPR52 activation led to a change in intracellular cAMP, as was described by other groups (8, 9). To do this, we utilized a bioluminescence resonance energy transfer 2 (BRET2)-based EPAC sensor, GFP10-EPAC-RlucII (Supplementary Fig. S6A) (22). This modified form of EPAC contains a luminescent donor and fluorescent acceptor that are in proximity of one another when EPAC is not bound to cAMP. However, the binding of cAMP induces a conformational shift that leads to increased distance between the donor and acceptor and a reduction in energy transfer. We found that increasing the amount of GPR52 and/or the concentration of FTBMT, synthetic agonist of GPR52, led to a reduction in BRET ratio in HEK293 cells (Supplementary Fig. S6B). Next, we introduced FTBMT and the EPAC sensor to MDA-MB-468 cells and found that the BRET values tended to increase at higher doses (Supplementary Fig. S6C). To investigate whether GPR52 may couple with $G_{\alpha q}$, we also quantified intracellular Ca^{2+} levels over a similar range of FTBMT doses using the Ca^{2+}-sensitive dye Fluo-4. Despite observing a robust response to the sarcoendoplasmic reticulum calcium ATPase (SERCA) inhibitor thapsigargin, we observed no change in baseline or thapsigargin-induced cytoplasmic Ca^{2+} levels (Supplementary Fig. S6D).

We next explored whether cAMP could modulate the phenotypes observed with GPR52 loss in MDA-MB-468 cells. GPR52 KO cells exhibited less rounding and were able to spread out on plastic with adenylyl cyclase inhibitor (ACi) treatment (Fig. 6D). Interestingly, the expression of snai1 was induced by forskolin (FSK, a direct activator of adenylyl cyclase) treatment of WT cells (Fig. 6E). We found that treatment of WT MDA-MB-231 cells with FSK led to formation of cell clusters with sharper borders than the vehicle control-treated cells and that GPR52 KO cells treated with ACi were more likely to overlap with one another and form smaller clusters than grow adjacent to one another in large clusters as they did with vehicle control (Fig. 6F). We also assessed the effect

of cAMP modulation on the ECM degradation pattern of invasive cells (Fig. 6G) and found that ACi treatment did increase the area of DQ normalized to the area of tdTomato without affecting the degree of co-localization of DQ and tdTomato (Fig. 6H-I).

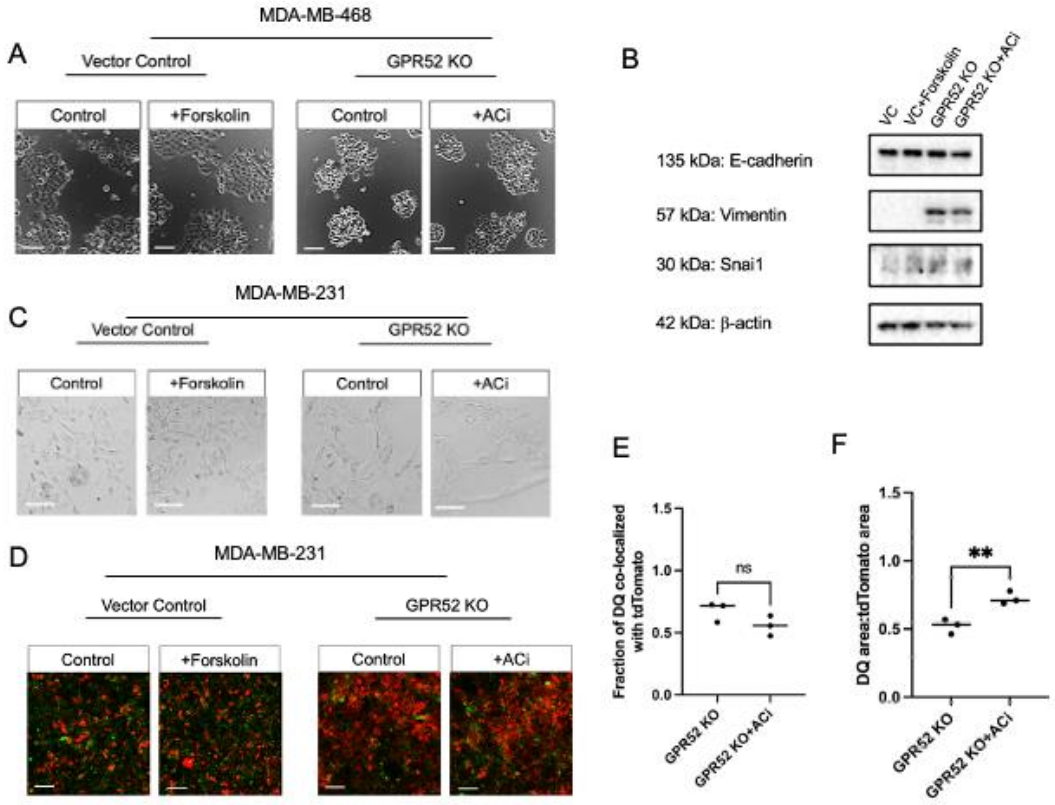


Figure 6. cAMP partially mediates phenotypes associated with GPR52 loss. A) Vector control (VC) and GPR52 KO (sgRNA2) MDA-MB-468 cells were cultured in monolayer and treated with 1 μM FSK, 1.4 μM ACi, or the vehicle control. Treatment was replaced every 72 hours over 6-8 days. Scalebar=100 μm. B) Western blotting of MDA-MB-468 cells that were treated as in (D) for 24 hours. C) VC and GPR52 KO (sgRNA2) MDA-MB-231 cells were cultured as in (D). Scalebar=100 μm. D) MDA-MB-231 cells were plated to confluency in monolayer and treated the next day with 1 μM FSK, 1.4 μM ACi, and the vehicle control treatments. Invasion assays were conducted as in **Fig. 4G**. Scalebar=100 μm. The fraction of DQ co-localized with tdTomato (E) and area of DQ normalized to the area of tdTomato (F) at z=10 μm at t=24 hours. n=3, Student's t-test, P-value<0.05; *P <0.05, **P <0.005, ns=not significant. Line=median.

Expression of the melanoma cell adhesion molecule (MCAM) is inversely related to GPR52 in breast cancer and is regulated by cAMP

RNA-sequencing analysis of MDA-MB-468 cells demonstrated that GPR52 loss is associated with differential expression of many cell adhesion molecules (CAMs) (Fig.

7A), including the melanoma cell adhesion molecule (MCAM). MCAM is highly expressed in large blood vessels, but recent studies have also described increased MCAM expression in certain cancer types and its promotion of cancer progression (23-25). Furthermore, in breast cancer, increased MCAM mRNA expression in resected tumors is associated with a reduction in overall survival (Fig. 7B) (12). Expression of MCAM and GPR52 mRNA are inversely correlated in resected breast tumors (Fig. 7C). We found that MCAM protein expression is increased in GPR52 KO MDA-MB-468 cells, and that 1.4 μM ACi treatment can reduce MCAM expression (Fig. 7D). Importantly, we also found that GPR52 KO is associated with increased expression of MCAM in MCF10A cells (Supplementary Fig. S7).

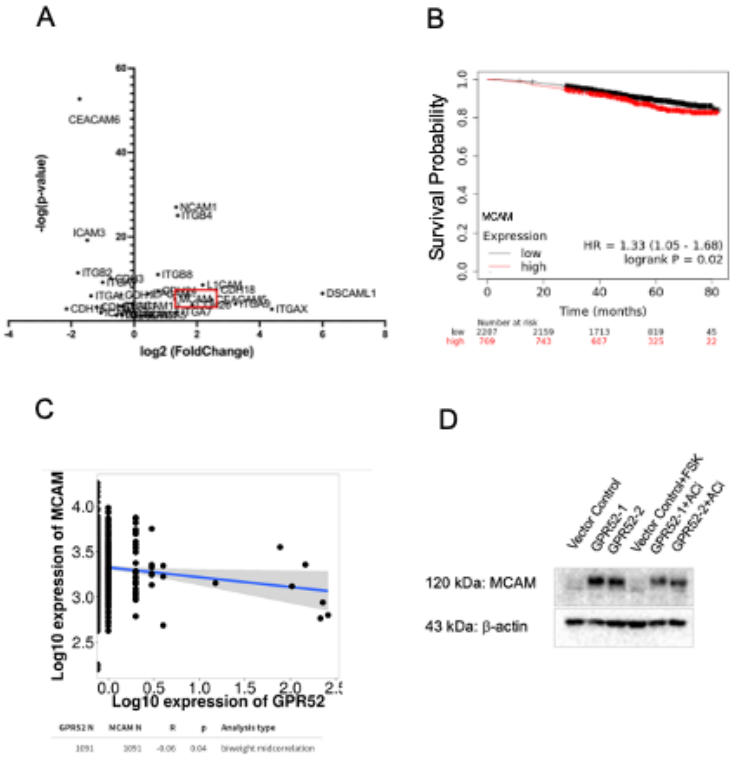


Figure 7. Expression of the melanoma cell adhesion molecule (MCAM) is inversely related to GPR52 in breast cancer and is regulated by cAMP. (A) CAMs that were differentially expressed in GPR52 KO MDA-MB-468 cells based on RNA-sequencing. P-value<0.05. n=3. (B) KMplot breast cancer RNA-seq webtool overall survival curves for patients with low versus high MCAM mRNA expression in resected tumors for all breast cancer subtypes. Low versus high cutoff was determined based on the maximum segregation between the groups. (C) TNMplot correlation of MCAM and GPR52 transcript expression in breast tumors from an RNA-sequencing dataset. (D) MDA-MB-468 cells were cultured in monolayer (TC-treated, non-PDL coated) and treated with forskolin (FSK), SQ22536 (ACi), or vehicle control for 24 hours. Cell lysates were probed for MCAM and beta-actin.

Loss of GPR52 is associated with increased TNBC burden in zebrafish

As we observed that loss of GPR52 led to changes in organization, cell-cell adhesion, and invasiveness in breast cancer cells, we wanted to determine whether these changes were associated with an increase in breast cancer burden *in vivo.* As our model of choice would also enable close monitoring of cancer cell organization and distribution, we utilized a zebrafish xenograft model and designed our study based on those previously described for human breast cancer cell lines (Fig. 8A) (26, 27). We utilized the TG(flk1:EGFP-NLS) zebrafish strain that constitutively expresses a green fluorescent protein in endothelial cells, allowing for monitoring of cancer cells in relation to the zebrafish vasculature (28).

We conducted two independent studies to compare the behavior of vector control MDA-MB-468 cells to GPR52 sgRNA1 and sgRNA2 cells, respectively. We found that WT and GPR52 KO MDA-MB-468 cells were detectable throughout the body of the zebrafish at 30 hours post-injection (hpi) (Fig. 8B). We also observed that WT and GPR52 KO cells circulated in the zebrafish bloodstream collectively and identified endothelial cells between cancer cells, suggesting interaction between the two cell types (Fig. 8C). The number of cancer foci did not differ significantly between groups at 30 hpi (Fig. 8D-E). However, the total cancer area, which we calculated as the sum of the area of cancer in the head (superior to the otolith) and trunk (distal to the injection site, not including the yolk sac) (Fig. 8F-I),) was significantly greater in the zebrafish injected with GPR52 KO cells at 30 hpi (Fig. 8J-K).

Increased clustering of cancer cells is associated with a reduction in sensitivity to cytotoxic chemotherapeutic drugs, particularly when used as single agents (2, 29). Doxorubicin is one of the most potent Food and Drug Administration-approved chemotherapeutic drugs and is used in breast cancer treatment (30, 31). We therefore designed a zebrafish xenograft therapeutic study that incorporated 8 μM doxorubicin or the vehicle control (milli-Q water) in the E3 water that zebrafish are maintained in at 5 hpi and then quantified cancer area at 30 hpi (32).

At 30 hpi, the total cancer area did not differ between the zebrafish receiving the vehicle or doxorubicin for the WT or GPR52 KO groups (Fig. 8L-M). However, doxorubicin did reduce the area of WT and GPR52 KO breast cancer cells in the head of the zebrafish, albeit to different extents (Fig. 8N-O).

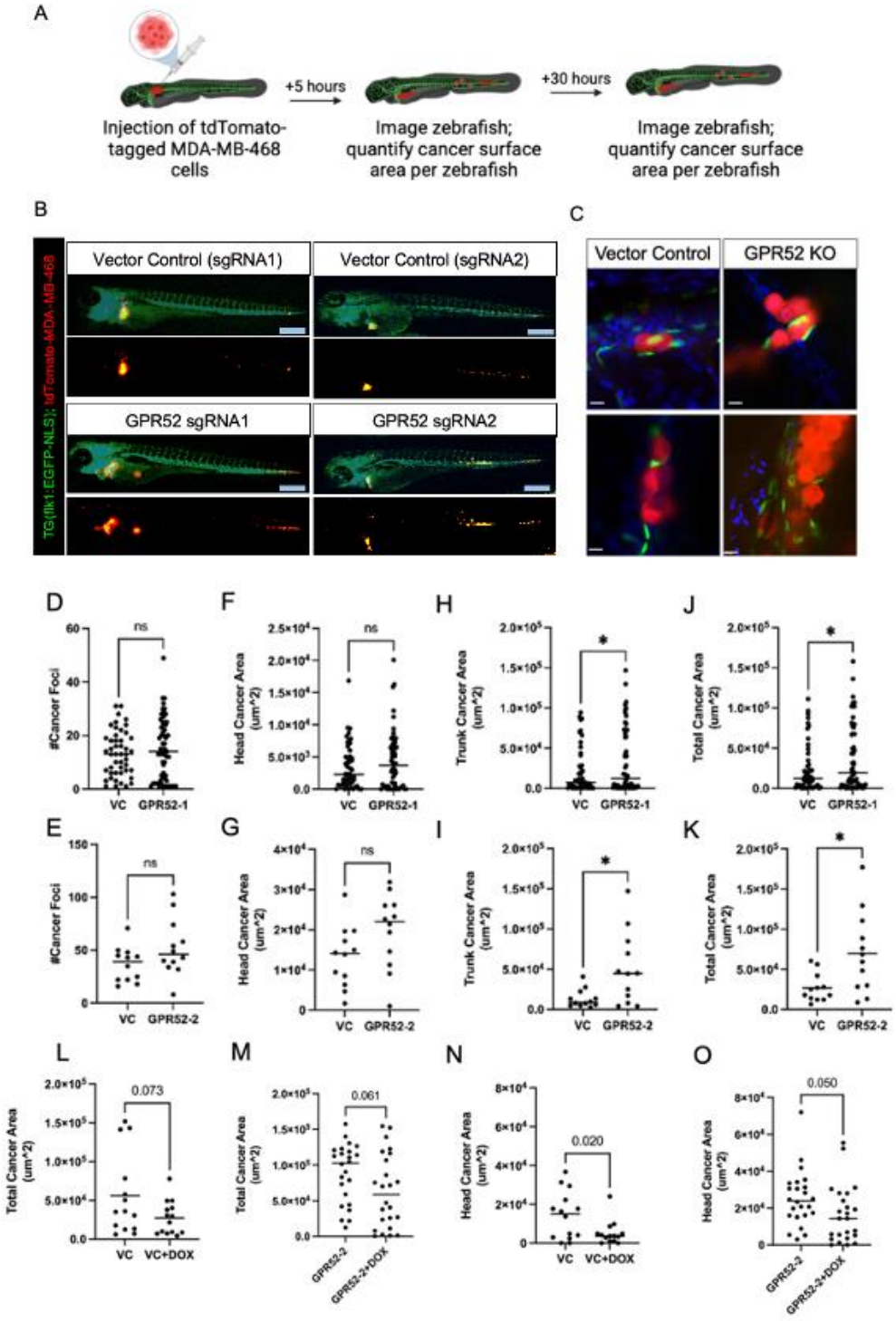


Figure 8. Loss of GPR52 is associated with increased TNBC burden in zebrafish. A) Schematic of zebrafish xenograft study. B) Tg(flk1:EGFP-NLS) zebrafish at 30 hpi. MDA-MB-468 cells (red) and endothelial cells (green) are visualized. Scalebar=300 μm. C) Visualization of the interaction between MDA-MB-468 cells (red) and endothelial cells (green) by confocal microscopy. Scalebar=10 μm. (D-E) The total number of

tdTomato foci per zebrafish. (F-G) The total area of tdTomato in the head and (H-I) trunk per zebrafish. (J-K) The sum of the area of tdTomato in the head and trunk is expressed as the total cancer area per zebrafish. VC=vector control. n=51 VC-1, n=59 GPR52-1, n=12 VC-2, n=12 GPR52-2. Student's t-test, P<0.05. *= *P*<0.05. ns=not significant. Line=median. (L-M) Zebrafish were treated with 8 μM doxorubicin or vehicle control (VC) and imaged at 30 hpi. The total area of tdTomato per zebrafish and (N-O) the area of tdTomato in the zebrafish head were determined. VC=vector control. DOX=doxorubicin. n=14 VC, n=24 GPR52-2. Student's t-test, P<0.05. ns=not significant. Line=median.

Discussion

The data presented herein demonstrate that GPR52 is a novel regulator of multicellular organization in breast cancer and its loss can promote features of cancer progression. First, we observed an increase in cell-cell interaction and cell clustering in 2D cultures with loss of GPR52 in MCF10A, MDA-MB-468, and MDA-MB-231 cell lines. TEM demonstrated that GPR52 loss in MDA-MB-468 is associated with a reduction in the length of the interface between two cells, but that GPR52 KO cells are more juxtaposed with one another than WT cells, which exhibit more intercellular space along their interface. Multicellular aggregation and increased cell-cell adhesion are mechanisms by which cancer cells can increase their metastatic potential (2). Moreover, the increase in disorganization and number of MDA-MB-468 spheroids with GPR52 loss are two hallmarks of cancer progression, with the latter suggesting a potential increase in stemness due to increased propensity to survive and proliferate in Matrigel (33). MDA-MB-231 cells did not demonstrate a change in the sphericity or roundness of the spheroids with GPR52 loss but did form smaller spheroids that tended to be more densely packed with cells, demonstrating that GPR52 loss also affects 3D multicellular organization in this cell line.

Furthermore, collective organization and invasion provide mechanisms for cancer cells to transmit survival signals and invade directionally, in some cases featuring a distinct leading front of cells that tend to be more mesenchymal and invasive (34). We also found that loss of GPR52 affected the pattern and extent to which cancer cells degraded the ECM. A computational model previously suggest that collective invasion requires less ECM digestion than single cell invasion (35). Linearization and alignment of collagen tracts have been hypothesized to promote the directed migration of cancer cells, and alignment of ECM fibers has been associated with a reduction in proteolytic degradation of the ECM (35, 36).

Many upstream regulators of EMT have been identified; however, induction of EMT or partial EMT with loss of a GPCR has not been previously described to our knowledge. Importantly, partial EMT has been documented in circulating tumor cells (37). The exclusivity of our observations of cell elongation and the development of a mesenchymal morphology on PDL- but not TC-treated plastic may suggest that integrin

activity can influence the extent of EMT in this cell line and is an interesting area for further exploration.

Our finding that GPR52 loss was associated with an increase in total cancer area in zebrafish at 30 hpi suggests increased survival and/or growth of GPR52 KO MDA-MB-468 cells. Our data also suggests reduction in sensitivity of cancer cells in the head to doxorubicin with GPR52 loss. Of note, the blood-brain barrier starts to develop at 3 days post-fertilization (dpf) in zebrafish (38). As doxorubicin has limited capacity to cross the blood-brain barrier, the efficacy of this chemotherapeutic has historically been limited to its effects outside of the CNS (39). However, blood-brain barrier permeable forms of doxorubicin have been developed and may have notable utility against head or brain metastases for certain molecular subtypes of breast cancer as demonstrated in this study (39).

We identified an increased cAMP signaling signature in GPR52 KO groups based on RNA-sequencing and proteomic studies and found that modulation of cAMP levels could induce or attenuate some, but not all, phenotypes associated with GPR52 loss. Increased intracellular cAMP levels have previously been associated with increased proliferation in 3D culture, increased invasiveness of breast cancer, and deposition of ECM components (13, 40). However, the role of cAMP in regulating the pattern of ECM degradation has not been previously characterized. Therefore, these mechanistic studies not only provide some insight into how GPR52 loss may influence breast cancer cell biology but also identify a role of cAMP in regulating breast cancer cell adhesion and ECM digestion. Treatment of HEK293 cells with FTBMT led to an increase in intracellular cAMP, as has been previously reported, but MDA-MB-468 cells demonstrated a lack of change or possible decrease in intracellular cAMP with FTBMT treatment and no change in calcium levels (41). The coupling of GPCRs to G proteins can be promiscuous, and one study estimates that 73% of GPCRs can activate multiple G proteins (42). Therefore, the interaction of GPR52 with different G proteins, particularly in different cell types, warrants further investigation.

Furthermore, we found that the melanoma cell adhesion molecule (MCAM), which is considered a potential biomarker and promoter of progression of many cancers, was increased in MCF10A and MDA-MB-468 cells, and that its expression could be reduced in GPR52 KO cells with adenylyl cyclase inhibitor treatment (43, 44). Thus, we identified a mechanism by which MCAM can be upregulated in breast cancer and a method to reduce expression of this driver of cancer aggression.

Our studies provide rationale for investigating a therapeutic effect of GPR52 agonism in breast cancer and encourage the investigation of the role of GPR52 in the progression of additional cancer types. Furthermore, our work identifies novel features of cell biology than can be regulated by a GPCR and broadens the scope of significance of this class of proteins in cancer biology.

Acknowledgements

This work was supported by the National Cancer Institute of the National Institutes of Health grant 1R01CA215797 (KAB), the Anne Moore Breast Cancer Research Fund (KAB), and the Emilie Lippmann and Janice Jacobs McCarthy Research Scholar Award in Breast Cancer (KAB). Sarah Hanif was supported by a Medical Scientist Training Program grant from the National Institute of General Medical Sciences of the National Institutes of Health under award number: T32GM152349 to the Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program. The content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Studies were conducted with the support and facilities provided by the Microscopy and Image Analysis Core Facility and the Genomics Resources Core Facility at Weill Cornell Medicine. All schematics were created in BioRender.

Materials and Methods

Cell culture

Human breast cell lines MDA-MB-468 (RRID:CVCL_0419), MDA-MB-231 (RRID:CVCL_0062), MCF7 (RRID:CVCL_0031), HS578T (CVCL_0332), T47D (CVCL_0553), and MCF10A (RRID:CVCL_0598) were purchased from the American Type Culture Collection (ATCC). MDA-MB-468, MDA-MB-231, MCF7, and HS578T were cultured in Dulbecco's Modified Eagle's Medium (DMEM, Gibco #11995-065) supplemented with 10% fetal bovine serum (FBS) (Gibco 26140079, Sigma-Aldrich F0926) and 1% penicillin-streptomycin (i.e. complete media). T47D cells were cultured in Roswell Park Memorial Institute (RPMI 1640, Thermo Fisher Scientific #11875) supplemented with 10% fetal bovine serum (FBS) (Gibco 26140079) and 1% penicillin-streptomycin (i.e. complete media). MCF10A cells were cultured in Dulbecco's Modified Eagle's Medium/F12 (Gibco, #11330-032) supplemented with 5% FBS, 1% penicillin-streptomycin, and the following growth factors: epidermal growth factor (20 ng/ml), hydrocortisone (0.5 mg/ml), cholera toxin (100 ng/ml), and insulin (10 µg/ml) (i.e. complete media). All cell lines were maintained at 37°C in a humidified atmosphere with 5% CO_2. All cell lines were authenticated by STR analysis and tested regularly for mycoplasma contamination using the Universal Mycoplasma Detection Kit (ATCC 30-1012K) and the MycoAlert® Mycoplasma Detection Kit (Lonza LT07-318).

Breast cancer survival curves

KM plotter RNA-seq breast cancer web tool (www.kmplot.com) (12) was used to obtain Kaplan-Meier survival curves. The associated hazard ratios, 95% confidence intervals, and logrank P-value were automatically determined by the webtool. Low versus high cutoff was determined based on the maximum segregation between the groups.

GPR52 and MCAM expression comparison in cancerous and non-cancerous human tissue

GPR52 mRNA expression in cancerous and non-cancerous tissues was analyzed across solid cancer types using RNA-seq datasets from the TNMplot database and web tool (https://www.tnmplot.com) (11). The following datasets were analyzed for this

analysis: adrenocortical carcinoma (“adrenal”), bladder urothelial carcinoma (“bladder”), breast invasive carcinoma (“breast”), colon adenocarcinoma (“colon”), esophageal carcinoma (“esophageal”), kidney renal clear cell carcinoma (“renal_cc”), kidney renal papillary cell carcinoma (“renal_pa”), liver hepatocellular carcinoma (“liver”), lung adenocarcinoma (“lung_ac”), lung squamous cell carcinoma (“lung_sc”), ovarian serous cystadenocarcinoma (“ovary”), pancreatic adenocarcinoma (“pancrease”), prostate adenocarcinoma (“prostate”), rectum adenocarcinoma (“rectum”), skin cutaneous melanoma (“skin”), stomach adenocarcinoma (“stomach”), testicular germ cell adenocarcinoma (“testis”), thyroid carcinoma (“thyroid”), uterus corpus endometrial carcinoma (“uterus_ec”). The comparison of GPR52 expression between normal, tumor, and metastatic breast-derived samples was determined based on a gene chip dataset available through TNM plot. Source data for GPR52 expression was collected in January 2024 and the correlation between GPR52 and MCAM expression in breast tumors was determined in June 2024.

RT-qPCR

Each breast cancer cell line was cultured in an individual well of a 6-well tissue culture-treated plate until 50-80% confluent. Total RNA was extracted using the RNeasy Mini Kit (Qiagen # 74106) and RNA purity was assessed with the Nanodrop 2000 instrument (Thermo Fisher Scientific). RNA was reverse transcribed using the qScript cDNA Synthesis kit (VWR # 101414-100) based on the manufacturer’s protocol. Quantitative polymerase chain reaction (PCR) was performed in triplicate from the RNA collected from each well using the Fast SYBR Green Master Mix (Thermo Fisher Scientific #4385612) following standard protocol. Quantification of GPR52 and RPL32 transcripts was carried out using the following primers:

GPR52 F: 5'- CGGGTCTTGGACAATCCAACTC -3',
GPR52 R: 5' - TGCTTCCTGATCCTTCACACAC - 3',
RPL32 F: 5' – CAGGGTTCGTAGAAGATTCA -3',
RPL32 R: 5' – CTTGGAGGAAACATTGTGAGCGATC -3'.

GPR52 expression for each human breast cell line was normalized to the housekeeping gene RPL32 and then normalized to the average expression of GPR52 across the human breast cell lines analyzed.

Generation of stable cell lines

We used transduction-based CRISPR-Cas9 gene editing to generate indels in *GPR52* in MDA-MB-468, MDA-MB-231, and MCF10A cells. For lentivirus production, HEK293 cells were plated in a six-well plate and transfected with a prepared mix in 150uL DMEM (with no supplements) containing 2.5 µg of pLenti-Cas9-P2A-Puro (a gift from Lukas Dow; RRID:Addgene_110837), 1.25 µg of PAX2, 1.25 µg of VSV-G, and 30 µl of polyethylenimine (1 mg/ml). The media was replaced the following day and changed 36 hours post-transfection to target cell collection medium. MDA-MB-468, MDA-MB-231, and MCF10A cells were plated in individual wells of a six-well plate and transduced with the pLenti-Cas9-P2A-Puro lentivirus generated above in serum- and antibiotic-free

media containing 8 μg/mL polybrene for 24 hours, then selected with 1-2.5 μg/mL puromycin in complete media (45).

The forward and reverse sgRNA cloning primers were as follows:

sgRNA1-F: CACCGCAAAACCATGGCGTAGCGA
sgRNA1-R: AAACTCGCTACGCCATGGTTTTGC
sgRNA2-F: CACCGTGAATGGTGTGCCACGTCT
sgRNA2-R: AAACAGACGTGGCACACCATTCAC

These primers were annealed and cloned following standard procedures using the BsmBI/EcoRI site of the pLenti-U6-sgRNA-tdTomato-P2A-BlasR (LRT2B) lentiviral vector (a gift from Lukas Dow; RRID:Addgene_110854) (45). MDA-MB-468, MDA-MB-231, and MCF10A cells stably expressing the pLenti-Cas9-P2A-Puro lentiviral vector were then transduced with the GPR52 sgRNA lentiviral vectors in serum- and antibiotic-free media containing 8 μg/mL polybrene for 24 hours and underwent blasticidin selection followed by limiting dilution assays to isolate single cells from the sgRNA transduced populations. The cells were then expanded in culture to generate clonal populations, and a subset of these cells was collected by centrifugation for DNA extraction and Sanger sequencing of the GPR52 sgRNA target region. Briefly, DNA was extracted from a cell pellet with the QiaAMP DNA Mini kit (Qiagen #56304) based on the manufacturer's protocol. The GPR52 guide target regions were then amplified by PCR using the following primers and purified with the QIAquick PCR Purification Kit (Qiagen #28104).

GPR52-1 F: 5'- AAACCTGGTTACCATGGTGAC -3',
GPR52-2 F: 5' - GATCTGGATCTACTCCTGCC - 3',
GPR52-1,-2 R: 5'- ATTATATAGGGGAGCCACAGC-3',

The amplified PCR products were submitted for Sanger sequencing (Genewiz, Azenta) to identify clones with indels in the GPR52 sgRNA target regions (Supplementary Fig. S2). The clonal populations generated from each sgRNA were expanded and utilized for downstream studies. A heterogeneous population of gSafe cells was used to represent the vector control population.

Transmission electron microscopy

MDA-MB-468 cells were cultured in suspension in non-tissue culture treated plates for 24-48 hours and then briefly centrifuged in 1.7mL tubes. The resulting pellet was washed then fixed with 4% paraformaldehyde, 2.5% glutaraldehyde, 0.02% picric acid in 0.1 M sodium cacodylate buffer, pH 7.3. The samples were then stained with uranyl acetate and dehydrated with a graded ethanol series. After dehydration, samples were covered by a layer of fresh resin, and embedding molds were inserted into the resin. After polymerization, samples were cut at 200 nm for screening by light microscopy and then at 65 nm to be mounted on grids for TEM under a JEOL JSM 1400 (JEOL, USA) electron microscope. The camera used is a Veleta, 2 K × 2 K CCD (EMSIS, GmbH, Muenster). Images were taken under 2500-20,000× magnification, and each field included extracellular space, cytoplasmic area, and nucleus. Cell-cell interface length was determined using 2500x images by drawing a straight line that extended the distance of cell-cell interface only between cells that exhibited extensive adhesion between their plasma membrane surfaces. Cell diameter was determined using 2500x images by drawing a straight line that spanned the midpoint of the aforementioned

analyzed cells. The free intercellular space was determined with Fiji software. An image threshold intensity was set to distinguish cells from the (white) background. The boundary between two cells was then outlined and the fraction of the area occupied by free space (met the threshold intensity) was determined.

3D Matrigel cell culture and analysis

MDA-MB-468, MDA-MB-231 and MCF10A cells were respectively resuspended in 1:1, 1:1, and 1:2 growth-factor-reduced Matrigel:serum-free DMEM and plated at a density of 10,000 cells/100 µL per well of a black optical-bottom 96-well plate (Corning Matrigel #356231). 100 µL of complete media was added to each well after solidification of the Matrigel, and media was replaced every 2-3 days. After 10 days (MDA-MB-468, -231) or 3 weeks (MCF10A) in culture, the cells were fixed with 10% formalin and stained with Hoechst 33342 (Santa Cruz Biotechnology # SC-495790). Cells were imaged with confocal microscopy (Zeiss LSM880) to detect cytoplasmic tdTomato and nuclear Hoechst signals. Images were then imported to Imaris Microscopy Image Analysis Software (OXFORD instruments). The Imaris "Surface" function was used to determine the area per spheroid, total area occupied per well, and the roundness of each spheroid based on the absolute intensity of tdTomato signal. For MDA-MB-468, the following formula for circularity was used to determine the roundness of each spheroid:

$$Circularity = \frac{4\pi \times Area}{Perimeter^2}$$

For MDA-MB-231, as images were obtained as z-stacks, the following formula for sphericity rather than circularity was used to determine roundness of spheroids:

$$\psi = \frac{\pi^{\frac{1}{3}}(6V_p)^{\frac{2}{3}}}{A_p}$$

V_p=spheroid volume, A_p=spheroid area

For MDA-MB-468, the total surface number calculated by Imaris was used to determine the total number of spheroids. For MDA-MB-231, the total number of spheroids was determined manually, as spheroids were closer together and could not be distinguished by the software. The area per spheroid for MDA-MB-468 was automatically determined by the software as the average area per surface, whereas for MDA-MB-231 the average spheroid area was determined by dividing the total tdTomato area by the number of manually-determined spheroids. A minimal fluorescence threshold was maintained to maximize the number of detected surfaces while eliminating background fluorescence. A final area cutoff of ≥1000 μm^2 was set for consideration of a structure for the analyses described above. This threshold was set in order to reduce detection of single cells or spheroids on other z-planes.

Matrigel invasion assays

MDA-MB-468 and MDA-MB-231 cells were plated at >80% confluency in optical-grade 96-well plates. A 50 µL solution of 1:1 growth-factor-reduced Matrigel:serum-free DMEM ± 1:40 DQ-Collagen IV (Thermo Fisher Scientific) was added above the monolayer and allowed to form a gel. 50 µL of DMEM containing 10% FBS was added above the gel once it had solidified. After 24 hours, confocal microscopy (Zeiss LSM880) was used to obtain a z-stack of tdTomato and DQ signal in each well at 10 µm intervals starting from

the base of the well. The Imaris “Surface” function was used to determine the surface area of tdTomato foci above a minimum fluorescence intensity threshold at established z-planes. For collective invasion analyses and class B determination at z=10um, the threshold intensity was auto-determined by Imaris per image; surfaces were identified based on local contrast of tdTomato signal with the non-fluorescent background. The Imaris “Surface” function was used to quantify the area of DQ at z=10 µm based on a maintained absolute value threshold across samples and co-localize the volume of DQ with the volume of tdTomato. that was adjusted to maximize detection across samples based on local contrast at z=10 µm.

Poly-D-lysine coating

PDL (Santa Cruz Biotechnology #SC-136156) was resuspended in milli-Q water at 1mg/mL and frozen in aliquots at -20°C. To develop coating solutions, PDL was resuspended in milli-Q water to a final concentration of 100 µg/mL and added to tissue-culture treated plates to cover the surface. Plates were then incubated for at least 30 minutes at room temperature or several hours at 4°C. Plates were then washed twice with PBS and used for cell culture.

Western blotting and analysis

Cells were cultured in monolayer in standard tissue-culture conditions as described above. At the time of collection, cells were scraped on ice with ice-cold PBS, centrifuged at 1500-2000 rpm, and the resulting cell pellet was snap-frozen in liquid nitrogen. On the day of western blotting, the cell pellet was lysed in ice-cold buffer (5 mmol/L HEPES, 137 mmol/L NaCl, 1 mmol/L $MgCl_2$, 1 mmol/L $CaCl_2$, 10 mmol/L NaF, 2 mmol/L EDTA, 10 mmol/L sodium pyrophosphate, 2 mmol/L $NaVO_4$, 1% NP-40, 10% glycerol) containing protease inhibitors (Thermo Fisher Scientific #78429), incubated on a rotating shaker at 4°C for 15 minutes, and centrifuged for 20 minutes at 4°C at 20,000 rpm. Cell extracts were denatured in buffer containing β-mercaptoethanol, run on NuPAGE 4–12% Bis-Tris protein gels (Thermo Fisher Scientific), and then transferred to nitrocellulose membranes. Membranes were blocked with 5% nonfat dry milk (Bio-Rad, #1706404) for one hour at room temperature (RT), washed with 1X Tris-Buffered Saline, 0.1% Tween (TBST), then incubated with primary antibodies in 4°C for at least 16 hours. The following primary antibodies were used: E-cadherin (RRID:AB_2291471), Snai1 (RRID:AB_2255011), Vimentin (RRID:AB_10695459), p-CREB (RRID:AB_2561044), CREB (RRID:AB_310268), and MCAM (RRID:AB_2143373). Membranes were then washed with TBST and incubated with secondary antibodies for 1-2 hours at room temperature. The following secondary antibodies were used: Anti-rabbit IgG HRP-linked Antibody (RRID:AB_2099233) and Anti-mouse IgG HRP-linked Antibody (RRID:AB_330924). Western Lightning Plus-ECL (Thermo Fisher Scientific #509049325) and ImageLab software (BioRad) were used for band detection. Membranes were stripped with Restore PLUS Western blot stripping buffer (Thermo-Fisher Scientific, #46430) for 15 minutes and then probed for β-actin expression as a loading control using an HRP-linked β-actin antibody (RRID:AB_262011).

Zebrafish maintenance

All animal work was performed under an approved IACUC protocol (2011-0100) from Weill Cornell Medicine with animal care under supervision of the Research Animal Resource Center. The TG(flk1:EGFP-NLS) zebrafish strain was generated by the Markus Affolter laboratory (Biozentrum, University of Basel, Switzerland) (Developmental Biology 316: 312-322, 2008) and kindly provided by Jesus Torres-Vazquez (New York University, New York, NY). Embryos were obtained by natural matings and raised at 28.5 °C with a 14-h light–10-h dark cycle. Zebrafish were maintained in E3 water (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$, 1 ppm methylene blue).

Zebrafish xenograft studies

tdTomato-tagged MDA-MB-468 cells were passed through a 40 µm filter and resuspended in PBS at a density of 400 cells/5nL. Two dpf zebrafish embryos were each injected with 5nL of the cancer cells in the perivitelline space or the duct of Cuvier, whichever was best accessible based on the orientation of the fish. Injected embryos were evaluated at two hpi and only zebrafish with tdTomato detectable distal to the injection site were maintained at 32 ^{o}C. Zebrafish were anesthetized with 0.02% tricaine and imaged at five hpi and 30 hpi with a Nikon SMZ1500 microscope mounted with a Nikon DS-FI3 camera. Cancer surface area per zebrafish was quantified using Imaris "Surface" function; two rectangles were drawn per zebrafish to encompass the area distal to the injection site excluding the yolk sac (i.e., trunk) and superior to the otolith (i.e., head), and the total surface area of tdTomato determined in each rectangle was combined to calculate the total cancer area. The number of tdTomato surfaces detected was used to determine the number of cancer foci. Zebrafish were fixed with 4% paraformaldehyde at study endpoint. For confocal imaging, fixed zebrafish were washed with PBS and then incubated for at least 12 hours at 4°C in PBS containing Hoechst 33342. Zebrafish were then mounted in 1% low-melt agarose and imaged using confocal microscopy (Zeiss LSM880).

For drug treatment studies, zebrafish with tdTomato detectable distal to the yolk sac were selected at two hpi, placed into individual wells of 24-well plates, and imaged at 5hpi. Zebrafish were then maintained in E3 water containing either 8 µM doxorubicin or the vehicle (milli-Q water). All zebrafish were imaged at 30 hpi and analyzed as above.

RNA-sequencing studies and computational analyses

MDA-MB-468 WT (parental) and GPR52 sgRNA1 KO cells (one mixed population and two clonal populations) were cultured in 6-well plates. Total RNA was extracted using QIAzol lysis reagent (Qiagen #79306) and the RNeasy Mini Kit (Qiagen #74106). Samples were submitted for RNA-sequencing at the Genomics Resources Core Facility (GRCF, Weill Cornell Medicine). Total RNA integrity was assessed with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). RNA concentrations were measured using the NanoDrop system (Thermo Fisher Scientific, Inc., Waltham, MA). RNA sample library and RNA-sequencing were performed by the Genomics Core Laboratory at Weill Cornell Medicine with Illumina TruSeq Stranded mRNA Sample Library Preparation kit (Illumina, San Diego, CA), according to the manufacturer's instructions. The normalized cDNA libraries were pooled and sequenced on Illumina NovaSeq 6000 sequencer with pair-end 50 cycles. The raw sequencing reads in BCL

format were processed through bcl2fastq 2.19 (Illumina) for FASTQ conversion and demultiplexing.

All reads were independently aligned with STAR_2.4.0f1 for sequence alignment against the human genome sequence build hg19, downloaded using the UCSC genome browser and SAMTOOLS v0.1.19 for sorting and indexing reads. Cufflinks (2.0.2) was used to estimate the expression values (FPKMS), and GENCODE v19 GTF file for annotation. The gene counts from htseq-count and DESeq2 Bioconductor package were used to identify DEGs. All DEGs with P-value <0.05 and false discovery rate <0.05 were uploaded to Ingenuity Pathway Analysis software (Qiagen) for further analysis.

Proteomic studies and computational analyses

Parental, vector control, and GPR52-1 and -2 KO MDA-MB-468, MDA-MB-231, and MCF10A cells were cultured in 10-cm tissue-cultured treated plates until 60-80% confluent. Cells were then scraped on ice with ice-cold PBS, centrifuged, and the resulting cell pellet was snap-frozen in liquid nitrogen. Cells were lysed by resuspension in Radio-Immunoprecipitation Assay (RIPA) supplemented with protease and phosphatase inhibitors. Soluble proteins were reduced by the addition of 0.5 M TCEP to a final concentration of 5mM followed by incubation at 55 ^{o}C for 30 minutes. Reduced samples were alkylated with the addition of 375 mM iodoacetamide to a final concentration of 10 mM followed by incubation in the dark at room temperature for 30 minutes. Ice cold acetone was added to each sample to a volume ratio of 5:1. Samples were vortexed and stored at -20 ^{o}C overnight. After precipitation, samples were centrifuged at 14,000 x *g* at 4 ^{o}C for 30 minutes to pellet the proteins. The supernatant was removed and the pellet was air dried on benchtop for 10 minutes. The proteins were resuspended in 50 mM TEAB pH 8 with 2 mM $CaCl_2$. Trypsin was added (500 ng) and the proteins were allowed to digest overnight at 37 ^{o}C with shaking at 500 RPM (Thermomixer, Eppendorf). The digestion was quenched with the addition of 10% formic acid to a final concentration of 1%. Digested samples were centrifuged at 10,000 x *g* for 10 minutes to remove particulates and the supernatant was transferred to a fresh tube and stored at -20 ^{o}C until phosphopeptide enrichment.

Samples were enriched for phosphorylated peptides using the SMOAC method. Briefly, digested samples were first enriched by the High Select™ Phosphopeptide Enrichment kit (Thermo Fisher Scientific) following the manufacturers protocol. The flow-through was applied to the High Select™ Fe-NTA Phosphopeptide Enrichment Kit (Thermo Fisher Scientific) following the manufacturers protocol. The flow through after the second enrichment became the global, unenriched sample and the elutes from both kits were pooled to generate the phosphopeptide enriched fraction. Peptide concentration was measured using a Nanodrop spectrophotometer (Thermo Scientific) at 205 nm.

Samples were injected using the Vanquish Neo (Thermo Fisher Scientific) nano-UPLC onto a C18 trap column (0.3 mm x 5 mm, 5 µm C18) using pressure loading. Peptides were eluted onto the separation column (PepMap™ Neo, 75 µm x 150 mm, 2 µm C18 particle size, Thermo Fisher Scientific) prior to elution directly to the mass spectrometer. Briefly, peptides were loaded and washed for 5 minutes at a flow rate of

0.350 µL/minutes at 2% B (mobile phase A: 0.1% formic acid in water, mobile phase B: 80% ACN, 0.1% formic acid in water). Peptides were eluted over 100 minutes from 2-25% mobile phase B before ramping to 40% B in 20 min. The column was washed for 15 minutes at 100% B before re-equilibrating at 2% B for the next injection. The nano-LC was directly interfaced with the Orbitrap Ascend Tribrid mass spectrometer (Thermo Fisher Scientific) using a silica emitter (20 µm i.d., 10 cm) equipped with a high field asymmetric ion mobility spectrometry (FAIMS) source. The data were collected by data dependent acquisition with the intact peptide detected in the Orbitrap at 120,000 resolving power from 375-1500 *m/z*. Peptides with charge +2-7 were selected for fragmentation by higher energy collision dissociation (HCD) at 28% NCE and were detected in the ion trap using rapid scan rate (global) or the Orbitrap at resolving power 30,000 (enriched). Dynamic exclusion was set to 60s after one instance. The mass list was shared between the FAIMS compensation voltages. FAIMS voltages were set at -45 (1.4 s), -60 (1 s), -75 (0.6 s) CV for a total duty cycle time of 3s. Source ionization was set at 1700 V with the ion transfer tube temperature of 305 ^{o}C. Raw files were searched against the human protein database downloaded from Uniprot on 05-05-2023 using SEQUEST in Proteome Discoverer 3.0. Abundances, abundance ratios, and P-values were exported to Microsoft Excel for further analysis. All proteins with differential GPR52 KO vs WT abundance ratios (P-value <0.05 and false discovery rate <0.05) were uploaded to Ingenuity Pathway Analysis software (Qiagen) for further analysis for each of the three cell lines. KEA3 software was used to rank kinase activity of the combined vector control and GPR52 KO populations across the three cell lines based on the respective phophoproteomes (20).

Analysis of TCGA-BRCA transcriptomic dataset

The Cancer Genome Atlas TCGA-BRCA database was used to obtain primary breast tumor gene expression data from 142 breast cancer patients. The downloaded data displayed gene ENSEMBL ID and FPKM-UQ normalized expression counts. Four patients were found to have mutations in the *Gpr52* gene and were not included in downstream analyses. The data for the remaining 138 patients were segregated based on “zero” and “non-zero” expression of GPR52 and then subsequently read into edgeR, a software package available from BiocManager for differential expression analysis of RNA-sequencing data (46). Briefly, dispersion was estimated to measure inter-library variability. Data were then fitted into a generalized linear regression model. Statistical significance testing (likelihood ratio test) between the cohorts was performed using the fitted model to compare differential gene expression between “zero” and “non-zero data”, creating the dataset “Non-zero/Zero”. All DEGs with P-value <0.05 and false discovery rate <0.05 were uploaded to Ingenuity Pathway Analysis software (Qiagen) for further analysis.

BRET biosensor studies

HEK 293 and MDA-MB-468 cells were maintained in DMEM with 10% FBS, respectively, and cultured at 37°C and 5% CO_2. For transfection, cells were plated at a density of 2.5×10^5 cells (HEK293) and 8×10^5 cells (MDA-MB-468) per well in 6-well plates (Thermo Scientific, 140,675). On the day of transfection, the media was replaced with DMEM containing 2.5% FBS (HEK 293) and Opti-MEM (MDA-MB-468) (Thermo

Fisher Scientific #31985062) withno antibiotics. HEK293 cells were transfected with Lipofectamine 2000 (Thermo Fisher Scientific #11668027). The GFP10-Epac-RlucII in pcDNA3.1 (generously provided by Dr. Michel Bouvier, Université de Montréal) and GPR52 cDNA ORF Clone GPR52 (Sino Biological # HG25891-UT) in pcDNA3.1 with added N-terminal flag tag were used in these studies:

> For HEK293 cells, per well: 1.5 µg of plasmid DNA (total) was added to 100 µl DMEM (no serum, no antibiotics) in one 1.5 mL Eppendorf tube. In another tube: 3 µl of Lipofectamine 2000 was added to 100 µl of DMEM (no serum, no antiobiotics).
>
> For MDA-MB-468 cells, per well: 2.5 µg of plasmid DNA (total) was added to 100 µl Opti-MEM (no serum, no antibiotics) in one 1.5 mL Eppendorf tube. In another tube: 5 µl of Lipofectamine 2000 or 3000 was added to 100 µl of Opti-MEM (no serum, no antibiotics).

After 5 minutes, the contents of the tubes were combined and gently mixed or vortexed for each cell line. The mixture was then incubated for 20-30 minutes at room temperature. The 200 µL DNA:Lipofectamine mixture was then added dropwise per well and the plate was swirled gently and incubated at 37 °C and 5% CO_2 for 4-5 hours. The media was then replaced with complete media.

Cells were detached the day following transfection with 0.25% trypsin–EDTA (Wisent) and plated at a density of 3×10^4 cells/well in a poly-l-ornithine (Sigma-Aldrich)-coated 96-well white bottom plate (Thermo Scientific, 236,105) for BRET analysis. After 24 hours incubation in the 96-well plate, the media was removed and cells were washed once with Krebs buffer (146 mM NaCl, 42 mM MgCl2, 10 mM HEPES pH 7.4, 1 g/L D-glucose) then incubated for 2 hours at room temperature in Krebs buffer. Coelenterazine 400A (Cedarlane) was added, to a final concentration of 2.5 µM, to each well and incubated for 5 minutes prior to basal reading. FTBMT was added to the indicated final concentration and the plate was read after a 15 or 30 minute stimulation. For MDA-MB-468, cells were co-stimulated with 1 µM forskolin.

BRET ratios were calculated as the emission at 515 nm/emission at 410 nm. For all experiments, ΔBRET = (BRET ratio from stimulated cells with FTBMT and forskolin) – (BRET ratio from vehicle treated with forskolin). Three technical replicates were used for all treatments. BRET experiments were performed using a Tristar2 plate reader (Berthold. Technologies GmbH & Co. KG). The normalized BRET ratio was computed as $BRET_{Stimuated}/BRET_{Vehicle}$. Dose response curves were plotted using non-linear regression. Data is represented as mean ± standard error (SE). For MDA-MB-468 studies, the experiment was conducted five times and the aggregate data from all experiments is presented. For HEK 293 cells, the experiment was conducted twice, and data combined from the two experiments is presented.

Intracellular calcium measurements

MDA-MB-468 cells were plated at a density of 5000 cells/well in a 96-well plate. After 24 hours, the cell-permeable calcium indicator Fluo-4 (Thermo Fisher Scientific #F14201) was added to each well. Cells were incubated for 45 minutes at 37 °C, then for 15 minutes at room temperature. Baseline fluorescence measurements were obtained (t=0) and FTBMT was then added to select wells. The plate was read at 30 second intervals for 10 minutes. Thapsigargin (Sigma-Aldrich) was added to cells as a positive control, and the plate was then read at 30 second intervals for an additional 10 minutes (47).

Drug preparation

Forskolin (Santa Cruz Biotechnology #SC-3562) and SQ22536 (Santa Cruz Biotechnology #SC-201572) were resuspended in ethanol at 1 mM and DMSO at 14mM, respectively, and frozen. SQ22536 was aliquoted to avoid freeze-thaw cycles. Ethanol and DMSO were used as vehicle controls in all experiments that incorporated these compounds.

Statistical analysis

All data are expressed as individual datapoints with a line at the median value or as mean ± SEM, unless indicated otherwise. Groups were compared using statistical tests for significance as described in the figure legends. A P value <0.05 was considered statistically significant. Statistical tests were performed with GraphPad Prism.

The data generated in this study are available upon request from the corresponding author.

References

1. Hagemeister, F. B., Jr., Buzdar, A. U., Luna, M. A., and Blumenschein, G. R. (1980) Causes of death in breast cancer: a clinicopathologic study. *Cancer* **46**, 162-167
2. Schuster, E., Taftaf, R., Reduzzi, C., Albert, M. K., Romero-Calvo, I., and Liu, H. (2021) Better together: circulating tumor cell clustering in metastatic cancer. *Trends Cancer* **7**, 1020-1032
3. Bakir, B., Chiarella, A. M., Pitarresi, J. R., and Rustgi, A. K. (2020) EMT, MET, Plasticity, and Tumor Metastasis. *Trends Cell Biol* **30**, 764-776
4. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu,

F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., Yao, A., Ye, J., Zhan, M., Zhang, W., Zhang, H., Zhao, Q., Zheng, L., Zhong, F., Zhong, W., Zhu, S., Zhao, S., Gilbert, D., Baumhueter, S., Spier, G., Carter, C., Cravchik, A., Woodage, T., Ali, F., An, H., Awe, A., Baldwin, D., Baden, H., Barnstead, M., Barrow, I., Beeson, K., Busam, D., Carver, A., Center, A., Cheng, M. L., Curry, L., Danaher, S., Davenport, L., Desilets, R., Dietz, S., Dodson, K., Doup, L., Ferriera, S., Garg, N., Gluecksmann, A., Hart, B., Haynes, J., Haynes, C., Heiner, C., Hladun, S., Hostin, D., Houck, J., Howland, T., Ibegwam, C., Johnson, J., Kalush, F., Kline, L., Koduru, S., Love, A., Mann, F., May, D., McCawley, S., McIntosh, T., McMullen, I., Moy, M., Moy, L., Murphy, B., Nelson, K., Pfannkoch, C., Pratts, E., Puri, V., Qureshi, H., Reardon, M., Rodriguez, R., Rogers, Y. H., Romblad, D., Ruhfel, B., Scott, R., Sitter, C., Smallwood, M., Stewart, E., Strong, R., Suh, E., Thomas, R., Tint, N. N., Tse, S., Vech, C., Wang, G., Wetter, J., Williams, S., Williams, M., Windsor, S., Winn-Deen, E., Wolfe, K., Zaveri, J., Zaveri, K., Abril, J. F., Guigo, R., Campbell, M. J., Sjolander, K. V., Karlak, B., Kejariwal, A., Mi, H., Lazareva, B., Hatton, T., Narechania, A., Diemer, K., Muruganujan, A., Guo, N., Sato, S., Bafna, V., Istrail, S., Lippert, R., Schwartz, R., Walenz, B., Yooseph, S., Allen, D., Basu, A., Baxendale, J., Blick, L., Caminha, M., Carnes-Stine, J., Caulk, P., Chiang, Y. H., Coyne, M., Dahlke, C., Deslattes Mays, A., Dombroski, M., Donnelly, M., Ely, D., Esparham, S., Fosler, C., Gire, H., Glanowski, S., Glasser, K., Glodek, A., Gorokhov, M., Graham, K., Gropman, B., Harris, M., Heil, J., Henderson, S., Hoover, J., Jennings, D., Jordan, C., Jordan, J., Kasha, J., Kagan, L., Kraft, C., Levitsky, A., Lewis, M., Liu, X., Lopez, J., Ma, D., Majoros, W., McDaniel, J., Murphy, S., Newman, M., Nguyen, T., Nguyen, N., Nodell, M., Pan, S., Peck, J., Peterson, M., Rowe, W., Sanders, R., Scott, J., Simpson, M., Smith, T., Sprague, A., Stockwell, T., Turner, R., Venter, E., Wang, M., Wen, M., Wu, D., Wu, M., Xia, A., Zandieh, A., and Zhu, X. (2001) The sequence of the human genome. *Science* **291**, 1304-1351

5. Hanson, M. A., and Stevens, R. C. (2009) Discovery of new GPCR biology: one receptor structure at a time. *Structure* **17**, 8-14
6. Hamm, H. E., and Gilchrist, A. (1996) Heterotrimeric G proteins. *Curr Opin Cell Biol* **8**, 189-196
7. Chaudhary, P. K., and Kim, S. (2021) An Insight into GPCR and G-Proteins as Cancer Drivers. *Cells* **10**
8. Komatsu, H., Maruyama, M., Yao, S., Shinohara, T., Sakuma, K., Imaichi, S., Chikatsu, T., Kuniyeda, K., Siu, F. K., Peng, L. S., Zhuo, K., Mun, L. S., Han, T. M., Matsumoto, Y., Hashimoto, T., Miyajima, N., Itoh, Y., Ogi, K., Habata, Y., and Mori, M. (2014) Anatomical transcriptome of G protein-coupled receptors leads to the identification of a novel therapeutic candidate GPR52 for psychiatric disorders. *PLoS One* **9**, e90134
9. Nishiyama, K., Suzuki, H., Harasawa, T., Suzuki, N., Kurimoto, E., Kawai, T., Maruyama, M., Komatsu, H., Sakuma, K., Shimizu, Y., and Shimojo, M. (2017) FTBMT, a Novel and Selective GPR52 Agonist, Demonstrates Antipsychotic-Like and

Procognitive Effects in Rodents, Revealing a Potential Therapeutic Agent for Schizophrenia. *J Pharmacol Exp Ther* **363**, 253-264
10. Yao, Y., Cui, X., Al-Ramahi, I., Sun, X., Li, B., Hou, J., Difiglia, M., Palacino, J., Wu, Z. Y., Ma, L., Botas, J., and Lu, B. (2015) A striatal-enriched intronic GPCR modulates huntingtin levels and toxicity. *Elife* **4**
11. Bartha, A., and Gyorffy, B. (2021) TNMplot.com: A Web Tool for the Comparison of Gene Expression in Normal, Tumor and Metastatic Tissues. *Int J Mol Sci* **22**
12. Gyorffy, B. (2021) Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. *Comput Struct Biotechnol J* **19**, 4101-4109
13. Au, C. C., Furness, J. B., Britt, K., Oshchepkova, S., Ladumor, H., Soo, K. Y., Callaghan, B., Gerard, C., Inghirami, G., Mittal, V., Wang, Y., Huang, X. Y., Spector, J. A., Andreopoulou, E., Zumbo, P., Betel, D., Dow, L., and Brown, K. A. (2020) Three-dimensional growth of breast cancer cells potentiates the anti-tumor effects of unacylated ghrelin and AZP-531. *Elife* **9**
14. Sriram, K., and Insel, P. A. (2018) G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol Pharmacol* **93**, 251-258
15. Hughes, C. S., Postovit, L. M., and Lajoie, G. A. (2010) Matrigel: a complex protein mixture required for optimal growth of cell culture. *Proteomics* **10**, 1886-1890
16. Guan, J. L., Trevithick, J. E., and Hynes, R. O. (1991) Fibronectin/integrin interaction induces tyrosine phosphorylation of a 120-kDa protein. *Cell Regul* **2**, 951-964
17. Kornberg, L., Earp, H. S., Parsons, J. T., Schaller, M., and Juliano, R. L. (1992) Cell adhesion or integrin clustering increases phosphorylation of a focal adhesion-associated tyrosine kinase. *J Biol Chem* **267**, 23439-23442
18. Jedeszko, C., Sameni, M., Olive, M. B., Moin, K., and Sloane, B. F. (2008) Visualizing protease activity in living cells: from two dimensions to four dimensions. *Curr Protoc Cell Biol* **Chapter 4**, Unit 4 20
19. Gonzalez, G. A., and Montminy, M. R. (1989) Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* **59**, 675-680
20. Kuleshov, M. V., Xie, Z., London, A. B. K., Yang, J., Evangelista, J. E., Lachmann, A., Shu, I., Torre, D., and Ma'ayan, A. (2021) KEA3: improved kinase enrichment analysis via data integration. *Nucleic Acids Res* **49**, W304-W316
21. Kang, T. H., Park, D. Y., Kim, W., and Kim, K. T. (2008) VRK1 phosphorylates CREB and mediates CCND1 expression. *J Cell Sci* **121**, 3035-3041
22. Martin, R. D., Sun, Y., Bourque, K., Audet, N., Inoue, A., Tanny, J. C., and Hebert, T. E. (2018) Receptor- and cellular compartment-specific activation of the cAMP/PKA pathway by alpha(1)-adrenergic and ETA endothelin receptors. *Cell Signal* **44**, 43-50
23. An, Y., Wei, N., Cheng, X., Li, Y., Liu, H., Wang, J., Xu, Z., Sun, Z., and Zhang, X. (2020) MCAM abnormal expression and clinical outcome associations are highly cancer dependent as revealed through pan-cancer analysis. *Brief Bioinform* **21**, 709-718
24. Wang, J., Tang, X., Weng, W., Qiao, Y., Lin, J., Liu, W., Liu, R., Ma, L., Yu, W., Yu, Y., Pan, Q., and Sun, F. (2015) The membrane protein melanoma cell adhesion

molecule (MCAM) is a novel tumor marker that stimulates tumorigenesis in hepatocellular carcinoma. *Oncogene* **34**, 5781-5795
25. Zhang, X., Wang, Z., Kang, Y., Li, X., Ma, X., and Ma, L. (2014) MCAM expression is associated with poor prognosis in non-small cell lung cancer. *Clin Transl Oncol* **16**, 178-183
26. Teng, Y., Xie, X., Walker, S., White, D. T., Mumm, J. S., and Cowell, J. K. (2013) Evaluating human cancer cell metastasis in zebrafish. *BMC Cancer* **13**, 453
27. Ren, J., Liu, S., Cui, C., and Ten Dijke, P. (2017) Invasive Behavior of Human Breast Cancer Cells in Embryonic Zebrafish. *J Vis Exp*
28. Blum, Y., Belting, H. G., Ellertsdottir, E., Herwig, L., Luders, F., and Affolter, M. (2008) Complex cell rearrangements during intersegmental vessel sprouting and vessel fusion in the zebrafish embryo. *Dev Biol* **316**, 312-322
29. Campenni, M., May, A. N., Boddy, A., Harris, V., and Nedelcu, A. M. (2020) Agent-based modelling reveals strategies to reduce the fitness and metastatic potential of circulating tumour cell clusters. *Evol Appl* **13**, 1635-1650
30. Tacar, O., Sriamornsak, P., and Dass, C. R. (2013) Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. *J Pharm Pharmacol* **65**, 157-170
31. Society, A. C. (2021) Chemotherapy for Breast Cancer.
32. Gopal, U., Monroe, J. D., Marudamuthu, A. S., Begum, S., Walters, B. J., Stewart, R. A., Washington, C. W., Gibert, Y., and Zachariah, M. A. (2023) Development of a Triple-Negative Breast Cancer Leptomeningeal Disease Model in Zebrafish. *Cells* **12**
33. Qu, Y., Han, B., Yu, Y., Yao, W., Bose, S., Karlan, B. Y., Giuliano, A. E., and Cui, X. (2015) Evaluation of MCF10A as a Reliable Model for Normal Human Mammary Epithelial Cells. *PLoS One* **10**, e0131285
34. Vilchez Mercedes, S. A., Bocci, F., Levine, H., Onuchic, J. N., Jolly, M. K., and Wong, P. K. (2021) Decoding leader cells in collective cancer invasion. *Nat Rev Cancer* **21**, 592-604
35. Kumar, S., Kapoor, A., Desai, S., Inamdar, M. M., and Sen, S. (2016) Proteolytic and non-proteolytic regulation of collective cell invasion: tuning by ECM density and organization. *Sci Rep* **6**, 19905
36. Han, W., Chen, S., Yuan, W., Fan, Q., Tian, J., Wang, X., Chen, L., Zhang, X., Wei, W., Liu, R., Qu, J., Jiao, Y., Austin, R. H., and Liu, L. (2016) Oriented collagen fibers direct tumor cell intravasation. *Proc Natl Acad Sci U S A* **113**, 11208-11213
37. Papadaki, M. A., Mala, A., Merodoulaki, A. C., Vassilakopoulou, M., Mavroudis, D., and Agelaki, S. (2022) Investigating the Role of CTCs with Stem/EMT-like Features in Metastatic Breast Cancer Patients Treated with Eribulin Mesylate. *Cancers (Basel)* **14**
38. Fleming, A., Diekmann, H., and Goldsmith, P. (2013) Functional characterisation of the maturation of the blood-brain barrier in larval zebrafish. *PLoS One* **8**, e77548
39. Rousselle, C., Clair, P., Lefauconnier, J. M., Kaczorek, M., Scherrmann, J. M., and Temsamani, J. (2000) New advances in the transport of doxorubicin through the blood-brain barrier by a peptide vector-mediated strategy. *Mol Pharmacol* **57**, 679-686

40. Insel, P. A., Murray, F., Yokoyama, U., Romano, S., Yun, H., Brown, L., Snead, A., Lu, D., and Aroonsakool, N. (2012) cAMP and Epac in the regulation of tissue fibrosis. *Br J Pharmacol* **166**, 447-456
41. Wang, P., Felsing, D. E., Chen, H., Stutz, S. J., Murphy, R. E., Cunningham, K. A., Allen, J. A., and Zhou, J. (2020) Discovery of Potent and Brain-Penetrant GPR52 Agonist that Suppresses Psychostimulant Behavior. *J Med Chem* **63**, 13951-13972
42. Masuho, I., Kise, R., Gainza, P., Von Moo, E., Li, X., Tany, R., Wakasugi-Masuho, H., Correia, B. E., and Martemyanov, K. A. (2023) Rules and mechanisms governing G protein coupling selectivity of GPCRs. *Cell Rep* **42**, 113173
43. Wu, Z., Wu, Z., Li, J., Yang, X., Wang, Y., Yu, Y., Ye, J., Xu, C., Qin, W., and Zhang, Z. (2012) MCAM is a novel metastasis marker and regulates spreading, apoptosis and invasion of ovarian cancer cells. *Tumour Biol* **33**, 1619-1628
44. Chen, J., Dang, Y., Feng, W., Qiao, C., Liu, D., Zhang, T., Wang, Y., Tian, D., Fan, D., Nie, Y., Wu, K., and Xia, L. (2020) SOX18 promotes gastric cancer metastasis through transactivating MCAM and CCL7. *Oncogene* **39**, 5536-5552
45. Zafra, M. P., Schatoff, E. M., Katti, A., Foronda, M., Breinig, M., Schweitzer, A. Y., Simon, A., Han, T., Goswami, S., Montgomery, E., Thibado, J., Kastenhuber, E. R., Sanchez-Rivera, F. J., Shi, J., Vakoc, C. R., Lowe, S. W., Tschaharganeh, D. F., and Dow, L. E. (2018) Optimized base editors enable efficient editing in cells, organoids and mice. *Nat Biotechnol* **36**, 888-893
46. Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oles, A. K., Pages, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L., and Morgan, M. (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* **12**, 115-121
47. Wrennall, J. A., Ahmad, S., Worthington, E. N., Wu, T., Goriounova, A. S., Voeller, A. S., Stewart, I. E., Ghosh, A., Krajewski, K., Tilley, S. L., Hickey, A. J., Sassano, M. F., and Tarran, R. (2022) A SPLUNC1 Peptidomimetic Inhibits Orai1 and Reduces Inflammation in a Murine Allergic Asthma Model. *Am J Respir Cell Mol Biol* **66**, 271-282