# 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], Caleb Kutz[4], Syeda Y. Jannath[2,5], Tabassum Fabiha[6], Bhavneet Bhinder[7], Michael P. Washburn[4,8], Dominic Devost[9], Shuchen Liu[2], Priya Bhardwaj[2], Todd Evans[3,] Pradeep K. Anand[10], Robert Tarran[10], Sailesh Palikhe[14], Olivier Elemento[6], Lukas E. 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. University of Kansas Cancer Center, Kansas City, KS, USA.
5. Hunter College, City University of New York, New York, NY, USA.
6. Columbia University Computational & Systems Biology Program, Sloan-Kettering Institute of Memorial Sloan-Kettering Cancer Center, New York, NY, USA.
7. Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, New York, USA.
8. Department of Cancer Biology, University of Kansas Medical 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:** Role of GPR52 in breast cancer

**Statement of Significance:**

We showed 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:**

Background: G protein-coupled receptors (GPCRs) are the largest class of membrane-bound receptors that transmit critical signals from extracellular to intracellular spaces. Transcriptomic data of resected breast tumors show that low mRNA expression of orphan

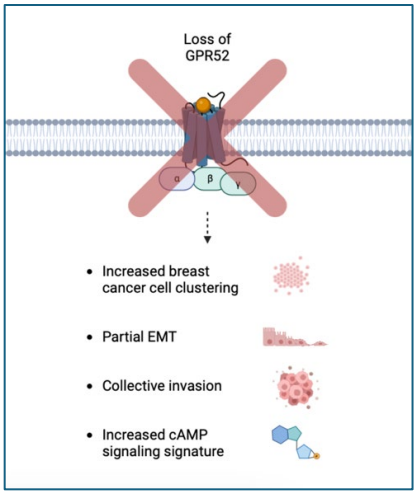
GPCR GPR52 correlates with reduced overall survival in patients with breast cancer, leading to the hypothesis that loss of GPR52 supports breast cancer progression.

Methods: CRISPR-Cas9 was used to knockout GPR52 in the 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. 2D and 3D *in vitro* studies, electron microscopy, Matrigel culture, and a zebrafish xenograft model were used to assess the morphology and behavior of GPR52 KO cells. RNA-sequencing and proteomic analyses were also conducted on these cell lines, and transcriptomic data from The Cancer Genome Atlas (TCGA) database were used to compare GPR52-null and wild-type (WT) signatures in breast cancer.

Results: 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, and zebrafish injected with GPR52 KO cells developed a greater total cancer area than those injected with control cells. RNA sequencing and proteomic analyses of GPR52-null breast cancer cells revealed an increased cAMP signaling signature. Consistently, we found that treatment of wild-type (WT) cells with forskolin, which stimulates the production of cAMP, induces phenotypic changes associated with GPR52 loss, and inhibition of cAMP production rescued some GPR52 KO phenotypes.

Conclusion: GPR52 is an orphan GPCR and its role in cancer progression has not been previously characterized. We found that GPR52 loss in breast cancer cells can lead to increased cell clustering, collective invasion, and EMT *in vitro*. These are features of increased cancer aggression. Our results reveal that GPR52 loss is a potential mechanism by which breast cancer progression may occur and support the investigation of GPR52 agonism as a therapeutic option for breast cancer.

**Graphical Abstract:**




**Keywords:** breast cancer, clustering, adhesion, collective invasion, cAMP, metastasis

**Background:**

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 cells undergo several adaptive processes to enhance their metastatic potential. These include changes

87 in cell-cell and cell-matrix adhesion, transitions between epithelial and mesenchymal cell
88 states, and the ability to degrade and invade tissues (2, 3). However, the upstream
89 regulators of these processes are not well characterized, which limits mechanistic
90 understanding and therapeutic intervention.
91
92 G protein-coupled receptors (GPCRs) are the largest protein family encoded by the
93 human genome (4). These receptors consist of an extracellular N-terminus followed by
94 seven transmembrane α-helices, which are connected by three intracellular and three
95 extracellular loops and a cytoplasmic C-terminal tail (5). Their transmembrane structure
96 enables the transmission of critical signals between extracellular and intracellular spaces.
97 Dissociation of the heterotrimeric G protein upon GPCR activation can regulate a diverse
98 array of downstream molecules, allowing for the regulation of various cell processes,
99 including proliferation, migration, adhesion, and metabolism (6, 7).
100
101 GPR52 is a structurally unique GPCR that is enriched in the basal ganglia and its
102 endogenous ligand remains unknown, rendering it an orphan receptor (8). It has garnered
103 increased attention in recent years owing to its potential as a neurotherapeutic target for
104 schizophrenia and Huntington's disease (9, 10). We examined GPR52 mRNA levels in
105 19 solid tumor types and determined that GPR52 is significantly downregulated in tumor
106 samples (11). However, the role of GPR52 in cancer progression has not been reported.
107 In patients with breast cancer, we found that GPR52 expression was further reduced in
108 metastases compared with that in the primary tumor (11). Low GPR52 mRNA expression
109 in resected breast tumors is also associated with a reduction in overall survival (12).

110

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

121

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

128

129 Overall, our results revealed that GPR52 loss is a potential mechanism by which
130 important processes in breast cancer progression, such as EMT and changes in
131 multicellular organization, may occur. These processes have long been implicated in
132 many solid tumors; however, critical and targetable upstream regulators have not been

133 identified. As GPCRs are the targets of more than 1/3 of FDA-approved small-molecule
134 drugs, these data support the investigation of GPR52 agonism as a viable therapeutic
135 approach in breast cancer (14).
136
137 **Results:**
138 **GPR52 expression is reduced in cancerous tissue compared to normal, and**
139 **inversely associated with breast cancer prognosis and metastatic potential**
140 To date, there have been no reports on the physiological role of GPR52 in any type of
141 cancer. However, GPR52 mRNA expression levels have been reported in many
142 transcriptome profiles of resected cancerous and noncancerous tissues. We compared
143 GPR52 mRNA expression levels in normal and tumor samples from the tissues from
144 which solid tumors arose (Fig. 1A) using the TNMplot webtool (11). We found that, in the
145 majority of these tissue types, GPR52 mRNA expression levels were lower in tumors than
146 in non-cancerous samples (Fig. 1A, $P<0.05$, indicated with an asterisk). Importantly, low
147 tumor GPR52 mRNA expression was found to be associated with a reduction in overall
148 survival in patients with breast cancer (Fig. 1B), with the impact of low GPR52 on survival
149 being more pronounced for triple-negative breast cancer (TNBC) (Fig. 1C) (12). GPR52
150 mRNA expression was also lower in metastatic nodes than in the primary resected tumor
151 (Fig. 1D, p=4.45e-17) (11), which is consistent with GPR52 the differentially expressed in
152 several human breast cell lines with varying metastatic potentials (Fig. 1E).
153
154 **GPR52 regulates cell-cell adhesion properties of cancerous and non-cancerous**
155 **breast epithelial cells**

156 We used CRISPR-Cas9 to generate indels in *GPR52* in two human breast cancer cell
157 lines (MDA-MB-231 and MDA-MB-468) and one non-cancerous breast epithelial cell line
158 (MCF10A), using two different guide RNAs (Supplementary Fig. S1). The morphology of
159 GPR52 KO cells was notably different from that of control cells. GPR52 KO MDA-MB-
160 231, MDA-MB-468, and MCF10A cell lines generated from both guide RNAs formed
161 clusters of tightly packed cells with clearly defined borders in the monolayer culture (Fig.
162 2A). To further investigate the effect of GPR52 loss on the interactions between cells, we
163 performed TEM of vector control and GPR52 KO MDA-MB-468 cells cultured in
164 suspension (Fig. 2B). GPR52 KO cells had a shorter cell-cell interface than WT cells (Fig.
165 2C), but they interacted more closely with one another along the length of their interface
166 than WT cells did (Fig. 2D). The cell diameter was not significantly different between the
167 groups, suggesting that the reduction in cell interface length in the GPR52 KO group was
168 not due to a reduction in cell size (Fig. 2E).
169

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

172 Given the rounded property of GPR52 KO cells, we sought to improve their adherence to
173 2D culture by coating tissue culture plates with poly-D-lysine (Fig. 3A). The culture of
174 GPR52 KO MDA-MB-468 cells on plastic coated with poly-D-lysine led to their elongation
175 and a mesenchymal phenotype. This was not observed in WT cells. Western blotting of
176 GPR52 KO MDA-MB-468 cells revealed upregulation of mesenchymal cell markers, such
177 as Snai1 and vimentin, but the cells continued to express E-cadherin (Fig. 3B).

178 Compared to WT cells, 3D culture of GPR52 KO MDA-MB-468 cells led to the formation
179 of disorganized spheroids with irregular borders (Fig. 3A-B, top rows; 3C). To further
180 define the impact of GPR52 loss on this invasive phenotype, we plated cells as a
181 monolayer and evaluated their invasion through a layer of Matrigel (Fig. 3E). Interestingly,
182 GPR52 KO cells tended to organize and invade collectively in large clusters (MDA-MB-
183 468) or sheets (MDA-MB-231) (Fig. 3F). We categorized all cancer foci with an area ≥
184 3140 $\mu m^2$ at z=10 μm as Class B structures (yellow outline). This threshold was selected
185 because the diameter of MDA-MB-468 and MDA-MB-231 cells was found to range from
186 to 10-20 μm, therefore, the cross-sectional area of approximately 10 cells was $10 \times \pi \times$
187 $r^2 = 10 \times \pi \times 10^2$=~3140). We then calculated the sum of the areas of all Class B
188 structures and divided this by the sum of the areas of all cancer foci at z=10 μm. We found
189 that the fraction of the area occupied by Class B structures was increased in the MDA-
190 MB-468 and MDA-MB-231 GPR52 KO groups, suggesting an increased propensity to
191 organize collectively (Fig. 3G-H).

192 Although collective invasion has been investigated previously, the effect of this behavior
193 on the ECM through which cells invade has not been well characterized. To assess the
194 pattern of ECM degradation, we incorporated dye-quenched (DQ)-collagen IV, which
195 contains sequestered fluorophores that are released following proteolysis of collagen IV,
196 a major component of the basement membrane that breast cancer cells invade (18). At
197 the 24-hour timepoint, the DQ signal was found to colocalize strongly with the GPR52 KO
198 cells, while it was more diffuse and less colocalized with the vector control cells for both
199 MDA-MB-468 and MDA-MB-231 cells (Fig. 3I). Furthermore, we normalized the area of
200 the DQ signal to the area occupied by cancer cells and found that it tended to be lower

201 for the GPR52 KO groups than for the vector control, suggesting a focal area of matrix
202 degradation (Fig. 3L-M).

203

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

206 Next, we conducted proteomic analyses of the WT and GPR52 KO MDA-MB-468, MDA-
207 MB-231, and MCF10A cell lines. Using Ingenuity Pathway Analysis (IPA, Qiagen), we
208 observed upregulation of signatures associated with cellular homeostasis, viability,
209 survival, proliferation, and migration in all three cell lines, whereas the signatures
210 associated with organismal death and cell death of tumor cell lines were both significantly
211 reduced (Fig. 4A). Several upstream regulators were predicted to be similarly altered
212 across the GPR52 KO lines (Fig. 4B). We were particularly intrigued by the signature
213 associated with the cAMP analog 8-bromo-cAMP, as GPR52 activation and function have
214 been associated with its regulation of intracellular cAMP levels (9, 10).

215

216 The TCGA-BRCA dataset reports RNA sequencing data of resected breast tumors from
217 a cohort of patients. Normalized GPR52 mRNA expression in TCGA-BRCA tumors was
218 visualized (Fig. 4C), and it was determined that 45.7% did not express GPR52, whereas
219 the remainder expressed non-zero levels of GPR52 (Fig. 4D). The RNA-sequencing
220 datasets were compared for the GPR52 non-expressing and -expressing cohorts and the
221 differentially expressed genes (DEGs) were imported into IPA. We also conducted RNA
222 sequencing of GPR52 KO and WT MDA-MB-468 cells that were cultured in a monolayer
223 and imported the DEGs into IPA. We identified several common signatures in both

224 datasets. Notably, upregulation of the signature associated with cAMP response element-
225 binding protein (CREB1), a transcription factor that is activated downstream of cAMP
226 signaling, was increased in both datasets (Fig. 4E) (19). Based on phosphoproteomic
227 analyses of MDA-MB-468, MDA-MB-231, and MCF10A cells, we identified kinases
228 predicted to be active in the vector control (Fig. 4F) and GPR52 KO (Fig. 4G) datasets
229 (20). The rank of each of the top kinases in the GPR52 KO cell lines was then compared
230 to that in the vector control group (Fig. 4H). This revealed the greatest increase in the
231 rank of VRK1, which regulates cell cycle progression via phosphorylation and activation
232 of CREB (21).
233

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

236 Next, we explored whether cAMP could modulate the phenotypes observed following
237 GPR52 loss in MDA-MB-468 cells. GPR52 KO cells exhibited less rounding and were
238 able to spread on plastic with adenylyl cyclase inhibitor (ACi) treatment (Fig. 5A; arrows).
239 Interestingly, the expression of Snai1 was induced by forskolin (FSK, a direct activator of
240 adenylyl cyclase) treatment of WT cells, whereas vimentin expression was reduced in
241 GPR52 KO cells treated with ACi (Fig. 5B). ACi treatment of GPR52 KO MDA-MB-231
242 cells also increased the area of DQ normalized to the area of tdTomato without affecting
243 the degree of colocalization of DQ and tdTomato (Fig. 5C-E).

244 To further explore the relationship between GPR52 and cAMP signaling, we next
245 determined whether GPR52 activation led to a change in intracellular cAMP levels, as
246 described by other groups (8, 9). To do this, we used a bioluminescence resonance

247 energy transfer 2 (BRET2)-based EPAC sensor, GFP10-EPAC-RlucII (Supplementary
248 Fig. S2A) (22). This modified form of EPAC contains a luminescent donor and a
249 fluorescent acceptor that are in proximity to one another when EPAC is not bound to
250 cAMP. However, the binding of cAMP induces a conformational shift that leads to an
251 increased distance between the donor and acceptor, and a reduction in energy transfer.
252 We found that increasing the amount of GPR52 and/or the concentration of FTBMT, a
253 synthetic GPR52 agonist, led to a reduction in the BRET ratio in HEK293 cells
254 (Supplementary Fig. S2B). Next, we introduced FTBMT and the EPAC sensor into MDA-
255 MB-468 cells and found that BRET values tended to increase at higher doses
256 (Supplementary Fig. S2C). To investigate whether GPR52 couples with $G_{aq}$, we also
257 quantified intracellular $Ca^{2+}$ levels over a similar range of FTBMT doses using the $Ca^{2+}$-
258 sensitive dye Fluo-4. Despite observing a robust response to the sarcoendoplasmic
259 reticulum calcium ATPase (SERCA) inhibitor thapsigargin, we observed no change in
260 baseline or thapsigargin-induced cytoplasmic $Ca^{2+}$ levels (Supplementary Fig. S2D).

261

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

264 RNA-sequencing analysis of MDA-MB-468 cells demonstrated that GPR52 loss is
265 associated with differential expression of many cell adhesion molecules (CAMs) (Fig. 6A),
266 including melanoma cell adhesion molecule (MCAM). MCAM is highly expressed in large
267 blood vessels, but recent studies have also described increased MCAM expression in
268 certain cancer types and its promotion of cancer progression (23-25). In breast cancer,
269 increased MCAM mRNA expression in resected tumors was associated with a reduction

270 in overall survival (Fig. 6B) (12). The expression of MCAM and GPR52 mRNA was
271 inversely correlated in resected breast tumors (Fig. 6C). We found that MCAM protein
272 expression was increased in GPR52 KO MDA-MB-468 cells and that 1.4 mM ACi
273 treatment reduced MCAM expression (Fig. 6D). This was consistent with the effects of
274 GPR52 loss in MCF10A cells (Supplementary Fig. S3).

275

276 **Loss of GPR52 is associated with increased TNBC burden in zebrafish**

277 As we observed that loss of GPR52 led to changes in organization, cell-cell adhesion,
278 and invasiveness of breast cancer cells, we wanted to determine whether these changes
279 were associated with an increase in breast cancer burden *in vivo.* To enable close
280 monitoring of cancer cell organization and distribution, we used a zebrafish xenograft
281 model (Fig. 7A) (26, 27). The TG(flk1:EGFP-NLS) zebrafish strain, which constitutively
282 expresses a green fluorescent protein in endothelial cells, allows cancer cells to be
283 monitored in relation to the zebrafish vasculature (28).

284 Two independent studies were performed to compare the behavior of the vector
285 control MDA-MB-468 cells with GPR52 sgRNA1 and sgRNA2 cells. We found that WT
286 and GPR52 KO MDA-MB-468 cells were detectable throughout the body of the zebrafish
287 at 30 h post-injection (hpi) (Fig. 7B). We also observed that WT and GPR52 KO cells
288 circulated collectively in the zebrafish bloodstream and identified endothelial cells
289 between cancer cells, suggesting an interaction between the two cell types (Fig. 7C). The
290 number of cancer foci did not differ significantly between groups at 30 hpi (Fig. 7D-E).
291 However, the total cancer area, which was calculated as the sum of the area of cancer in

292 the head (superior to the otolith) and trunk (distal to the injection site, not including the
293 yolk sac) (Fig. 7F-I),) was significantly greater in zebrafish injected with GPR52 KO cells
294 (Fig. 7J-K).

295 Increased clustering of cancer cells is associated with reduced sensitivity to
296 cytotoxic chemotherapeutic drugs, particularly when used as a single agent (2, 29).
297 Doxorubicin is one of the most potent chemotherapeutic drugs approved by the Food and
298 Drug Administration and is used in breast cancer treatment (30, 31). Therefore, we
299 designed a zebrafish xenograft therapeutic study that incorporated 8 μM doxorubicin or
300 the vehicle control (Milli-Q water) in E3 water maintained at 5 hpi and then quantified the
301 cancer area at 30 hpi (32). At 30 hpi, there was a trend towards reduced total cancer area
302 in doxorubicin-treated animals for both WT and GPR52 KO groups (Fig. 7L-M).
303 Doxorubicin caused a significant reduction in the breast cancer cell area in the head of
304 the zebrafish for both WT and GPR52 KO cells, with noticeable residual disease in
305 animals xenografted with GPR52 KO cells (Fig. 7N-O).

306

307 **Discussion**

308 The data presented herein demonstrate that GPR52 is a novel regulator of
309 multicellular organization in breast cancer and its loss can promote features of cancer
310 progression. First, we observed an increase in cell-cell interactions and cell clustering in
311 2D cultures with GPR52 loss in MCF10A, MDA-MB-468, and MDA-MB-231 cell lines.
312 TEM demonstrated that GPR52 loss in MDA-MB-468 cells is associated with a reduction
313 in the length of the interface between the two cells, but that GPR52 KO cells are more

314 juxtaposed with one another than WT cells, which exhibit more intercellular space along
315 their interface. Multicellular aggregation and increased cell-cell adhesion are mechanisms
316 by which cancer cells can increase their metastatic potential (2). Moreover, the increase
317 in disorganization and number of MDA-MB-468 spheroids with GPR52 loss are two
318 hallmarks of cancer progression, with the latter suggesting a potential increase in
319 stemness due to an increased propensity to survive and proliferate in Matrigel (33).

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

329 Many upstream regulators of EMT have been identified; however, to our
330 knowledge, induction of EMT or partial EMT with the loss of a GPCR has not been
331 previously described. Importantly, partial EMT has been documented in circulating tumor
332 cells (37). The exclusivity of our observations of cell elongation and the development of
333 mesenchymal morphology on poly-D-lysine, but not tissue culture-treated plastic,
334 suggests that integrin activity can influence the extent of EMT in this cell line and is an
335 interesting area for further exploration.

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

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

359 interaction between GPR52 and different G proteins, particularly in different cell types,
360 warrants further investigation.

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

367 **Conclusion:**

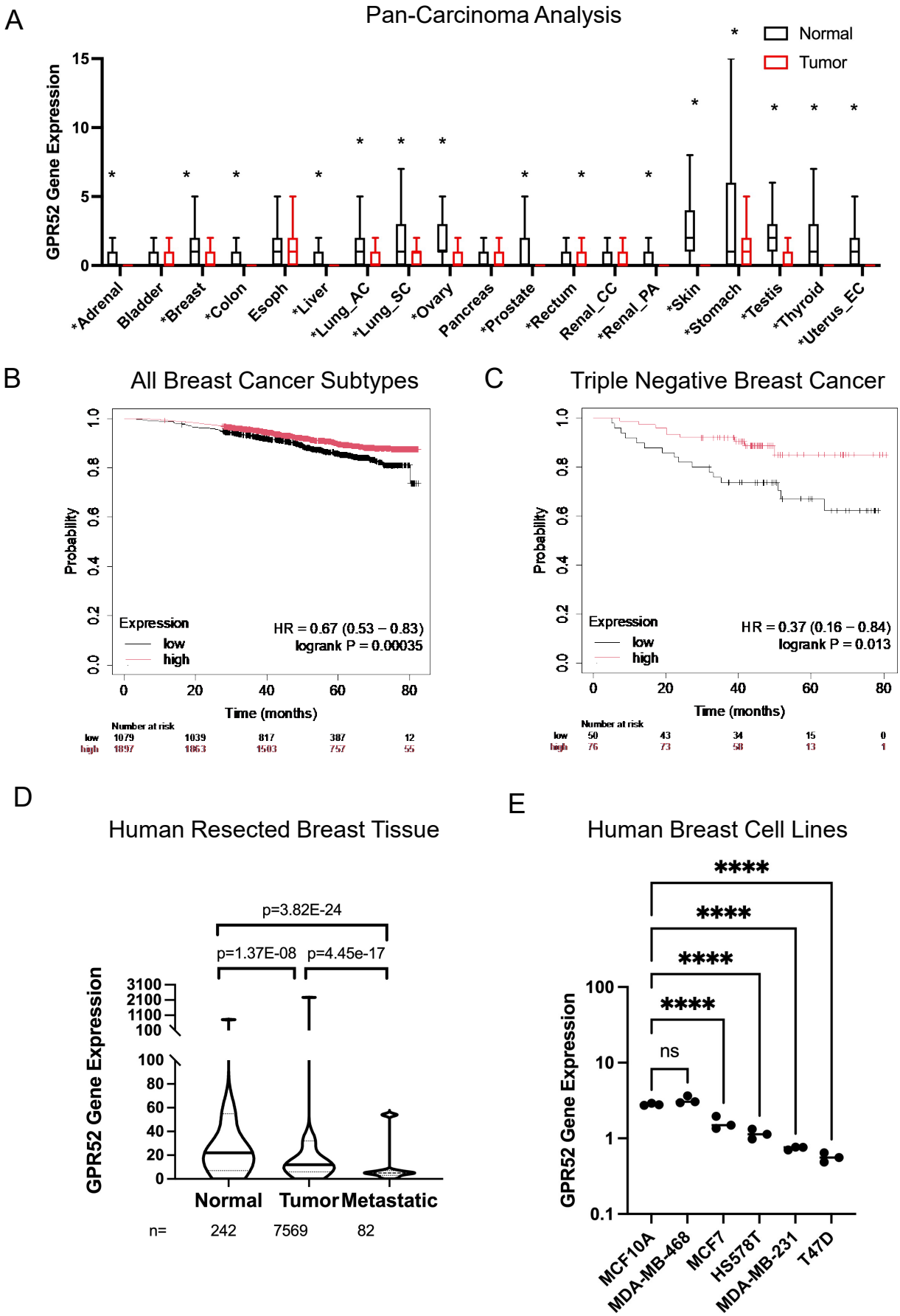
368 Our study identified GPR52 as a regulator of cancer cell clustering, collective
369 organization, invasion, and EMT. Using a zebrafish xenograft model, we demonstrated
370 that loss of GPR52 is associated with increased tumor burden. RNA sequencing and
371 proteomic analyses of WT and GPR52-null breast cancer cells demonstrated differences
372 in many gene and protein signatures, including increased cAMP signaling in GPR52-null
373 groups. We showed that inhibition of cAMP production can rescue some phenotypes
374 associated with GPR52 loss. Thus, we provide a rationale for investigating the therapeutic
375 effect of GPR52 agonism in breast cancer and encourage investigation of the role of
376 GPR52 in the progression of additional cancer types. Furthermore, our work identifies
377 novel features of cell biology that can be regulated by GPCR and broadens the scope of
378 significance of this class of proteins in cancer biology.

379

380 **Figures:**

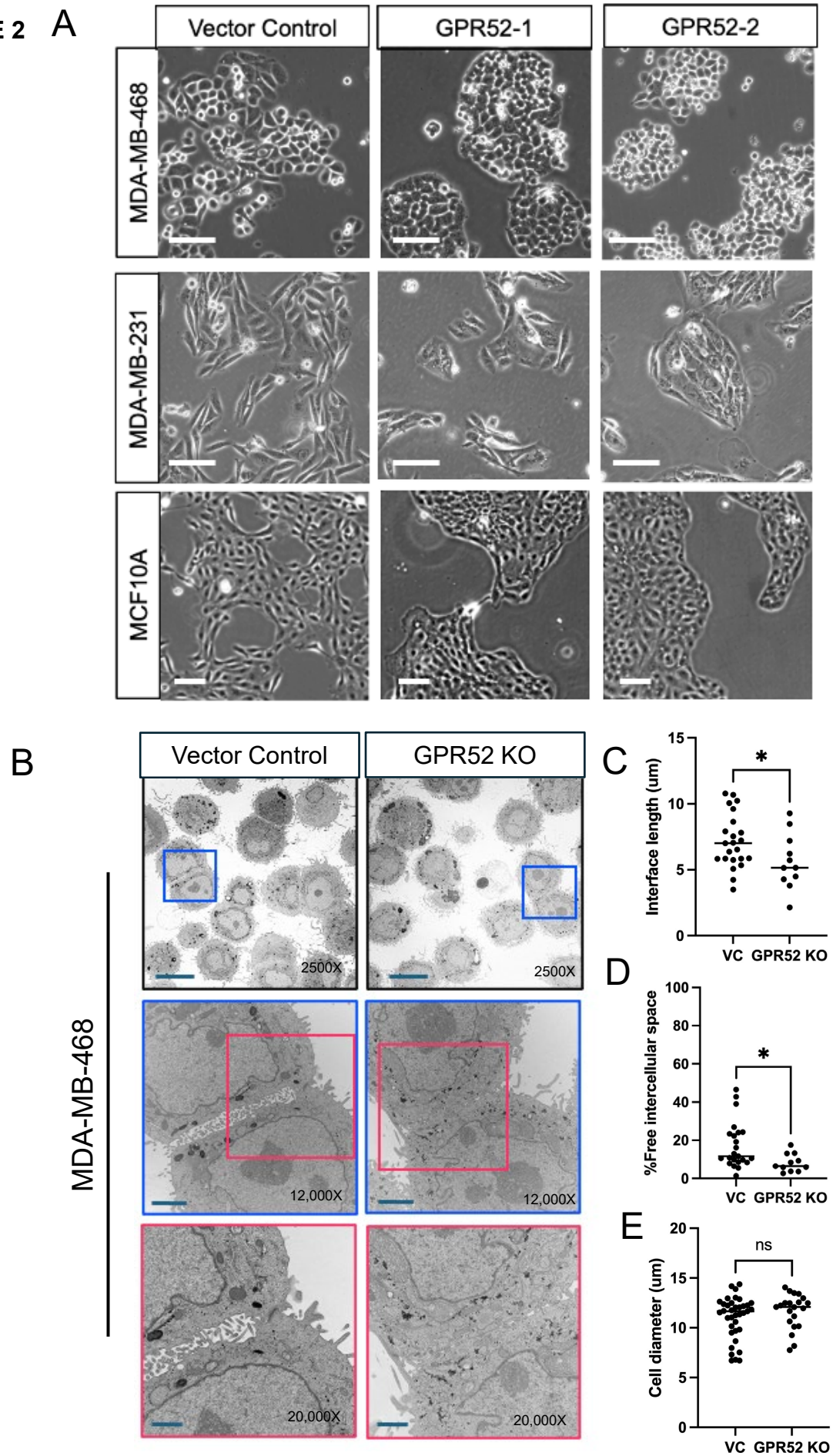
381

**FIGURE 1**


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
GPR52 Gene Expression
ns
MCF10A
MDA-MB-468
MCF7
HS578T
MDA-MB-231
T47D

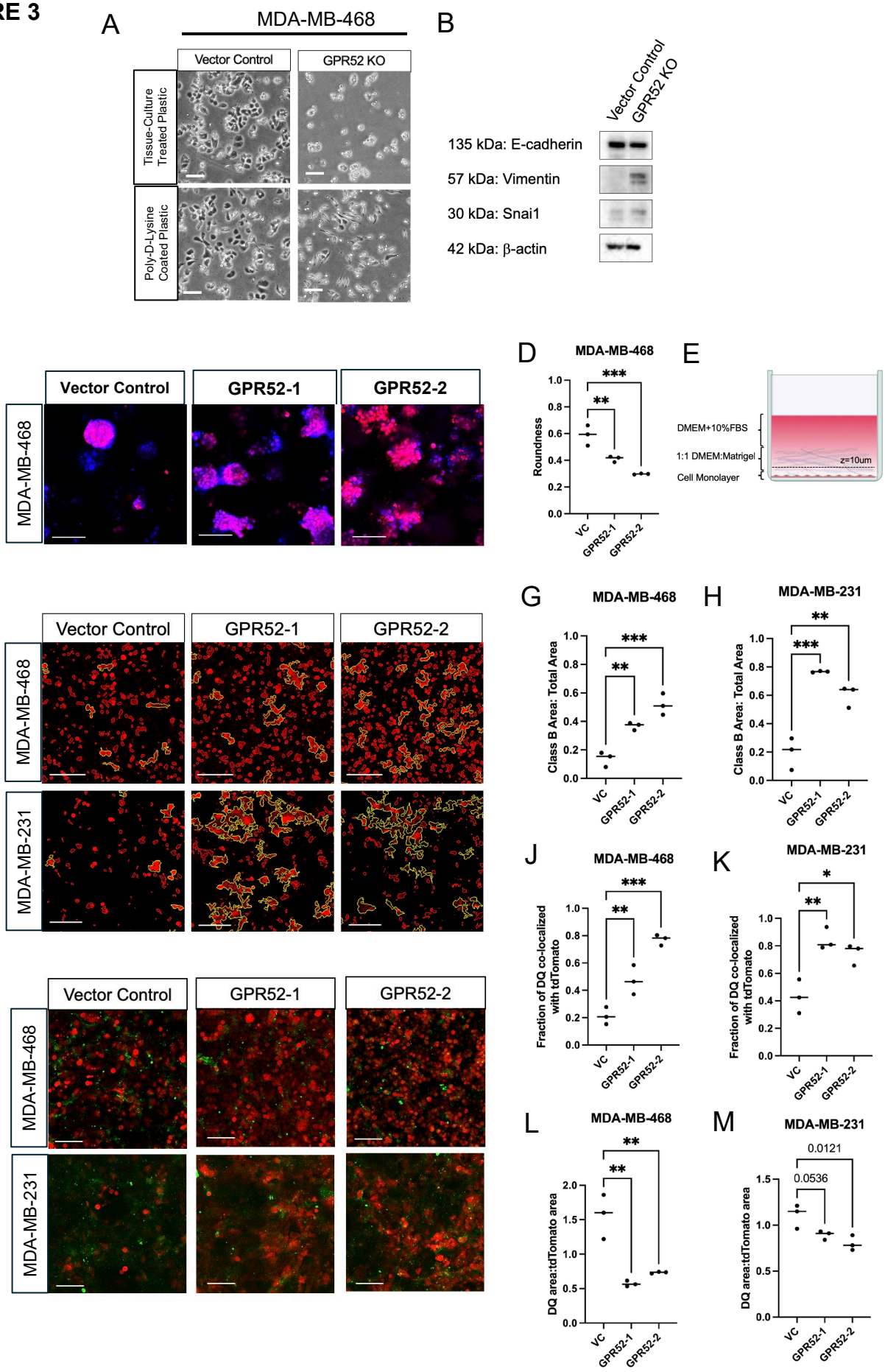
382 **Figure 1. Low GPR52 expression is associated with increased breast cancer**
383 **progression and reduced survival probability of breast cancer patients.** A) The
384 median GPR52 mRNA expression, interquartile range, minimum value, and upper
385 whisker are plotted for normal (black) and tumor (red) tissues. Mann-Whitney U tests were
386 conducted to determine statistical significance. *$P$ <0.05. AC=adenocarcinoma,
387 SC=squamous cell, CC=clear cell, PA=papillary cell, EC=endometrial carcinoma. B)
388 KMplot breast cancer overall survival curves for patients with low versus high GPR52
389 mRNA expression in resected tumors for (B) all breast cancer subtypes and (C) triple
390 negative breast cancer. D) GPR52 mRNA expression collected from a gene chip dataset
391 of non-cancerous breast tissue, primary tumor, and metastases of individuals with breast
392 cancer (un-paired). Data are presented as median with upper and lower quartiles and
393 minimum and maximum values. One-way ANOVA, p<0.05. E) GPR52 mRNA expression
394 was determined by QPCR and normalized to the housekeeping gene RPL32. The
395 normalized GPR52 expression was then divided by the average expression of GPR52
396 across the cell lines. n=3, line=median. One-way ANOVA, P-value<0.05; ****$P$ <0.00005,
397 ns=not significant.
398
399

**FIGURE 2**


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
Interface length (um)
15
10
5
0
*
VC
GPR52 KO
D
%Free intercellular space
100
80
60
40
20
0
*
VC
GPR52 KO
E
Cell diameter (um)
20
15
10
5
0
ns
VC
GPR52 KO

400 **Figure 2. GPR52 regulates cell-cell adhesion properties of cancerous and non-**
401 **cancerous breast epithelial cells.** A) Human breast cell lines were grown on tissue
402 culture-treated plastic under standard cell culture conditions and imaged at 50-60%
403 confluency with light microscopy. Scalebar=100 μm. B) MDA-MB-468 cells cultured in
404 suspension were visualized with TEM from 2500-20,000x magnification. Scalebars:
405 2500x=10 mm; 12,000x=2 mm, 20,000x=1 mm. C) The length of the interface between
406 cells, D) fraction of the cell-cell interface that was occupied by free space, and E) cell
407 diameter were determined using 2500x images. Student's t-test, P-value<0.05; *$P$ <0.05,
408 ns=not significant. Line=median. VC=vector control.
409
410

**FIGURE 3**


A
MDA-MB-468
Vector Control
GPR52 KO
Tissue-Culture Treated Plastic
Poly-D-Lysine Coated Plastic
B
Vector Control
GPR52 KO
135 kDa: E-cadherin
57 kDa: Vimentin
30 kDa: Snai1
42 kDa: β-actin
C
Vector Control
GPR52-1
GPR52-2
MDA-MB-468
D
MDA-MB-468
Roundness
VC
GPR52-1
GPR52-2
E
DMEM+10%FBS
1:1 DMEM:Matrigel
Cell Monolayer
z=10um
F
Vector Control
GPR52-1
GPR52-2
MDA-MB-468
MDA-MB-231
G
MDA-MB-468
Class B Area: Total Area
H
MDA-MB-231
Class B Area: Total Area
I
Vector Control
GPR52-1
GPR52-2
MDA-MB-468
MDA-MB-231
J
MDA-MB-468
Fraction of DQ co-localized with tdTomato
K
MDA-MB-231
Fraction of DQ co-localized with tdTomato
L
MDA-MB-468
DQ area:tdTomato area
M
MDA-MB-231
0.0121
0.0536
DQ area:tdTomato area

411 **Figure 3. GPR52 loss is associated with EMT, collective invasion, and an altered**
412 **pattern of ECM digestion in breast cancer.** A) Vector control and GPR52 KO (sgRNA2)
413 MDA-MB-468 cells were cultured on tissue-culture treated plates that were untreated (top
414 row) or treated with 1mg/mL poly-D-lysine coated plastic (bottom row). Scalebar=100
415 mm. B) Western blotting of MDA-MB-468 cells cultured on poly-D-lysine. C) Cells were
416 cultured in Matrigel for 10 days under complete media conditions. Cytoplasmic tdTomato
417 (red) and 1:1000 nuclear Hoechst 33342 stain (blue) allow visualization of MDA-MB-468
418 WT (vector control) and GPR52 KO (GPR52-1, GPR52-2) spheroids. Scalebar=100 mm.
419 D) The roundness of spheroids was determined based on the tdTomato signal. E)
420 Schematic of invasion assay. FBS=fetal bovine serum, DMEM= Dulbecco's Modified
421 Eagle Medium. F) Z-stacks obtained at t=24 hours were assessed at z=10 mm. All Class
422 B structures are outlined in yellow. Scalebar=150 μm. The fraction of the area occupied
423 by Class B structures for (G) MDA-MB-468 and (H) MDA-MB-231. VC=vector control.
424 n=3, One-way ANOVA, P-value<0.05. *$P$ <0.05, **$P$ <0.005, ns=not significant.
425 Line=median. I) Representative images are shown of tdTomato-tagged cancer cells (red)
426 and DQ-collagen (green) at t=24 hours at z=10 μm. Scalebar=100 mm. The fraction of
427 DQ co-localized with tdTomato at z=10 mm (J-K) and area of DQ normalized to the area
428 of tdTomato (J-K) were determined for MDA-MB-468 and MDA-MB-231, respectively.
429 n=3, One-way ANOVA, P-value<0.05. *$P$ <0.05, **$P$ <0.005, ***$P$ <0.0005, ns=not
430 significant. Line=median.
431
432

# FIGURE 4

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 cell*** | 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 |

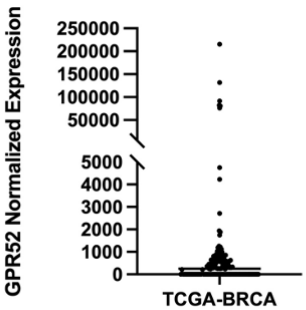
negative z-score  positive z-score

B

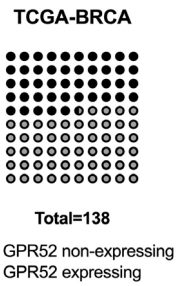
| Upstream Regulators | MDA-MB-468 | MDA-MB-231 | MCF10A |
|---|---|---|---|
| ST1926 | -3.86 | -1.195 | -1.726 |
| beta-estradiol | 2.087 | 2.078 | 2.373 |
| CD437 | -3.559 | -1.134 | -1.667 |
| medroxyprogesterone acetate | 1.511 | 2.034 | 2.183 |
| Z-LLL-CHO | 2.647 | N/A | 2.77 |
| HRAS | -2.538 | -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 UUUCAGU | -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

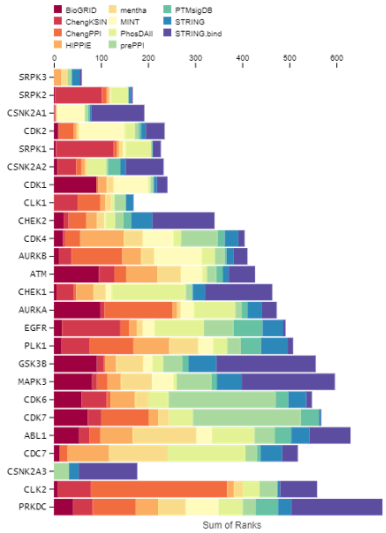



E

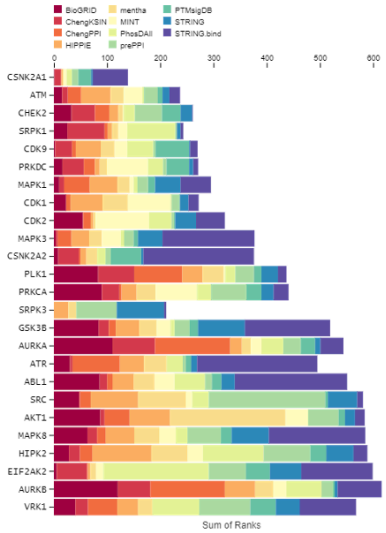
| Upstream Regulators | MDA-MB-468 | TCGA-BRCA |
|---|---|---|
| epicatechin | 2.07E+00 | 2.701 |
| FOXA2 | 2.685 | 2.008 |
| N-CoR | 1.698 | -2.646 |
| TNC | 1.941 | -2.219 |
| CDC20 | -1.89 | -2.236 |
| propylthiouracil | -1.769 | -2.331 |
| 10E-12Z-octadecadienoic acid | 1.681 | -2.275 |
| TNF | -1.014 | -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



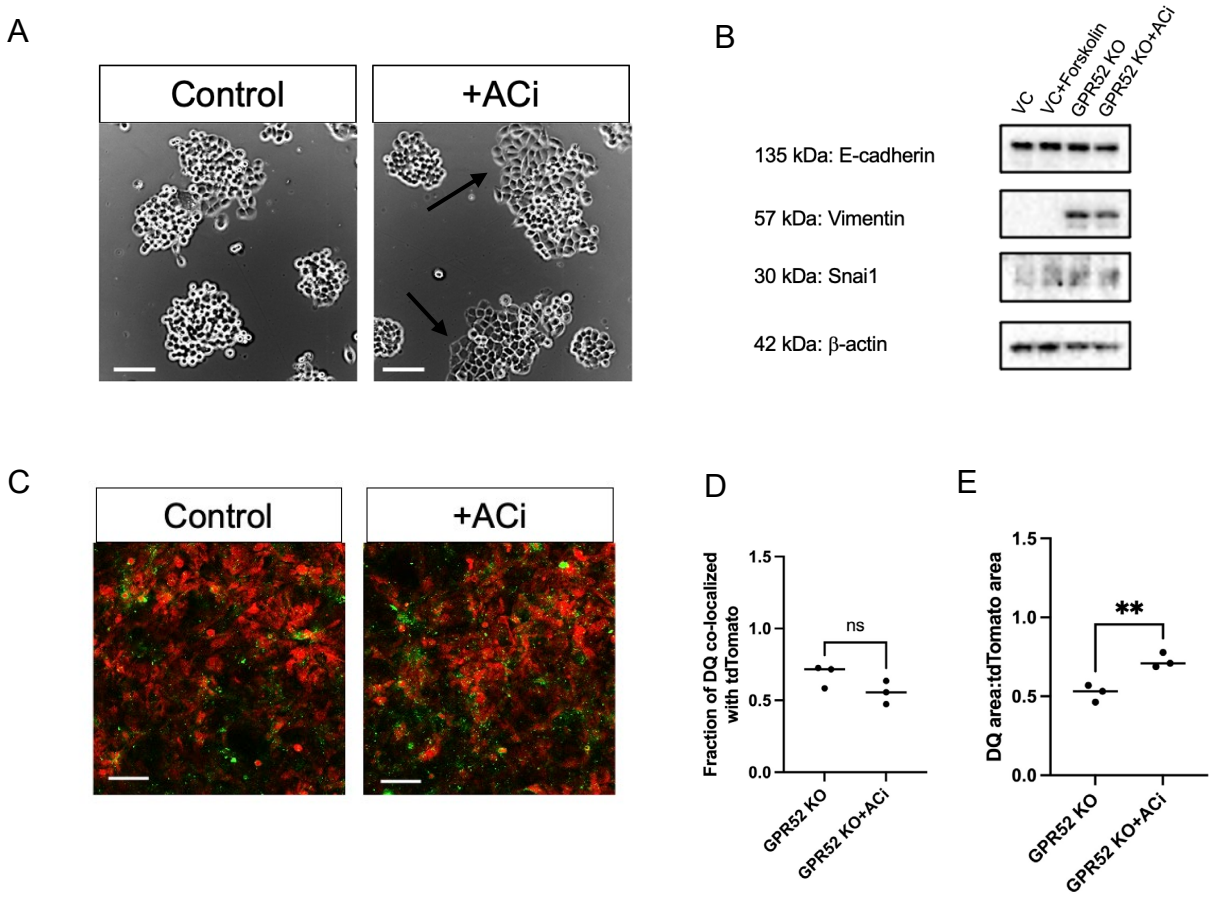

H

| 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

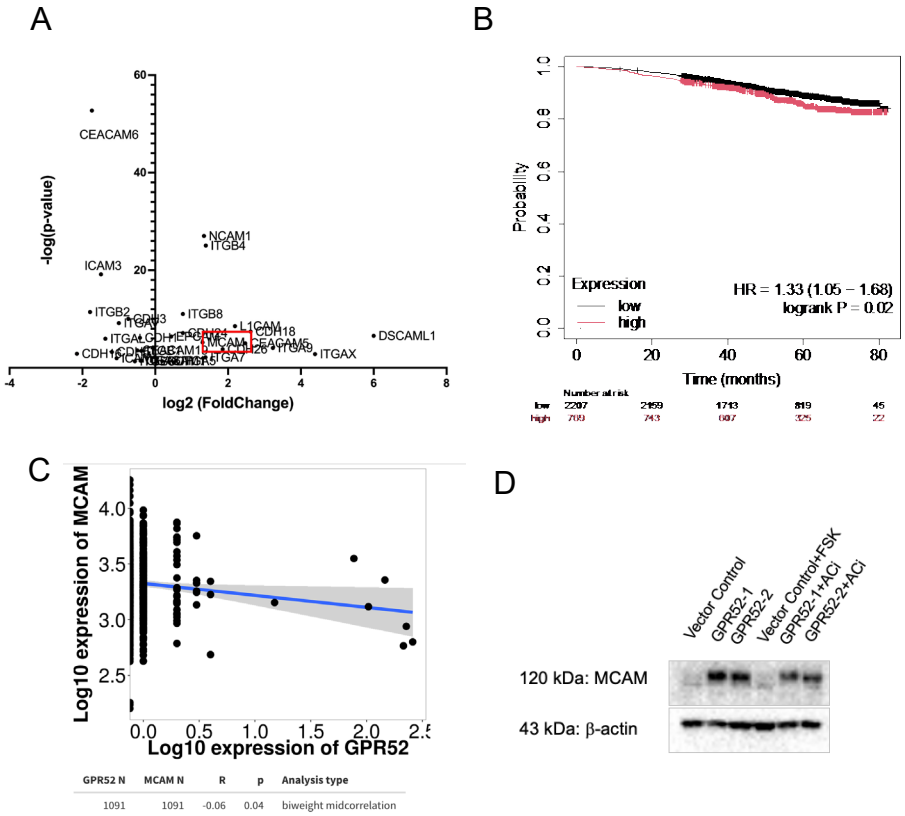
433 **Figure 4. Proteomic and transcriptomic analyses of GPR52-null breast cell lines**
434 **and resected breast tumors.** (A-B) Differentially expressed proteins in the GPR52 KO
435 cell lines with P-value <0.05, FDR<0.05, and fold-change greater than two were imported
436 into IPA for each of the three cell lines. The common predicted upstream regulators (A)
437 and top diseases and functions (B) with a P-value <0.05 are depicted. Pathways are
438 marked N/A (not-applicable) if the z-score is not determined. (C) Normalized expression
439 of GPR52 in TCGA-BRCA cohort resected breast tumors D) Graphical representation of
440 proportion of patients with undetected (black) and detected (grey) GPR52 mRNA in
441 resected breast tumors. E) DEGs in GPR52 KO MDA-MB-468 cells and the GPR52-null
442 TCGA-BRCA cohort with P-value <0.05, FDR<0.05, and fold-change greater than two
443 were imported into IPA for pathway analysis. Upstream regulators predicted to be
444 responsible for the DEGs observed in both datasets with a P-value <0.05 are depicted.
445 (F-H) Phosphoproteomic analysis was conducted on the groups outlined in Fig 5A. KEA3
446 software was used to rank kinases that were predicted to be active in the vector control
447 (F) and GPR52 KO (G) cell lines. H) The difference in rank of kinases which were reported
448 for both the GPR52 KO and vector control groups.
449
450

**FIGURE 5**


A
Control
+ACi
B
VC
VC+Forskolin
GPR52 KO
GPR52 KO+ACi
135 kDa: E-cadherin
57 kDa: Vimentin
30 kDa: Snai1
42 kDa: β-actin
C
Control
+ACi
D
Fraction of DQ co-localized with tdTomato
1.5
1.0
0.5
0.0
ns
GPR52 KO
GPR52 KO+ACi
E
DQ area:tdTomato area
1.5
1.0
0.5
0.0
**
GPR52 KO
GPR52 KO+ACi

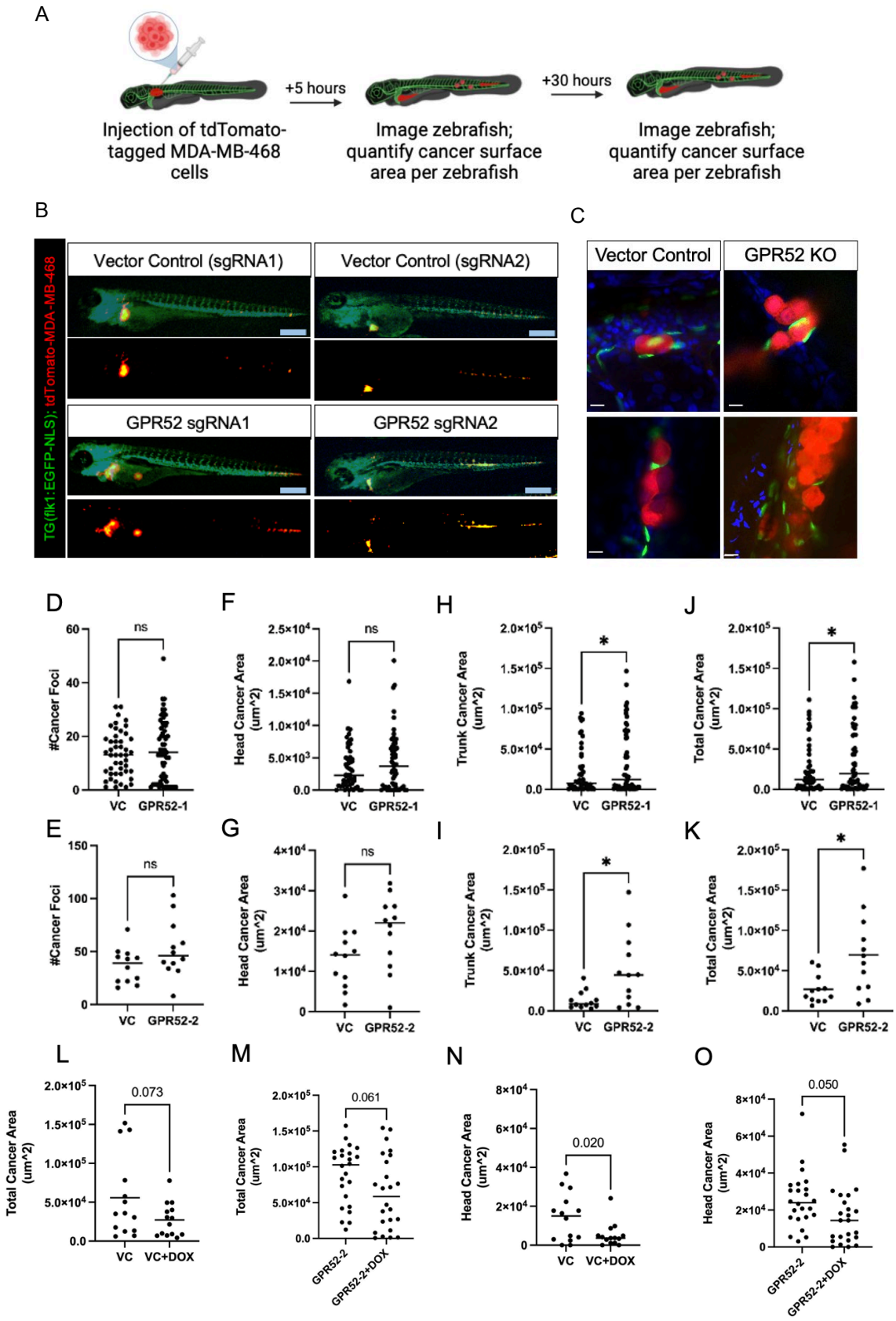
451 **Figure 5. cAMP partially mediates phenotypes associated with GPR52 loss.** A)
452 GPR52 KO (sgRNA2) MDA-MB-468 cells were cultured in monolayer and treated with
453 1.4 μM ACi or the vehicle control. Treatment was replaced every 72 hours over 6-8 days.
454 Scalebar=100 mm. B) Western blotting of MDA-MB-468 cells treated for 24 hours. C)
455 GPR52 KO MDA-MB-231 cells were plated to confluency in monolayer and treated the
456 next day with 1.4 mM ACi or vehicle control. Invasion assays were conducted as in **Fig.**
457 **3**. Scalebar=100 mm. The fraction of DQ co-localized with tdTomato (D) and area of DQ
458 normalized to the area of tdTomato (E) at z=10 mm at t=24 hours. n=3, Student's t-test,
459 P-value<0.05; *$P$ <0.05, **$P$ <0.005, ns=not significant. Line=median.
460
461

# FIGURE 6


A
-log(p-value)
log2 (FoldChange)
CEACAM6
NCAM1
ITGB4
ICAM3
ITGB2
ITGB8
L1CAM
CDH18
DSCAML1
ITGAX
B
Probability
Time (months)
Expression
low
high
HR = 1.33 (1.05 – 1.68)
logrank P = 0.02
Number at risk
C
Log10 expression of MCAM
Log10 expression of GPR52
GPR52 N
MCAM N
R
p
Analysis type
1091
1091
-0.06
0.04
biweight midcorrelation
D
Vector Control
GPR52-1
GPR52-2
Vector Control+FSK
GPR52-1+ACi
GPR52-2+ACi
120 kDa: MCAM
43 kDa: β-actin

462 **Figure 6. Expression of the melanoma cell adhesion molecule (MCAM) is inversely**
463 **related to GPR52 in breast cancer and is regulated by cAMP.** (A) CAMs that were
464 differentially expressed in GPR52 KO MDA-MB-468 cells based on RNA-sequencing. P-
465 value<0.05. n=3. (B) KMplot breast cancer RNA-seq webtool overall survival curves for
466 patients with low versus high MCAM mRNA expression in resected tumors for all breast
467 cancer subtypes. Low versus high cutoff was determined based on the maximum
468 segregation between the groups. (C) TNMplot correlation of MCAM and GPR52 transcript
469 expression in breast tumors from an RNA-sequencing dataset. (D) MDA-MB-468 cells
470 were cultured in monolayer (TC-treated, non-PDL coated) and treated with forskolin
471 (FSK), SQ22536 (ACi), or vehicle control for 24 hours. Cell lysates were probed for MCAM
472 and beta-actin.
473
474

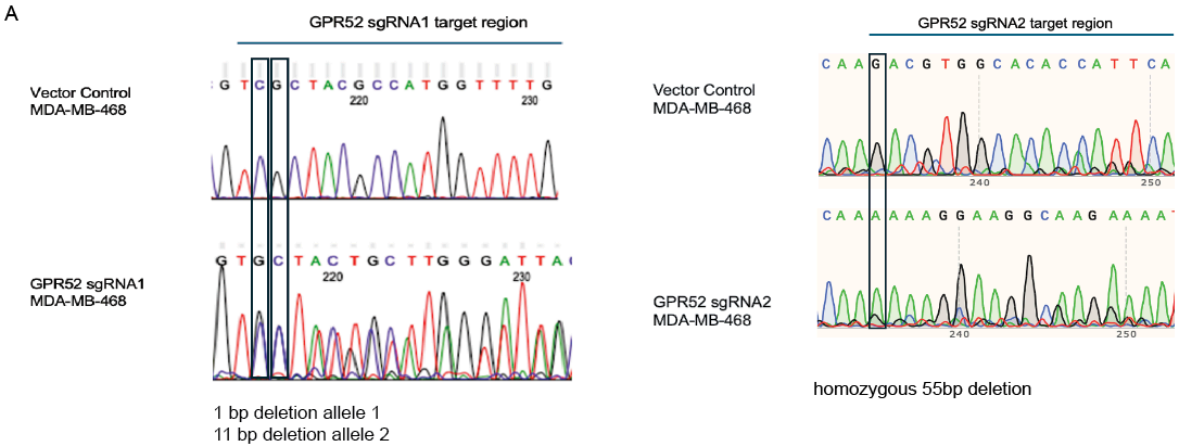
# FIGURE 7


A
+5 hours
+30 hours
Injection of tdTomato-tagged MDA-MB-468 cells
Image zebrafish; quantify cancer surface area per zebrafish
Image zebrafish; quantify cancer surface area per zebrafish
B
TG(flk1:EGFP-NLS); tdTomato-MDA-MB-468
Vector Control (sgRNA1)
Vector Control (sgRNA2)
GPR52 sgRNA1
GPR52 sgRNA2
C
Vector Control
GPR52 KO
D
#Cancer Foci
ns
VC
GPR52-1
E
#Cancer Foci
ns
VC
GPR52-2
F
Head Cancer Area (um^2)
ns
VC
GPR52-1
G
Head Cancer Area (um^2)
ns
VC
GPR52-2
H
Trunk Cancer Area (um^2)
*
VC
GPR52-1
I
Trunk Cancer Area (um^2)
*
VC
GPR52-2
J
Total Cancer Area (um^2)
*
VC
GPR52-1
K
Total Cancer Area (um^2)
*
VC
GPR52-2
L
Total Cancer Area (um^2)
0.073
VC
VC+DOX
M
Total Cancer Area (um^2)
0.061
GPR52-2
GPR52-2+DOX
N
Head Cancer Area (um^2)
0.020
VC
VC+DOX
O
Head Cancer Area (um^2)
0.050
GPR52-2
GPR52-2+DOX

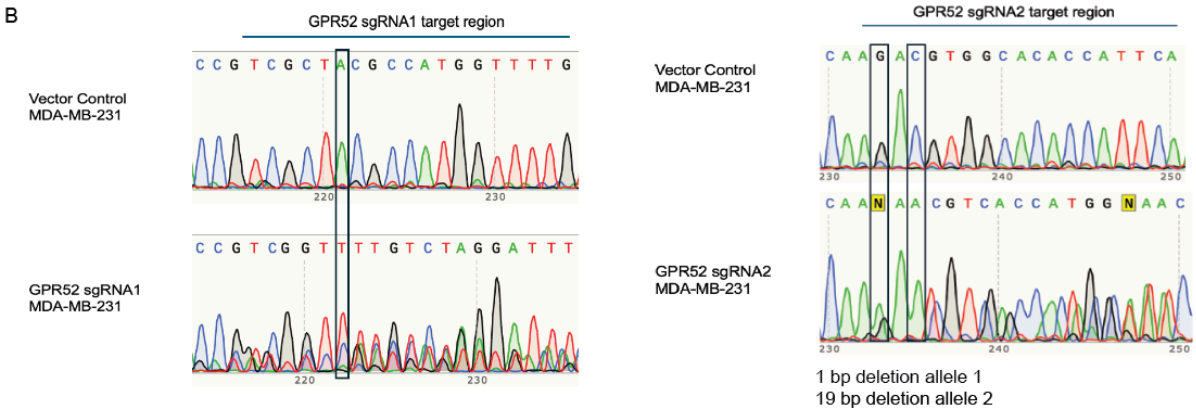
475 **Figure 7. Loss of GPR52 is associated with increased TNBC burden in zebrafish.**
476 A) Schematic of zebrafish xenograft study. B) Tg(flk1:EGFP-NLS) zebrafish at 30 hpi.
477 MDA-MB-468 cells (red) and endothelial cells (green) are visualized. Scalebar=300 μm.
478 C) Visualization of the interaction between MDA-MB-468 cells (red) and endothelial cells
479 (green) by confocal microscopy. Scalebar=10 μm. (D-E) The total number of tdTomato
480 foci per zebrafish. (F-G) The total area of tdTomato in the head and (H-I) trunk per
481 zebrafish. (J-K) The sum of the area of tdTomato in the head and trunk is expressed as
482 the total cancer area per zebrafish. VC=vector control. n=51 VC-1, n=59 GPR52-1, n=12
483 VC-2, n=12 GPR52-2. Student's t-test, P<0.05. *= *P*<0.05. ns=not significant.
484 Line=median. (L-M) Zebrafish were treated with 8 mM doxorubicin or vehicle control (VC)
485 and imaged at 30 hpi. The total area of tdTomato per zebrafish and (N-O) the area of
486 tdTomato in the zebrafish head were determined. VC=vector control. DOX=doxorubicin.
487 n=14 VC, n=24 GPR52-2. Student's t-test, P<0.05. ns=not significant. Line=median.
488
489
490
491
492
493
494
495
496
497 **Supplemental figures:**

# FIGURE S1

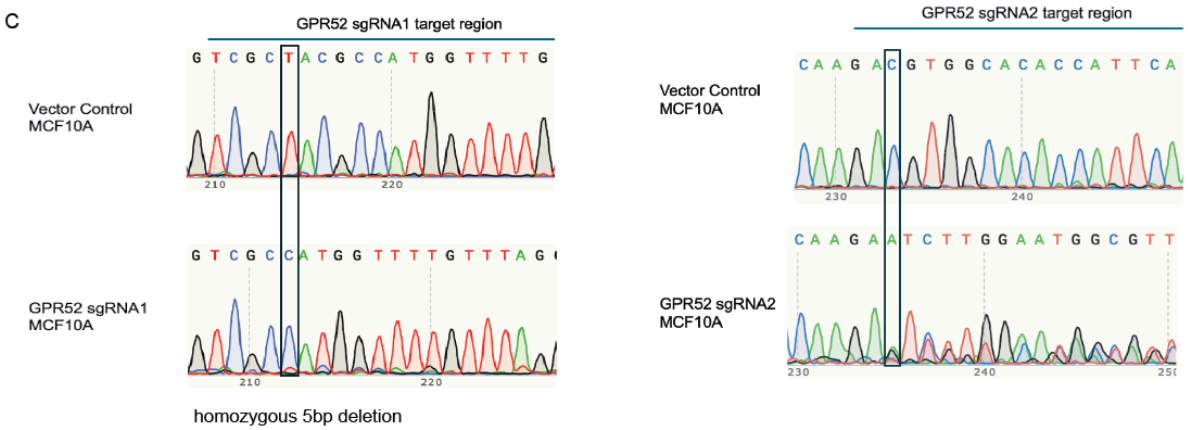
A




B

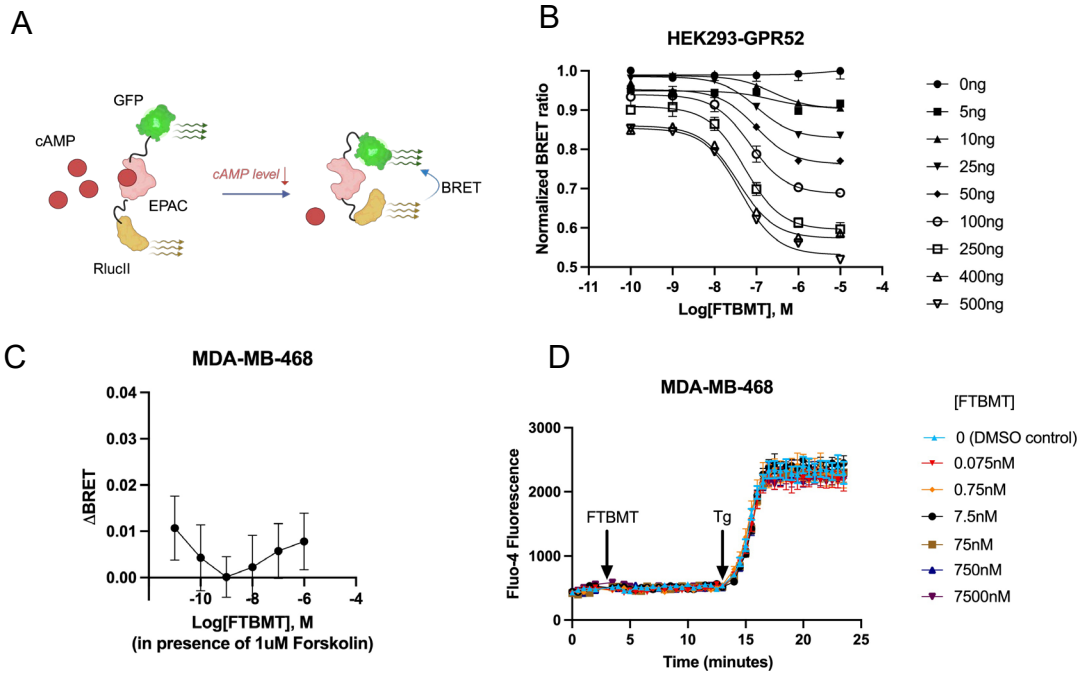



C



498 **Supplementary Figure S1. Sanger sequencing of breast cell lines to confirm indels**
499 **in the *GPR52* gene.** MDA-MB-468 (A), MDA-MB-231 (B), and MCF10A (C) were
500 transduced with Cas9 followed by a GPR52-targeted sgRNA or the empty vector. DNA
501 was extracted from cell pellets and sequenced using GPR52 primers described in
502 *Materials and Methods.* Indels in the sgRNA target regions were confirmed for cell lines
503 transduced with GPR52 sgRNA1 and sgRNA2, and the nucleotide position indicating loss
504 of the WT allele is outlined in black.
505
506

# FIGURE S2


A
GFP
cAMP
EPAC
RlucII
cAMP level
BRET
B
HEK293-GPR52
Normalized BRET ratio
Log[FTBMT], M
0ng
5ng
10ng
25ng
50ng
100ng
250ng
400ng
500ng
C
MDA-MB-468
ΔBRET
Log[FTBMT], M
(in presence of 1uM Forskolin)
D
MDA-MB-468
Fluo-4 Fluorescence
FTBMT
Tg
Time (minutes)
[FTBMT]
0 (DMSO control)
0.075nM
0.75nM
7.5nM
75nM
750nM
7500nM

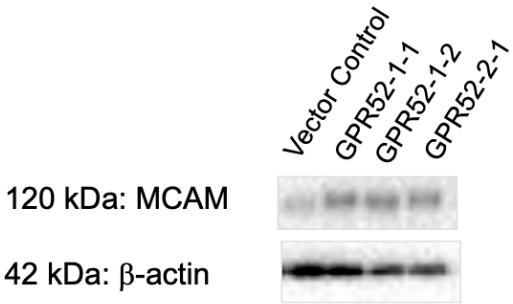
507 **Supplementary Figure S2. Treatment with GPR52 agonist FTBMT in HEK293 and**
508 **MDA-MB-468 cells**

509 A) Schematic of BRET-based GFP10-EPAC1-RlucII cAMP sensor. B) HEK293 cells
510 transfected with GPR52 and GFP-EPAC1-RlucII were treated with varying concentrations
511 of FTBMT for 15 minutes. Datapoints=mean, error bars=SEM. Data is pooled from 1-2
512 experiments, n=2 technical replicates per experiment. C) MDA-MB-468 cells transfected
513 with the EPAC1 sensor were treated with increasing concentrations of FTBMT for 30
514 minutes in the presence of 1 μM forskolin to assess for possible $G_{\alpha i}$ coupling.
515 Datapoints=mean, error bars=SEM. Data was pooled from five independent experiments,
516 n=3 technical replicates per experiment. D) MDA-MB-468 cells were plated at a density
517 of 5,000 cells/well in a 96-well plate. After 24 hours, the cell-permeable calcium indicator
518 Fluo-4 was added to each well. Baseline fluorescence measurements were then obtained
519 (t=0) and FTBMT was added at varying concentrations to achieve the final concentrations
520 listed. The plate was read at 30 second intervals for 10 minutes. 100 nM Thapsigargin
521 (Tg) was then added to all wells as a positive control for [$Ca^{2+}$] detection. The plate was
522 then read for 30 second intervals for an additional 10 minutes. Datapoints=mean, error
523 bars=SEM. n=8.

524

525

**FIGURE S3**


Vector Control
GPR52-1-1
GPR52-1-2
GPR52-2-1
120 kDa: MCAM
42 kDa: β-actin

526 **Supplementary Figure S3. Detection of MCAM in MCF10A cell lysate.** MCF10A
527 vector control, two unique isogeneic GPR52 sgRNA1-transduced cell lines, and GPR52-
528 2 cells were cultured in monolayer to 50-85% confluency. Cell lysates were collected as
529 described in *Materials and Methods* and probed for MCAM and beta-actin.
530
531 **List of abbreviations:**
532

| | |
|---|---|
| KO | Knockout |
| GPCR | G protein-coupled receptor |
| TEM | Transmission electron microscopy |
| PDL | Poly-D-lysine |
| FDA | Federal drug administration |
| EMT | Epithelial-to-mesenchymal transition |
| QPCR | Quantitative polymerase chain reaction |
| WT | Wild-type |
| CREB | cAMP response element binding protein |
| BRET | Bioluminescence resonance energy transfer |
| TCGA | The cancer genome atlas |
| FSK | Forskolin |
| ACi | Adenylyl cyclase inhibitor |
| DEG | Differentially expressed gene |

| | |
|---|---|
| HPI | Hours post-injection |
| MCAM | Melanoma cell adhesion molecule |

533

534 **Declarations:**

535 **Ethics approval and consent to participate:**

536 All animal experiments were performed with the approval of the Institutional Animal Care
537 and Use Committee (IACUC) protocol #2011-0100 from Weill Cornell Medicine with
538 animal care under the supervision of the Research Animal Resource Center. Approval
539 date: September 19, 2023.

540

541 **Consent for publication:**

542 Not applicable

543

544 **Availability of data and materials:**

545 The datasets used and/or analyzed during the current study are available from the
546 corresponding author upon reasonable request. Proteomics and phoshphopeptide-
547 enriched data were submitted to the MassIVE Proteomics Repository. Data are available
548 via MassIVE (MSV000065698, MSV000095699) or ProteomeXchange (PXD055208,
549 PXD055209) for global and phosphopeptide-enriched data, respectively.

550

551

552 **Competing interests:**

553 The authors declare that they have no competing interests

**Funding:**

This work was supported by the National Cancer Institute of the National Institutes of Health grant 1R01CA215797 (KAB), Anne Moore Breast Cancer Research Fund (KAB), Emilie Lippmann and Janice Jacobs McCarthy Research Scholar Award in Breast Cancer (KAB), and University of Kansas Cancer Center grant P30CA168524 (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 (in this format): 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.

**Author's Contributions:**

SZH conceptualized, designed, and wrote the manuscript with KAB. SZH performed the majority of the experiments in this study, except for those specified hereafter. CCA has developed CRISPR-Cas9 tools for GPR52 expression. IT performed the cancer cell injections in zebrafish. SYJ and TF performed RNA-seq analyses for the TCGA datasets. BB and OE performed RNA-seq analysis for MDA-MB-468 WT vs. GPR52 KO cells. MW performed proteomic analyses. DD and THE conducted the BRET studies. SL assisted with the tissue culture. PB optimized and provided the protocols for MCF10A culture. TE-supervised zebrafish studies. PKA, RT, and SP conducted calcium fluo-4 experiments.

576 LD provided the CRISPR plasmids, tools, and study advice. JB provided intellectual
577 advice and materials for conducting this research.
578
579 **Acknowledgements:**
580 Studies were conducted with the support and facilities provided by the Microscopy and
581 Image Analysis Core Facility and the Genomics Resources Core Facility at Weill Cornell
582 Medicine, and the Mass Spectrometry and Proteomics Core Facility at Kansas University
583 Medical Center. All schematics were created using BioRender.
584
585
586 **Materials and Methods:**
587
588 **Cell culture**
589 Human breast cell lines MDA-MB-468 (RRID:CVCL_0419), MDA-MB-231
590 (RRID:CVCL_0062), MCF7 (RRID:CVCL_0031), HS578T (CVCL_0332), T47D
591 (CVCL_0553), and MCF10A (RRID:CVCL_0598) were purchased from the American
592 Type Culture Collection (ATCC). MDA-MB-468, MDA-MB-231, MCF7, and HS578T cells
593 were cultured in Dulbecco's Modified Eagle's medium (DMEM, Gibco #11995-065)
594 supplemented with 10% fetal bovine serum (FBS) (Gibco 26140079, Sigma-Aldrich
595 F0926) and 1% penicillin-streptomycin ( complete medium). T47D cells were cultured in
596 Roswell Park Memorial Institute medium (RPMI 1640, Thermo Fisher Scientific #11875)
597 supplemented with 10% fetal bovine serum (FBS) (Gibco 26140079) and 1% penicillin-
598 streptomycin ( complete medium). MCF10A cells were cultured in Dulbecco's Modified

599 Eagle's Medium/F12 (Gibco, #11330-032) supplemented with 5% FBS, 1% penicillin-
600 streptomycin, and the following growth factors: epidermal growth factor (20 ng/ml),
601 hydrocortisone (0.5 mg/ml), cholera toxin (100 ng/ml), and insulin (10 μg/ml) (i.e.,
602 complete media). All cell lines were maintained at 37°C in a humidified atmosphere
603 containing 5% $CO_2$. All cell lines were authenticated by STR analysis and tested regularly
604 for mycoplasma contamination using the Universal Mycoplasma Detection Kit (ATCC 30-
605 1012K) and MycoAlert Mycoplasma Detection Kit (Lonza LT07-318).
606
607 **Breast cancer survival curves**
608 The KM plotter RNA-seq breast cancer web tool (www.kmplot.com) (12) was used to
609 obtain Kaplan-Meier survival curves. The associated hazard ratios, 95% confidence
610 intervals, and log-rank P-values were automatically determined using the web tool. Low
611 versus high cut-off values were determined based on the maximum segregation between
612 the groups.
613
614 **GPR52 and MCAM expression comparison in cancerous and non-cancerous**
615 **human tissue**
616 GPR52 mRNA expression in cancerous and non-cancerous tissues was analyzed across
617 solid cancer types using RNA-seq datasets from the TNMplot database and a web tool
618 (https://www.tnmplot.com) (11). The following datasets were analyzed for this analysis:
619 adrenocortical carcinoma ("adrenal"), bladder urothelial carcinoma ("bladder"), breast
620 invasive carcinoma ("breast"), colon adenocarcinoma ("colon"), esophageal carcinoma
621 ("esophageal"), kidney renal clear cell carcinoma ("renal_cc"), kidney renal papillary cell

622 carcinoma (“renal_pa”), liver hepatocellular carcinoma (“liver”), lung adenocarcinoma
623 (“lung_ac”), lung squamous cell carcinoma (“lung_sc”), ovarian serous
624 cystadenocarcinoma (“ovary”), pancreatic adenocarcinoma (“pancrease”), prostate
625 adenocarcinoma (“prostate”), rectum adenocarcinoma (“rectum”), skin cutaneous
626 melanoma (“skin”), stomach adenocarcinoma (“stomach”), testicular germ cell
627 adenocarcinoma (“testis”), thyroid carcinoma (“thyroid”), uterus corpus endometrial
628 carcinoma (“uterus_ec”). GPR52 expression between normal, tumor, and metastatic
629 breast-derived samples was compared based on a gene chip dataset available through
630 the TNM plot. Source data for GPR52 expression were collected in January 2024, and
631 the correlation between GPR52 and MCAM expression in breast tumors was determined
632 in June 2024.
633
634 **RT-qPCR**
635 Each breast cancer cell line was cultured in an individual well of a 6-well tissue culture-
636 treated plate until 50-80% confluent. Total RNA was extracted using an RNeasy Mini Kit
637 (Qiagen # 74106), and RNA purity was assessed using a Nanodrop 2000 instrument
638 (Thermo Fisher Scientific). RNA was reverse-transcribed using the qScript cDNA
639 Synthesis kit (VWR # 101414-100) based to the manufacturer’s protocol. Quantitative
640 polymerase chain reaction (PCR) was performed in triplicate from the RNA collected from
641 each well using Fast SYBR Green Master Mix (Thermo Fisher Scientific #4385612 )
642 following the standard protocol. GPR52 and RPL32 transcripts were quantified using the
643 following primers.
644

645 GPR52 F: 5'- CGGGTCTTGGACAATCCAACTC -3',

646 GPR52 R: 5' - TGCTTCCTGATCCTTCACACAC - 3',

647 RPL32 F: 5' – CAGGGTTCGTAGAAGATTCA -3',

648 RPL32 R: 5' – CTTGGAGGAAACATTGTGAGCGATC -3'.

649

650 GPR52 expression in each human breast cell line was normalized to the housekeeping
651 gene RPL32, and then normalized to the average expression of GPR52 across the human
652 breast cell lines analyzed.

653

654 **Generation of stable cell lines**

655 Transduction-based CRISPR-Cas9 gene editing was used to generate indels in *GPR52*
656 in MDA-MB-468, MDA-MB-231, and MCF10A cells. For lentivirus production, HEK293
657 cells were plated in a six-well plate and transfected with a prepared mix in 150uL DMEM
658 (with no supplements) containing 2.5 µg of pLenti-Cas9-P2A-Puro (a gift from Lukas Dow;
659 RRID:Addgene_110837), 1.25 µg of PAX2, 1.25 µg of VSV-G, and 30 µl of
660 polyethylenimine (1 mg/ml). The medium was replaced the following day and changed 36
661 h post-transfection to the target cell collection medium. MDA-MB-468, MDA-MB-231, and
662 MCF10A cells were plated in individual wells of a six-well plate and transduced with the
663 pLenti-Cas9-P2A-Puro lentivirus generated above in serum- and antibiotic-free media
664 containing 8 mg/mL polybrene for 24 h, and then selected with 1-2.5 mg/mL puromycin
665 in complete media (45).

666 The forward and reverse sgRNA cloning primers were as follows:

sgRNA1-F: CACCGCAAAACCATGGCGTAGCGA

sgRNA1-R: AAACTCGCTACGCCATGGTTTTGC

sgRNA2-F: CACCGTGAATGGTGTGCCACGTCT

sgRNA2-R: AAACAGACGTGGCACACCATTCAC

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

680 GPR52-1 F: 5'- AAACCTGGTTACCATGGTGAC -3',

681 GPR52-2 F: 5' - GATCTGGATCTACTCCTGCC - 3',

682 GPR52-1,-2 R: 5'- ATTATATAGGGGAGCCACAGC-3',

683 The amplified PCR products were subjected to Sanger sequencing (Genewiz, Azenta) to
684 identify clones with indels in GPR52 sgRNA target regions (Supplementary Fig. S1). The
685 clonal populations generated from each sgRNA were expanded and used in downstream

686 studies. A heterogeneous population of gSafe cells was used as the vector control
687 population.
688
689 **Transmission electron microscopy**
690 MDA-MB-468 cells were cultured in suspension in non-tissue culture-treated plates for
691 24-48 hours and then briefly centrifuged in 1.7mL tubes. The resulting pellet was washed
692 and fixed with 4% paraformaldehyde, 2.5% glutaraldehyde, and 0.02% picric acid in 0.1 M
693 sodium cacodylate buffer (pH 7.3). The samples were stained with uranyl acetate and
694 dehydrated using a graded ethanol series. After dehydration, the samples were covered
695 with a layer of fresh resin, and embedded molds were inserted into the resin. After
696 polymerization, the samples were cut at 200 nm for screening by light microscopy and
697 then at 65 nm to be mounted on grids for TEM under a JEOL JSM 1400 (JEOL, USA)
698 electron microscope. The camera used was a Veleta 2 K × 2 K CCD (EMSIS, GmbH,
699 Muenster). Images were taken at 2500-20,000× magnification, and each field included
700 the extracellular space, cytoplasmic area, and nucleus. Cell-cell interface length was
701 determined using 2500x images by drawing a straight line that extended the distance of
702 the cell-cell interface only between cells that exhibited extensive adhesion between their
703 plasma membrane surfaces. The cell diameter was determined using 2500x images by
704 drawing a straight line that spanned the midpoint of the aforementioned analyzed cells.
705 The free intercellular space was determined using Fiji software. The image threshold
706 intensity was set to distinguish the cells from the (white) background. The boundary
707 between the two cells was then outlined, and the fraction of the area occupied by free
708 space (met the threshold intensity) was determined.

709

710 **3D Matrigel cell culture and analysis**

711 MDA-MB-468 cells were resuspended in 1:1 growth-factor-reduced Matrigel:serum-free
712 DMEM and plated at a density of 10,000 cells/100 µL per well in a black optical-bottom
713 96-well plate (Corning Matrigel #356231). Complete medium (100 µL) was added to each
714 well after solidification of the Matrigel, and the medium was replaced every 2-3 days. After
715 10 days of culture, the cells were fixed with 10% formalin and stained with Hoechst 33342
716 (Santa Cruz Biotechnology # SC-495790). Cells were imaged using a confocal
717 microscope (Zeiss LSM880) to detect cytoplasmic tdTomato and nuclear Hoechst signals.
718 Images were then imported to Imaris Microscopy Image Analysis Software (OXFORD
719 instruments). The Imaris "Surface" function was used to determine the area per spheroid,
720 total area occupied per well, and roundness of each spheroid based on the absolute
721 intensity of the tdTomato signal. For MDA-MB-468 cells, the following formula for
722 circularity was used to determine the roundness of each spheroid.

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

723

724

725 **Matrigel invasion assays**

726 MDA-MB-468 and MDA-MB-231 cells were plated at >80% confluence in optical grade
727 96-well plates. A 50 µL solution of 1:1 growth-factor-reduced Matrigel:serum-free DMEM
728 ± 1:40 DQ-Collagen IV (Thermo Fisher Scientific) was added above the monolayer and
729 allowed to form a gel. Fifty microliters of DMEM containing 10% FBS was added above
730 the gel once it solidified. After 24 h, confocal microscopy (Zeiss LSM880) was used to
731 obtain a z-stack of tdTomato and DQ signals in each well at 10 µm intervals starting from

732 the base of the well. The Imaris “Surface” function was used to determine the surface
733 area of tdTomato foci above the minimum fluorescence intensity threshold at the
734 established z-planes. For collective invasion analyses and class B determination at
735 z=10um, the threshold intensity was auto-determined by Imaris per image. Surfaces were
736 identified based on the local contrast of the tdTomato signal with the non-fluorescent
737 background. The Imaris “Surface” function was used to quantify the area of DQ at z=10
738 µm based on a maintained absolute value threshold across samples and colocalize the
739 volume of DQ with the volume of tdTomato.  that was adjusted to maximize detection
740 across samples based on the local contrast at z=10 µm.
741
742 **Poly-D-lysine coating**
743 PDL (Santa Cruz Biotechnology #SC-136156) was resuspended in Milli-Q water at
744 1mg/mL and frozen in aliquots at -20°C. To develop coating solutions, PDL was
745 resuspended in Milli-Q water to a final concentration of 100 µg/mL and added to tissue
746 culture-treated plates to cover the surface. plates were then incubated for at least 30 min
747 at room temperature or several hours at 4°C. The plates were washed twice with PBS
748 and used for cell culture experiments.
749
750 **Western blotting and analysis**
751 Cells were cultured in monolayers under standard tissue culture conditions as described
752 above. At the time of collection, cells were scraped on ice with ice-cold PBS and
753 centrifuged at 1500-2000 rpm, and the resulting cell pellet was snap-frozen in liquid
754 nitrogen. On the day of western blotting, the cell pellet was lysed in ice-cold buffer (5 mM

755 HEPES, 137 mmol/L NaCl, 1 mmol/L $MgCl_2$, 1 mmol/L $CaCl_2$, 10 mmol/L NaF, 2 mmol/L
756 EDTA, 10 mmol/L sodium pyrophosphate, 2 mmol/L $NaVO_4$, 1% NP-40, 10% glycerol)
757 containing protease inhibitors (Thermo Fisher Scientific #78429), incubated on a rotating
758 shaker at 4°C for 15 min, and centrifuged for 20 min at 4°C at 20,000 rpm. Cell extracts
759 were denatured in buffer containing β-mercaptoethanol, run on NuPAGE 4–12% Bis-Tris
760 protein gels (Thermo Fisher Scientific), and then transferred to nitrocellulose membranes.
761 Membranes were blocked with 5% nonfat dry milk (Bio-Rad, #1706404) for one hour at
762 room temperature (RT), washed with 1X Tris-Buffered Saline, 0.1% Tween (TBST), and
763 incubated with primary antibodies at 4°C for at least 16 h. The following primary
764 antibodies were used: E-cadherin (RRID:AB_2291471), Snai1 (RRID:AB_2255011),
765 vimentin (RRID:AB_10695459), p-CREB (RRID:AB_2561044), CREB
766 (RRID:AB_310268) and MCAM (RRID:AB_2143373). The membranes were then
767 washed with TBST and incubated with secondary antibodies for 1-2 hours at room
768 temperature. The secondary antibodies used were anti-rabbit IgG HRP-linked antibody
769 (RRID:AB_2099233) and anti-mouse IgG HRP-linked antibody (RRID:AB_330924).
770 Western Lightning Plus ECL (Thermo Fisher Scientific #509049325) and ImageLab
771 software (Bio-Rad) were used for band detection. Membranes were stripped with Restore
772 PLUS Western blot stripping buffer (Thermo Fisher Scientific, #46430) for 15 min and then
773 probed for β-actin expression as a loading control using an HRP-linked β-actin antibody
774 (RRID:AB_262011).
775
776 **Zebrafish maintenance**

777 All animal experiments were performed under an approved IACUC protocol (2011-0100)
778 from Weill Cornell Medicine, with animal care under the supervision of the Research
779 Animal Resource Center. The TG(flk1:EGFP-NLS) zebrafish strain was generated by the
780 Markus Affolter Laboratory (Biozentrum, University of Basel, Switzerland)
781 (Developmental Biology 316:312-322, 2008) and kindly provided by Jesus Torres-
782 Vazquez (New York University, New York, NY). Embryos were obtained by natural
783 matings and raised at 28.5 °C with a 14-h light–10-h dark cycle. Zebrafish were
784 maintained in E3 water (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$, 1
785 ppm methylene blue).
786
787 **Zebrafish xenograft studies**
788 tdTomato-tagged MDA-MB-468 cells were passed through a 40 µm filter and
789 resuspended in PBS at a density of 400 cells/5nL. Two dpf zebrafish embryos were each
790 injected with 5nL of cancer cells in the perivitelline space or the duct of Cuvier, whichever
791 was best accessible based on the orientation of the fish. Injected embryos were evaluated
792 at two hpi and only zebrafish with detectable tdTomato distal to the injection site were
793 maintained at 32 ºC. Zebrafish were anesthetized with 0.02% tricaine and imaged at five
794 hpi and 30 hpi using a Nikon SMZ1500 microscope mounted with a Nikon DS-FI3 camera.
795 Cancer surface area per zebrafish was quantified using Imaris “Surface” function; two
796 rectangles were drawn per zebrafish to encompass the area distal to the injection site
797 excluding the yolk sac (i.e., trunk) and superior to the otolith (i.e., head), and the total
798 surface area of tdTomato determined in each rectangle was combined to calculate the
799 total cancer area. The number of detected tdTomato surfaces was used to determine the

800 number of cancer foci. Zebrafish were fixed with 4% paraformaldehyde at the study
801 endpoint. For confocal imaging, fixed zebrafish were washed with PBS and then
802 incubated for at least 12 h at 4°C in PBS containing Hoechst 33342. The zebrafish were
803 then mounted in 1% low-melt agarose and imaged using a confocal microscope (Zeiss
804 LSM880).
805 For drug treatment studies, zebrafish with detectable tdTomato distal to the yolk sac were
806 selected at two hpi, placed into individual wells of 24-well plates, and imaged at 5 hpi.
807 The zebrafish were then maintained in E3 water containing either 8 μM doxorubicin or a
808 vehicle (Milli-Q water). All zebrafish were imaged at 30 hpi and analyzed as described
809 above.
810
811 **RNA-sequencing studies and computational analyses**
812 MDA-MB-468 WT (parental) and GPR52 sgRNA1 KO cells (one mixed population and
813 two clonal populations) were cultured in 6-well plates. Total RNA was extracted using the
814 QIAzol lysis reagent (Qiagen #79306) and RNeasy Mini Kit (Qiagen #74106). Samples
815 were subjected to RNA sequencing at the Genomics Resources Core Facility (GRCF,
816 Weill Cornell Medicine). Total RNA integrity was assessed with a 2100 Bioanalyzer
817 (Agilent Technologies, Santa Clara, CA). RNA concentration was measured using a
818 NanoDrop system (Thermo Fisher Scientific, Inc., Waltham, MA, USA). RNA sample
819 library and RNA sequencing were performed by the Genomics Core Laboratory at Weill
820 Cornell Medicine with the Illumina TruSeq Stranded mRNA Sample Library Preparation
821 kit (Illumina, San Diego, CA), according to the manufacturer's instructions. The
822 normalized cDNA libraries were pooled and sequenced on an Illumina NovaSeq 6000

823 sequencer with 50 paired-end cycles. The raw sequencing reads in BCL format were
824 processed using bcl2fastq 2.19 (Illumina) for FASTQ conversion and demultiplexing.
825
826 All reads were independently aligned with STAR_2.4.0f1 for sequence alignment against
827 the human genome sequence build hg19 downloaded using the UCSC genome browser
828 and SAMTOOLS v0.1.19 for sorting and indexing reads. Cufflinks (2.0.2) was used to
829 estimate the expression values (FPKMS) and GENCODE v19 GTF files for annotation.
830 Gene counts from the htseq-count and DESeq2 Bioconductor packages were used to
831 identify DEGs. All DEGs with P-value <0.05 and false discovery rate <0.05 were uploaded
832 to Ingenuity Pathway Analysis software (Qiagen) for further analysis.
833
834 **Proteomic studies and computational analyses**
835 Parental, vector control, and GPR52-1 and -2 KO MDA-MB-468, MDA-MB-231, and
836 MCF10A cells were cultured in 10-cm tissue-culture-treated plates until 60-80% confluent.
837 Cells were then scraped on ice with ice-cold PBS and centrifuged, and the resulting cell
838 pellet was snap-frozen in liquid nitrogen. Cells were lysed by resuspension in
839 radioimmunoprecipitation assay (RIPA) buffer supplemented with protease and
840 phosphatase inhibitors. Soluble proteins were reduced by the addition of TCEP (0.5 M)
841 to a final concentration of 5mM followed by incubation at 55 $^{o}$C for 30 min. The reduced
842 samples were alkylated by the addition of 375 mM iodoacetamide to a final concentration
843 of 10 mM, followed by incubation in the dark at room temperature for 30 min. Ice-cold
844 acetone was added to each sample at a volume ratio of 5:1. The samples were vortexed
845 and stored at -20 $^{o}$C overnight. After precipitation, the samples were centrifuged at 14,000

846 × *g* and 4 ºC for 30 min to pellet the proteins. The supernatant was removed, and
847 the pellet was air-dried on a benchtop for 10 min. Proteins were resuspended in 50 mM
848 TEAB (pH 8) containing 2 mM $CaCl_2$. Trypsin was added (500 ng), and the proteins were
849 allowed to digest overnight at 37 ºC with shaking at 500 RPM (Thermomixer, Eppendorf).
850 Digestion was quenched by adding 10% formic acid to a final concentration of 1%.
851 Digested samples were centrifuged at 10,000 × *g* for 10 min to remove particulates, and
852 the supernatant was transferred to a fresh tube and stored at -20 ºC until phosphopeptide
853 enrichment.
854
855 The samples were enriched for phosphorylated peptides using SMOAC. Briefly, digested
856 samples were first enriched using the High Select™ Phosphopeptide Enrichment Kit
857 (Thermo Fisher Scientific) following the manufacturer's protocol. The flow-through was
858 applied to the High Select™ Fe-NTA Phosphopeptide Enrichment Kit (Thermo Fisher
859 Scientific), following the manufacturer's protocol. After the second enrichment, the flow-
860 through became the global, unenriched sample, and the elutes from both kits were pooled
861 to generate the phosphopeptide-enriched fraction. The peptide concentration was
862 measured at 205 nm using a NanoDrop spectrophotometer (Thermo Scientific).
863
864 Samples were injected using the Vanquish Neo (Thermo Fisher Scientific) nano-UPLC
865 onto a C18 trap column (0.3 mm x 5 mm, 5 µm C18) using pressure loading. Peptides
866 were eluted onto a separation column (PepMap™ Neo, 75 µm × 150 mm, 2 µm C18
867 particle size, Thermo Fisher Scientific) prior to elution directly onto the mass
868 spectrometer. Briefly, peptides were loaded and washed for 5 min at a flow rate of 0.350

869 μL/min in 2% B (mobile phase A: 0.1% formic acid in water; mobile phase B: 80% ACN,
870 0.1% formic acid in water). Peptides were eluted over 100 minutes from 2-25% mobile
871 phase B before ramping to 40% B for 20 min. The column was washed for 15 min at
872 100% B before reequilibration at 2% B for the next injection. The nano-LCs were directly
873 interfaced with an Orbitrap Ascend Tribrid mass spectrometer (Thermo Fisher Scientific)
874 using a silica emitter (20 μm i.d., 10 cm) equipped with a high-field asymmetric ion mobility
875 spectrometry (FAIMS) source. Data were collected by data-dependent acquisition with
876 the intact peptide detected in the Orbitrap at 120,000 resolving power from to 375-
877 1500 *m/z*. Peptides with charge +2-7 were selected for fragmentation by higher-energy
878 collision dissociation (HCD) at 28% NCE and were detected in the ion trap using a rapid
879 scan rate (global) or Orbitrap at a resolving power of 30,000 (enriched). Dynamic
880 exclusion was set to 60s after one instance. The mass list is shared among the FAIMS
881 compensation voltages. FAIMS voltages were set at -45 (1.4 s), -60 (1 s), -75 (0.6 s) CV
882 for a total duty cycle time of 3s. Source ionization was set at 1700 V with an ion transfer
883 tube temperature of 305 $^{o}$C. Raw files were searched against the human protein database
884 downloaded from UniProt on 05-05-2023 using SEQUEST in Proteome Discoverer 3.0.
885 The abundances, abundance ratios, and P-values were exported to Microsoft Excel for
886 further analysis. All proteins with differential GPR52 KO vs. WT abundance ratios (P-
887 value <0.05, false discovery rate <0.05) were uploaded to the Ingenuity Pathway Analysis
888 software (Qiagen) for further analysis of each of the three cell lines. KEA3 software was
889 used to rank the kinase activity of the combined vector control and GPR52 KO populations
890 across the three cell lines based on their respective phosphoproteomes (20).
891

892 **Analysis of TCGA-BRCA transcriptomic dataset**

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

907

908 **BRET biosensor studies**

909 HEK 293 and MDA-MB-468 cells were maintained in DMEM with 10% FBS, respectively,
910 and cultured at 37°C and 5% $CO_2$. For transfection, cells were plated at a density of
911 $2.5 \times 10^5$ cells (HEK293) and $8 \times 10^5$ cells (MDA-MB-468) per well in 6-well plates
912 (Thermo Scientific, 140,675). On the day of transfection, the medium was replaced with
913 DMEM containing 2.5% FBS (HEK 293) and Opti-MEM (MDA-MB-468) (Thermo Fisher
914 Scientific #31985062) without antibiotics. HEK293 cells were transfected using

915 Lipofectamine 2000 (Thermo Fisher Scientific, #11668027). The GFP10-Epac-RlucII in
916 pcDNA3.1 (generously provided by Dr. Michel Bouvier, Université de Montréal) and
917 GPR52 cDNA ORF Clone GPR52 (Sino Biological # HG25891-UT) in pcDNA3.1, with
918 added N-terminal flag tag, were used in these studies.
919
920 For HEK293 cells, per well: 1.5 µg of plasmid DNA (total) was added to 100 µl DMEM (no
921 serum, no antibiotics) in a 1.5 mL Eppendorf tube. In another tube, 3 µl of Lipofectamine
922 2000 was added to 100 µl of DMEM (no serum or antibiotics).
923 For MDA-MB-468 cells, per well: 2.5 µg of plasmid DNA (total) was added to 100 µl Opti-
924 MEM (no serum, no antibiotics) in a 1.5 mL Eppendorf tube. In another tube, 5 µl of
925 Lipofectamine 2000 or 3000 was added to 100 µl of Opti-MEM (no serum or antibiotics).
926
927 After 5 min, the contents of the tubes were combined and gently mixed or vortexed for
928 each cell line. The mixture was incubated for 20-30 minutes at room temperature. The
929 200 µL DNA:Lipofectamine mixture was then added dropwise per well, and the plate was
930 swirled gently and incubated at 37 °C and 5% $CO_2$ for 4-5 hours. The medium was then
931 replaced with a complete medium.
932
933 Cells were detached the day following transfection with 0.25% trypsin–EDTA (Wisent)
934 and plated at a density of $3 \times 10^4$ cells/well in a poly-l-ornithine (Sigma-Aldrich)-coated
935 96-well white bottom plate (Thermo Scientific, 236,105) for BRET analysis. After 24 h of
936 incubation in the 96-well plate, the media was removed and cells were washed once with
937 Krebs buffer (146 mM NaCl, 42 mM MgCl2, 10 mM HEPES pH 7.4, 1 g/L D-glucose) and

938 incubated for 2 h at room temperature in Krebs buffer. Coelenterazine 400A (Cedarlane)
939 was added, to a final concentration of 2.5 μM, to each well, followed by incubation for
940 5 min prior to the basal reading. FTBMT was added to the indicated final concentration
941 and the plate was read after 15 or 30 min of stimulation. MDA-MB-468 cells were co-
942 stimulated with 1 mM forskolin.
943
944
945 The BRET ratios were calculated as the emission at 515 nm and 410 nm. For all
946 experiments, ΔBRET = (BRET ratio from cells stimulated with FTBMT and forskolin) –
947 (BRET ratio from vehicle treated with forskolin). Three technical replicates were used for
948 all the treatments. The BRET experiments were performed using a Tristar2 plate reader
949 (Berthold. Technologies GmbH and Co. KG). The normalized BRET ratio was
950 computed as $BRET_{Stimuated}/BRET_{Vehicle.}$ Dose-response curves were plotted using non-
951 linear regression. Data are presented as the mean ± standard error (SE). For MDA-MB-
952 468 cells, the experiment was conducted five times, and the aggregate data from all
953 experiments are presented. For HEK 293 cells, the experiment was conducted twice, and
954 the combined data from the two experiments are presented.
955
956 **Intracellular calcium measurements**
957 MDA-MB-468 cells were seeded at a density of 5000 cells/well in a 96-well plate. After 24
958 h, the cell-permeable calcium indicator Fluo-4 (Thermo Fisher Scientific, #F14201) was
959 added to each well. cells were incubated for 45 min at 37 °C, then for 15 min at room
960 temperature. Baseline fluorescence measurements were obtained (t=0), and FTBMT was

961 added to select the wells. The plate was read at 30 second intervals for 10 min.
962 Thapsigargin (Sigma-Aldrich) was added to cells as a positive control, and the plate was
963 then read at 30 second intervals for an additional 10 min (47).
964

965 **Drug preparation**

966 Forskolin (Santa Cruz Biotechnology #SC-3562) and SQ22536 (Santa Cruz
967 Biotechnology #SC-201572) were resuspended in ethanol (1 mM) and DMSO (14 mM),
968 and frozen. SQ22536 was aliquoted to avoid freeze-thaw cycles. Ethanol and DMSO
969 were used as vehicle controls in all experiments that incorporated these compounds.
970

971 **Statistical analysis**

972 All data were expressed as individual data points with a line at the median value or as the
973 mean ± SEM, unless indicated otherwise. The groups were compared using statistical
974 tests for significance, as described in the figure legends. Statistical significance was set
975 at $P < 0.05$. Statistical tests were performed using the GraphPad Prism software.
976

977 The data generated in this study are available upon request from the corresponding
978 author.
979

980 **References:**

981

982 1. Hagemeister FB, Jr., Buzdar AU, Luna MA, Blumenschein GR. Causes of death in
983 breast cancer: a clinicopathologic study. Cancer. 1980;46(1):162-7.

984 2. Schuster E, Taftaf R, Reduzzi C, Albert MK, Romero-Calvo I, Liu H. Better
985 together: circulating tumor cell clustering in metastatic cancer. Trends Cancer.
986 2021;7(11):1020-32.
987 3. Bakir B, Chiarella AM, Pitarresi JR, Rustgi AK. EMT, MET, Plasticity, and Tumor
988 Metastasis. Trends Cell Biol. 2020;30(10):764-76.
989 4. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The
990 sequence of the human genome. Science. 2001;291(5507):1304-51.
991 5. Hanson MA, Stevens RC. Discovery of new GPCR biology: one receptor structure
992 at a time. Structure. 2009;17(1):8-14.
993 6. Hamm HE, Gilchrist A. Heterotrimeric G proteins. Curr Opin Cell Biol.
994 1996;8(2):189-96.
995 7. Chaudhary PK, Kim S. An Insight into GPCR and G-Proteins as Cancer Drivers.
996 Cells. 2021;10(12).
997 8. Komatsu H, Maruyama M, Yao S, Shinohara T, Sakuma K, Imaichi S, et al.
998 Anatomical transcriptome of G protein-coupled receptors leads to the identification of a
999 novel therapeutic candidate GPR52 for psychiatric disorders. PLoS One.
1000 2014;9(2):e90134.
1001 9. Nishiyama K, Suzuki H, Harasawa T, Suzuki N, Kurimoto E, Kawai T, et al. FTBMT,
1002 a Novel and Selective GPR52 Agonist, Demonstrates Antipsychotic-Like and
1003 Procognitive Effects in Rodents, Revealing a Potential Therapeutic Agent for
1004 Schizophrenia. J Pharmacol Exp Ther. 2017;363(2):253-64.
1005 10. Yao Y, Cui X, Al-Ramahi I, Sun X, Li B, Hou J, et al. A striatal-enriched intronic
1006 GPCR modulates huntingtin levels and toxicity. Elife. 2015;4.
1007 11. Bartha A, Gyorffy B. TNMplot.com: A Web Tool for the Comparison of Gene
1008 Expression in Normal, Tumor and Metastatic Tissues. Int J Mol Sci. 2021;22(5).
1009 12. Gyorffy B. Survival analysis across the entire transcriptome identifies biomarkers
1010 with the highest prognostic power in breast cancer. Comput Struct Biotechnol J.
1011 2021;19:4101-9.
1012 13. Au CC, Furness JB, Britt K, Oshchepkova S, Ladumor H, Soo KY, et al. Three-
1013 dimensional growth of breast cancer cells potentiates the anti-tumor effects of unacylated
1014 ghrelin and AZP-531. Elife. 2020;9.
1015 14. Sriram K, Insel PA. G Protein-Coupled Receptors as Targets for Approved Drugs:
1016 How Many Targets and How Many Drugs? Mol Pharmacol. 2018;93(4):251-8.
1017 15. Hughes CS, Postovit LM, Lajoie GA. Matrigel: a complex protein mixture required
1018 for optimal growth of cell culture. Proteomics. 2010;10(9):1886-90.
1019 16. Guan JL, Trevithick JE, Hynes RO. Fibronectin/integrin interaction induces
1020 tyrosine phosphorylation of a 120-kDa protein. Cell Regul. 1991;2(11):951-64.
1021 17. Kornberg L, Earp HS, Parsons JT, Schaller M, Juliano RL. Cell adhesion or integrin
1022 clustering increases phosphorylation of a focal adhesion-associated tyrosine kinase. J
1023 Biol Chem. 1992;267(33):23439-42.
1024 18. Jedeszko C, Sameni M, Olive MB, Moin K, Sloane BF. Visualizing protease activity
1025 in living cells: from two dimensions to four dimensions. Curr Protoc Cell Biol.
1026 2008;Chapter 4:Unit 4 20.
1027 19. Gonzalez GA, Montminy MR. Cyclic AMP stimulates somatostatin gene
1028 transcription by phosphorylation of CREB at serine 133. Cell. 1989;59(4):675-80.

1029 20. Kuleshov MV, Xie Z, London ABK, Yang J, Evangelista JE, Lachmann A, et al.
1030 KEA3: improved kinase enrichment analysis via data integration. Nucleic Acids Res.
1031 2021;49(W1):W304-W16.
1032 21. Kang TH, Park DY, Kim W, Kim KT. VRK1 phosphorylates CREB and mediates
1033 CCND1 expression. J Cell Sci. 2008;121(Pt 18):3035-41.
1034 22. Martin RD, Sun Y, Bourque K, Audet N, Inoue A, Tanny JC, et al. Receptor- and
1035 cellular compartment-specific activation of the cAMP/PKA pathway by alpha(1)-
1036 adrenergic and ETA endothelin receptors. Cell Signal. 2018;44:43-50.
1037 23. An Y, Wei N, Cheng X, Li Y, Liu H, Wang J, et al. MCAM abnormal expression and
1038 clinical outcome associations are highly cancer dependent as revealed through pan-
1039 cancer analysis. Brief Bioinform. 2020;21(2):709-18.
1040 24. Wang J, Tang X, Weng W, Qiao Y, Lin J, Liu W, et al. The membrane protein
1041 melanoma cell adhesion molecule (MCAM) is a novel tumor marker that stimulates
1042 tumorigenesis in hepatocellular carcinoma. Oncogene. 2015;34(47):5781-95.
1043 25. Zhang X, Wang Z, Kang Y, Li X, Ma X, Ma L. MCAM expression is associated with
1044 poor prognosis in non-small cell lung cancer. Clin Transl Oncol. 2014;16(2):178-83.
1045 26. Teng Y, Xie X, Walker S, White DT, Mumm JS, Cowell JK. Evaluating human
1046 cancer cell metastasis in zebrafish. BMC Cancer. 2013;13:453.
1047 27. Ren J, Liu S, Cui C, Ten Dijke P. Invasive Behavior of Human Breast Cancer Cells
1048 in Embryonic Zebrafish. J Vis Exp. 2017(122).
1049 28. Blum Y, Belting HG, Ellertsdottir E, Herwig L, Luders F, Affolter M. Complex cell
1050 rearrangements during intersegmental vessel sprouting and vessel fusion in the zebrafish
1051 embryo. Dev Biol. 2008;316(2):312-22.
1052 29. Campenni M, May AN, Boddy A, Harris V, Nedelcu AM. Agent-based modelling
1053 reveals strategies to reduce the fitness and metastatic potential of circulating tumour cell
1054 clusters. Evol Appl. 2020;13(7):1635-50.
1055 30. Tacar O, Sriamornsak P, Dass CR. Doxorubicin: an update on anticancer
1056 molecular action, toxicity and novel drug delivery systems. J Pharm Pharmacol.
1057 2013;65(2):157-70.
1058 31. Chemotherapy for Breast Cancer [Internet]. American Cancer Society. 2021.
1059 Available from: https://www.cancer.org/cancer/types/breast-
1060 cancer/treatment/chemotherapy-for-breast-cancer.html. Accessed July 2024.
1061 32. Gopal U, Monroe JD, Marudamuthu AS, Begum S, Walters BJ, Stewart RA, et al.
1062 Development of a Triple-Negative Breast Cancer Leptomeningeal Disease Model in
1063 Zebrafish. Cells. 2023;12(7).
1064 33. Qu Y, Han B, Yu Y, Yao W, Bose S, Karlan BY, et al. Evaluation of MCF10A as a
1065 Reliable Model for Normal Human Mammary Epithelial Cells. PLoS One.
1066 2015;10(7):e0131285.
1067 34. Vilchez Mercedes SA, Bocci F, Levine H, Onuchic JN, Jolly MK, Wong PK.
1068 Decoding leader cells in collective cancer invasion. Nat Rev Cancer. 2021;21(9):592-604.
1069 35. Kumar S, Kapoor A, Desai S, Inamdar MM, Sen S. Proteolytic and non-proteolytic
1070 regulation of collective cell invasion: tuning by ECM density and organization. Sci Rep.
1071 2016;6:19905.
1072 36. Han W, Chen S, Yuan W, Fan Q, Tian J, Wang X, et al. Oriented collagen fibers
1073 direct tumor cell intravasation. Proc Natl Acad Sci U S A. 2016;113(40):11208-13.

1074 37. Papadaki MA, Mala A, Merodoulaki AC, Vassilakopoulou M, Mavroudis D, Agelaki
1075 S. Investigating the Role of CTCs with Stem/EMT-like Features in Metastatic Breast
1076 Cancer Patients Treated with Eribulin Mesylate. Cancers (Basel). 2022;14(16).
1077 38. Fleming A, Diekmann H, Goldsmith P. Functional characterisation of the
1078 maturation of the blood-brain barrier in larval zebrafish. PLoS One. 2013;8(10):e77548.
1079 39. Rousselle C, Clair P, Lefauconnier JM, Kaczorek M, Scherrmann JM, Temsamani
1080 J. New advances in the transport of doxorubicin through the blood-brain barrier by a
1081 peptide vector-mediated strategy. Mol Pharmacol. 2000;57(4):679-86.
1082 40. Insel PA, Murray F, Yokoyama U, Romano S, Yun H, Brown L, et al. cAMP and
1083 Epac in the regulation of tissue fibrosis. Br J Pharmacol. 2012;166(2):447-56.
1084 41. Wang P, Felsing DE, Chen H, Stutz SJ, Murphy RE, Cunningham KA, et al.
1085 Discovery of Potent and Brain-Penetrant GPR52 Agonist that Suppresses
1086 Psychostimulant Behavior. J Med Chem. 2020;63(22):13951-72.
1087 42. Masuho I, Kise R, Gainza P, Von Moo E, Li X, Tany R, et al. Rules and
1088 mechanisms governing G protein coupling selectivity of GPCRs. Cell Rep.
1089 2023;42(10):113173.
1090 43. Wu Z, Wu Z, Li J, Yang X, Wang Y, Yu Y, et al. MCAM is a novel metastasis marker
1091 and regulates spreading, apoptosis and invasion of ovarian cancer cells. Tumour Biol.
1092 2012;33(5):1619-28.
1093 44. Chen J, Dang Y, Feng W, Qiao C, Liu D, Zhang T, et al. SOX18 promotes gastric
1094 cancer metastasis through transactivating MCAM and CCL7. Oncogene.
1095 2020;39(33):5536-52.
1096 45. Zafra MP, Schatoff EM, Katti A, Foronda M, Breinig M, Schweitzer AY, et al.
1097 Optimized base editors enable efficient editing in cells, organoids and mice. Nat
1098 Biotechnol. 2018;36(9):888-93.
1099 46. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al.
1100 Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods.
1101 2015;12(2):115-21.
1102 47. Wrennall JA, Ahmad S, Worthington EN, Wu T, Goriounova AS, Voeller AS, et al.
1103 A SPLUNC1 Peptidomimetic Inhibits Orai1 and Reduces Inflammation in a Murine
1104 Allergic Asthma Model. Am J Respir Cell Mol Biol. 2022;66(3):271-82.
1105