Oncogene. Author manuscript; available in PMC 2017 December 15.

Published in final edited form as:

Oncogene. 2017 June 15; 36(24): 3490–3503. doi:10.1038/onc.2016.492.

The Cdc42/Rac1 regulator CdGAP is a novel E-cadherin transcriptional co-repressor with Zeb2 in breast cancer

Yi He[1], **Jason J. Northey**[2,9], **Ariane Pelletier**[3], **Zuzana Kos**[4], **Liliane Meunier**[5], **Benjamin Haibe-Kains**[6,7,8], **Anne-Marie Mes-Masson**[5], **Jean-François Côté**[3], **Peter M. Siegel**[2], and **Nathalie Lamarche-Vane**[1]

[1]Cancer Research Program, Research Institute of the McGill University Health Center, Department of Anatomy and Cell Biology, McGill University, Montreal, Quebec, Canada

[2]Goodman Cancer Research Centre, McGill University, Montreal, Quebec

[3]Institut de recherches cliniques de Montréal, Montreal, Quebec

[4]Department of Pathology and Laboratory Medicine, University of Ottawa, Ottawa, Ontario, Canada

[5]Centre de recherche du Centre Hospitalier de l'Université de Montréal (CR/CHUM), Montreal, Quebec

[6]Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario

[7]Department of Medical Biophysics, University of Toronto, Toronto, Ontario

[8]Department of Computer Science, University of Toronto, Toronto, Ontario

Abstract

The loss of E-cadherin causes dysfunction of the cell-cell junction machinery, which is an initial step in epithelial-to-mesenchymal transition (EMT), facilitating cancer cell invasion and the formation of metastases. A set of transcriptional repressors of *E-cadherin (CDH1)* gene expression, including Snail1, Snail2 and Zeb2 mediate E-cadherin down-regulation in breast cancer. However, the molecular mechanisms underlying the control of E-cadherin expression in breast cancer progression remain largely unknown. Here, by using global gene expression approaches, we uncover a novel function for Cdc42 GTPase-activating protein (CdGAP) in the regulation of expression of genes involved in EMT. We found that CdGAP used its proline-rich domain to form a functional complex with Zeb2 to mediate the repression of E-cadherin expression in ErbB2-transformed breast cancer cells. Conversely, knockdown of CdGAP expression led to a decrease of the transcriptional repressors Snail1 and Zeb2, and this correlated with an increase in E-cadherin levels, restoration of cell-cell junctions, and epithelial-like

Corresponding Author: Nathalie Lamarche-Vane, RI-MUHC, 1001 Boul. Décarie, Room E02.6230, Montreal, Quebec, Canada H4A 3J1 Tel: 514-934-1934 ext 76166 Nathalie.lamarche@mcgill.ca.

[9]Current address: Department of Surgery, University of California San Francisco, California, USA

Data deposition: The RNA sequencing data have been deposited in the NCBI bioproject database (accession code GSE86301).

Conflict of interest The authors declare no conflict of interest.

morphological changes. *In vivo*, loss of CdGAP in ErbB2-transformed breast cancer cells impaired tumor growth and suppressed metastasis to lungs. Finally, CdGAP was highly expressed in basal-type breast cancer cells, and its strong expression correlated with poor prognosis in breast cancer patients. Together, these data support a previously unknown nuclear function for CdGAP where it cooperates in a GAP-independent manner with transcriptional repressors to function as a critical modulator of breast cancer through repression of E-cadherin transcription. Targeting Zeb2-CdGAP interactions may represent novel therapeutic opportunities for breast cancer treatment.

Keywords

epithelial-to-mesenchymal transition; E-cadherin; CdGAP/ARHGAP31; transcriptional repressor; breast tumorigenesis

Introduction

Metastasis is the leading cause of death in breast cancer patients. The epithelial-to-mesenchymal transition (EMT) plays a crucial role in metastasis and is highly critical for tumor cell dissemination. One of the hallmarks of EMT is the loss of E-cadherin expression.[1, 2] Therefore, defining the regulatory mechanisms of E-cadherin expression is essential to develop more effective therapeutic strategies to control metastatic cell migration.

Overexpression of the ErbB2 receptor tyrosine kinase has been widely associated with poor prognosis in breast cancer.[3, 4] ErbB2-transformed mammary epithelial cells form aggressively growing breast tumors that metastasize to the lungs.[5] A quantitative expression profile of Rho GTPases and their regulators in ErbB2-induced mouse breast tumors has identified Rac1 and its negative regulator Cdc42 GTPase-activating protein (CdGAP, also known as ARHGAP31) as the major GTPase and RhoGAP expressed in these tumors.[6] CdGAP regulates both Cdc42 and Rac1 activities, but not RhoA.[7, 8] CdGAP is also a substrate of ERK/GSK-3 and mediates cross talk between the Ras/MAP kinase pathway and the regulation of Rac1 activity.[8, 9] Previous studies have shown that CdGAP is a serum-inducible gene [9] and gain-of-function mutations in the *CdGAP* gene have been found in patients with the rare developmental Adams-Oliver syndrome (AOS), characterized by the combination of aplasia cutis congenita (ACC) and terminal transverse limb defects (TTLD).[10, 11] Importantly, CdGAP is required for transforming growth factor β (TGFβ)- and ErbB2-induced breast cancer cell motility and invasion.[12] Furthermore, a complete loss of E-cadherin expression was impaired in CdGAP-depleted cells during TGFβ-stimulated EMT.[12] However, the mechanism by which CdGAP regulates E-cadherin expression remains unknown. In this study, we report a previously uncharacterized GAP-independent role for CdGAP. CdGAP forms a functional protein complex with the transcriptional repressor Zeb2 to regulate E-cadherin expression in a GAP-independent manner. We correlate this nuclear function with the ability of CdGAP to promote ErbB2-mediated tumor growth and metastasis to the lungs. In addition, high expression of CdGAP correlates with poor prognosis in breast cancer patients. Taken together, this work demonstrates that CdGAP acts as a positive modulator of breast tumorigenesis, offering novel therapeutic perspectives for the treatment of breast cancer.

Results

Identification of a CdGAP-null gene signature in the TGFβ signaling pathway

To define how CdGAP regulates E-cadherin expression, we generated CdGAP-deficient stable pooled ErbB2-expressing mouse mammary tumor cells (Supplementary Figure 1a and b) and performed next generation RNA sequencing (RNA-seq) on these cells, which revealed 1694 differentially expressed genes (fold change>1.6; Adjusted *P* value<0.01; of −16,000 transcripts sequenced) (Supplementary Figure 2a, Supplementary Table 1). Global analysis of the expression data revealed genes linked to the TGFβ pathway to be associated with the depletion of CdGAP, including a subset of genes encoding the transcriptional factors Snail1 (ref. 13), Zeb2 (ref. 14), Twist2, ID2 and TGF*β* target genes, including E-cadherin (*Cdh1*), occludin (*Ocln*), and fibronectin1 (*Fn1*) (Figures 1a and b). Our analysis revealed that molecular pathways critical for cancer metastasis, including cell adhesion, basement membrane, angiogenesis, and cell junction were the most significantly affected biological processes in CdGAP-depleted cells (Figure 1c). Differential mRNA expression of *Cdh1*, *Snai1* and *Zeb2* was validated by Quantitative PCR (Q-PCR) and protein level by western blotting (Figures 2a–d). Moreover, increases of *Fn1* and *Ocln* mRNA levels were confirmed by Q-PCR, while *Zeb1* mRNA showed no significant change in CdGAP-depleted cells (Supplementary Figure 2b).

In good agreement with the changes observed in the expression of TGFβ target genes, the characterization of CdGAP-depleted cells revealed epithelial-like morphological changes. In contrast to diffuse cytoplasmic E-cadherin localization in control breast cancer cells, E-cadherin staining was enriched at cell-cell junctions in CdGAP-depleted cells (Figure 3a). E-cadherin fluorescence intensity measured along a 7-μm segment showed a shift of E-cadherin to the cell periphery in CdGAP-depleted cells compared to control cells (Figure 3b). Actin enrichment characterized by membrane ruffles was also observed at the leading edge of CdGAP-depleted cells, even though active Rac-GTP levels were not significantly altered compared to control cells (Figures 3c and d). Therefore, these results demonstrate that the loss of CdGAP restores E-cadherin localization at cell-cell junctions in breast cancer cells.

CdGAP represses the E-cadherin promoter activity in a GAP-independent manner

We next hypothesized that CdGAP could directly repress E-cadherin expression through a transcriptional mechanism. To accomplish this, we employed luciferase reporter constructs controlled by E-cadherin promoter that were introduced into ErbB2-expressing breast cancer cells or in HEK293 cells. In agreement with our hypothesis, CdGAP-depleted breast cancer cells showed a 5-fold increase in E-cadherin promoter activity compared to control cells (Figures 4a and b). The expression of human CdGAP (hCdGAP) in CdGAP-depleted cells was able to partially restore E-cadherin transcriptional repression, without significantly modifying the endogenous levels of Zeb2 expression when compared to cells transfected with empty vector (Figures 4a and b). However, 48 hours post-transfection, the expression of hCdGAP was able to restore the levels of Zeb2 in CdGAP-depleted cells comparable to control cells (Figure 4c). CdGAP consists of an N-terminal GAP domain, a basic central region followed by a proline-rich domain (PRD).[12] Interestingly, expression of only the PRD

or a version of CdGAP lacking the GAP domain (CdGAP*Δ*GAP) was sufficient to mediate E-cadherin repression (Figures 4a and b). In contrast, truncated CdGAP (1-683), which is expressed in patients with Adams-Oliver Syndrome,[10] did not rescue E-cadherin repression in CdGAP-deficient cells (Figures 4a and b). Consistent with this, overexpression of CdGAP or the PRD in HEK293 cells showed a significant 2-fold reduction in E-cadherin promoter activity whereas CdGAP (1-683) was unable to repress E-cadherin (Figures 4d and e). Zeb2 expression in HEK293 cells repressed E-cadherin[15] whereas co-expression of Zeb2 and CdGAP did not further increase E-cadherin repression compared to Zeb2 alone (Figures 4f and g). These results suggest that Zeb2 and CdGAP do not synergistically repress E-cadherin promoter activity. In addition, Zeb2 was able to repress E-cadherin promoter activity in both control (shCON) and CdGAP-depleted cells (shCdGAP) compared to their respective E.V. controls, although Zeb2 expression in CdGAP-depleted cells was 3-fold less efficient to repress E-cadherin compared to control cells expressing Zeb2 (p value = 0.043) (Figures 4h and i). Nevertheless, this is suggesting that Zeb2 can repress E-cadherin promoter activity independently of CdGAP. Furthermore, similar to Zeb2, CdGAP expression was less efficient to repress the E-box mutated E-cadherin promoter in HEK293 cells (Figures 4j and k), demonstrating that CdGAP requires the E-box elements binding to Zeb2 to fully repress E-cadherin transcription.

CdGAP interacts with Zeb2 and the promoter region of the *Cdh1* gene in breast cancer cells

We next performed a series of experiments to mechanistically address how CdGAP functions, in concert with Zeb2, to suppress E-cadherin expression. Endogenous CdGAP associated with Zeb2 in ErbB2-expressing breast cancer cells (Figure 5a). To delineate the regions within CdGAP that enable the association with Zeb2, CdGAP deletion mutants were expressed with Flag-Zeb2 in HEK293 cells and the association was assessed by co-immunoprecipitation. CdGAP, CdGAP-PRD or CdGAP-ΔGAP but not CdGAP (1-683) associated with Zeb2 (Figure 5b). Thus, these results demonstrate that an intact PRD is required to suppress E-cadherin expression and mediate the interaction between CdGAP and Zeb2.

Since Zeb2 localizes exclusively in the nucleus in several cell lines,[16] we next determined whether Zeb2 influences the subcellular localization of GFP-CdGAP expressed in HEK293 cells by confocal fluorescence microscopy (Figures 5c and d and Supplementary Figure 3) and subcellular fractionation experiments (Figure 5e). We found three patterns of GFP-CdGAP cellular localization: 57.78 (± 0.48)% of GFP-CdGAP-expressing cells displayed nuclear localization of CdGAP, while 18.75 (± 0.78)% of these cells exhibited CdGAP localization in the cytoplasm. Finally, 23.77 (± 1.21)% of GFP-CdGAP-expressing cells showed a mixed nuclear and cytoplasmic localization (Figures 5c and d). However, when GFP-CdGAP and myc-Zeb2 were co-expressed in cells, we observed a significant relocalization of GFP-CdGAP from the cytoplasm to the nucleus. Indeed, 73.35% of GFP-CdGAP and Myc-Zeb2 co-expressing cells displayed nuclear CdGAP localization, while cytoplasmic localization was reduced to 5.58% of co-expressing cells (Figure 5d and Supplementary Figure 3). Accordingly, the majority of GFP-CdGAP was detected in the nuclear fraction following subcellular fractionation experiments and overexpression of Myc-

Zeb2 increased the levels of GFP-CdGAP in the nuclear fraction (Figure 5e). We next assessed whether CdGAP interacts with the E-cadherin promoter by binding to the central core 5′-CACCTG-3′ (E-box) element[17] using chromatin immunoprecipitation (ChIP) assays. CdGAP interacted with the E-cadherin promoter and co-expression with Zeb2 did not further increase the binding to the promoter (Figure 5f). Taken together, these results demonstrate that CdGAP localizes with Zeb2 in the nucleus and is capable of interacting with the E-cadherin proximal promoter region to repress transcription.

CdGAP is required for ErbB2-induced tumor growth and metastasis to the lungs

To further examine the role of CdGAP *in vivo*, pooled ErbB2-expressing control and CdGAP-depleted tumor cells were injected into the mammary fat pads of athymic mice. 6 weeks post-injection, the loss of CdGAP led to a 75% reduction in primary tumor volume compared to tumors formed in mice injected with control cells (Figure 6a). Histological analysis of the primary tumors confirmed reduced CdGAP expression and increased E-cadherin expression in CdGAP-depleted compared to control tumors (Figures 6b and c). To better understand how CdGAP influenced tumor growth, we performed TUNEL staining and examined the cell proliferation Ki-67 marker by immunohistochemistry (IHC). Cell proliferation was reduced by 2.5-fold in CdGAP-depleted tumors (Figure 6d). In contrast, CdGAP depletion did not induce cell apoptosis in these tumors (Supplementary Figure 4a). Next, we determined whether CdGAP is required to promote lung metastasis of ErbB2-expressing breast cancer cells. An equal number of ErbB2-expressing control or CdGAP-depleted cells was injected directly into the mouse lateral tail vein. After 4 weeks, 90% of the mice injected with control cells developed multiple lung metastases, while only 10% of the mice injected with CdGAP-depleted cells formed small metastatic lesions (Figure 6e). Altogether, these results indicate that CdGAP is necessary for the growth and metastasis of ErbB2-expressing breast cancer cells.

High expression level of CdGAP is associated with poor clinical outcome for breast cancer patients

To assess the clinical relevance of CdGAP in human breast cancer, we examined CdGAP protein expression in a panel of human breast cancer epithelial cell lines. As shown in Figure 7a, CdGAP expression was detected in normal mammary epithelial MCF10A cells but was absent or barely detectable in MCF7 or T47D cells derived from luminal A breast cancers. Except for the basal A BT20 cell line, CdGAP was highly expressed in cells derived from basal-like tumors, in particular within claudin-low basal-like cancer cells.[18, 19] We also found an inverse correlation between CdGAP and E-cadherin expression levels in luminal A and basal B breast cancer cell lines, consistent with our results obtained in ErbB2-expressing breast cancer cells (Figure 2b). Furthermore, CdGAP overexpression in MCF7 cells, which have low CdGAP expression, significantly increased SNAIL1 and ZEB2 mRNA expression levels while decreasing E-Cadherin mRNA and protein levels (Figures 7b and c). Conversely, siRNA-mediated CdGAP knockdown in MDA-MB-231 cells, which express high levels of CdGAP, decreased SNAIL1 and ZEB2 mRNA levels while increasing E-Cadherin mRNA and protein levels (Figures 7d and e). In addition, the loss of CdGAP in MDA-MB-231 cells induced a morphological change from a spindle-shaped, mesenchymal form in control cells to a rounded, epithelial-like morphology in CdGAP-depleted MDA-

MB-231 cells (Figure 7f). A cortical pattern of F-actin staining, which is a hallmark of the epithelial morphology,[1] was also observed in CdGAP-depleted MDA-MB-231 cells (Figure 7f). Consistent with the increased E-Cadherin levels detected in CdGAP-depleted MDA-MB-231 cells (Figures 7d and e), we observed an increased E-Cadherin fluorescence intensity in CdGAP-depleted MDA-MB-231 cells compared to control cells (Figure 7g). We also examined the effect of Myc-Zeb2 overexpression on endogenous CdGAP localization in MDA-MB-231 cells by subcellular fractionation experiments (Figure 7h). Endogenous CdGAP was predominantly found in the cytoplasmic fraction of MDA-MB-231 cell lysates and overexpression of Myc-Zeb2 in MDA-MB-231 cells led to an increase of endogenous CdGAP in the nuclear fraction (Figure 7h). Next, we performed a molecular subtype analysis of transcriptional profiles using a large compendium of breast cancer datasets with tumor gene expression profiles and clinical annotations (n=3666).[20] A significantly higher level of CdGAP mRNA expression was found in $ER^{-}/HER2^{-}$ basal-like breast cancer patients compared to $Her2^{+}$, Luminal A or Luminal B subtypes (Kruskal-Wallis test p=1.1E-35) (Figure 8a). Moreover, using the cBioPortal for Cancer Genomics (www.cbioportal.org),[21] we found a strong correlation (Pearson=0.74, Spearman=0.81) between CdGAP and Zeb2 mRNA expression, based on the RSEM (RNA-Seq by Expectation-Maximization) values of the transcripts in 963 samples from a breast invasive carcinoma study (TCGA, provisional) (Figure 8b), further supporting the molecular interaction between CdGAP and Zeb2 observed in breast cancer cells (Figure 5a).

We next performed IHC on a panel of breast tumors using tissue microarray analysis (TMA) to assess CdGAP expression in breast cancer patients (n=353). 74% of the breast tumor specimens showed a moderate to strong cytoplasmic CdGAP expression, while 26% displayed weak to no CdGAP expression (Figure 8c). Consistent with CdGAP localization in human breast cancer cells, CdGAP was detected both in the cytoplasm and nucleus of breast tumors (Figure 8d). 51.3% of breast tumors showed predominantly a cytoplasmic localization and 43.9% of the tumors presented a mixture of cytoplasmic and nuclear localization, while only 4.8% of the tumors exhibited an exclusive nuclear localization of CdGAP (Figures 8d and e). We then examined the relationship between CdGAP mRNA expression and the survival probability in breast cancer patients according to clinical subtypes using a large dataset of microarray analysis linked to clinical outcome (n=1190).[20] As shown in Kaplan-Meier survival analysis of patients classified based on CdGAP expression, high CdGAP expression was associated with poor disease-free survival in all subtypes of breast cancer patients (n=1190, logrank P=4.1E-4), and specifically in $ER^{+}/HER2^{-}$ low proliferation (Luminal A) subtype (n=333, logrank P=7.1E-4) (Figure 8f and Supplementary Figure 4b). Overall, these results indicate that CdGAP is frequently overexpressed in breast cancer tumors and high CdGAP expression is associated with a poor prognosis for the patients.

Discussion

During the process of epithelial-to-mesenchymal transition in cancer cells, the loss of E-cadherin expression, which results in the disruption of adherens junctions, correlates with increased cell migration, invasion, cancer metastasis and a poor clinical outcome in patients with cancer.[15, 22] The study presented here identifies a gene signature composed of the

TGFβ signaling pathway and EMT genes under the control of CdGAP in ErbB2-expressing breast cancer cells. We demonstrate that CdGAP acts as a novel co-transcriptional repressor with Zeb2 to suppress E-cadherin expression in breast cancer cells. Consequently, this work reveals an unexpected GAP-independent role for CdGAP as a positive modulator of breast tumorigenesis and metastasis (Figure 9). This is in marked contrast with the paradigm that GAPs are often seen as tumor suppressor proteins through their negative regulation of Rho GTPases.[23] Notably, the RhoGAP protein DLC1 (deleted in liver cancer 1) serves as a tumor suppressor protein by inactivating RhoA in various cancer types.[24, 25] On the other hand, overexpression of DLC1 prevents cell migration, invasion and metastatic cancer progression.[24] However, recent studies have also suggested that DLC1 can function as a scaffolding protein to affect cell survival and cell motility in a GAP-independent manner.[26] ARHGAP30, a GAP for RhoA and Rac1, [27] has been found to suppress cell migration and invasion in colorectal cancer through the regulation of p53 acetylation independently of its GAP activity.[28] Here, we demonstrate that the mechanism by which CdGAP binds with Zeb2 and represses E-cadherin promoter activity does not involve its GAP domain but the proline-rich domain (PRD). Therefore, this is suggesting that the modulation of Rac1/Cdc42 GTPase activities is not required for CdGAP transcriptional function and its ability to induce breast tumorigenesis and metastasis to the lungs. This work also supports our previous findings that the CdGAP-PRD is sufficient to mediate TGFβ-induced cell migration and invasion in breast cancer cells. [12]

The truncated CdGAP (1-683) mutant protein, lacking the C-terminal tail and part of the PRD, was first identified in autosomal-dominant AOS patients and was shown to display increased GAP activity *in vitro* towards Cdc42, suggesting that the disease mutations in the *CdGAP* gene behave as dominant gain-of-function alleles.[10] In the present study, we demonstrate that this AOS-related CdGAP mutant is not able to repress E-cadherin transcription, raising the interesting possibility that additional regulatory mechanisms independent of Rac1/Cdc42 regulation may also explain the rare developmental abnormalities of AOS patients. Furthermore, mutations in several genes of the Notch signaling pathway, including EOGT, RBPJ, Notch1 and the Notch ligand Dll4, have recently been identified in AOS patients.[29–32] In addition, a recent study revealed a VEGF-regulated Snail1-Dll4/Notch axis to control vascular development.[33] VEGF modulates Dll4 expression by co-inducing the transcriptional repressor Snail1 via the ERK/AKT-dependent inactivation of GSK3β.[33] Of interest, we show here that CdGAP regulates the levels of Snail1 transcripts in breast cancer cells. Genes of the Notch signaling pathway, including Notch4, Dll4 and MAML2, were also identified under the control of CdGAP (Supplementary Table 1). It is noteworthy that we have recently found CdGAP as a critical regulator of vascular development and VEGF-mediated angiogenesis.[34] In this way these findings support an emerging link between the novel transcriptional activity of CdGAP and the potential regulation of the Notch signaling pathway in the control of vascular development, angiogenesis and cancer metastasis.

In patients with breast cancer, we found a significant high expression level of CdGAP in ER^{-}/HER2^{-} basal-like breast cancer patients. Furthermore, CdGAP expression strongly correlated with Zeb2 expression in patient samples from a breast invasive carcinoma study. Thus, these data strongly support the novel nuclear function of CdGAP in breast cancer

cells. Furthermore, CdGAP expression is inversely correlated with poor prognosis, in particular in the luminal A subtype, which tends to be the subtype with fairly high survival rates and low recurrence rates. [35,36] Therefore, we propose that CdGAP could serve as a novel biomarker to identify luminal A subtype patients at higher risk and targeting CdGAP-Zeb2 interactions may help to define novel therapeutic strategies to treat breast cancer patients.

Materials and Methods

Cell culture, DNA constructs, and transfections

NMuMG-derived ErbB2-expressing mammary tumor explant cells were cultured as previously described. [5, 12] Briefly, activated forms of the ErbB-2 receptor were individually transfected into an immortalized mammary epithelial cell line (NMuMG), along with an empty vector control. Pooled stable transfected cells were injected into the mammary fat pads of athymic mice. Activated ErbB2-expressing tumor cells were explanted back into culture from ErbB2-expressing mammary tumors. ErbB2-expressing tumor explants cells with stable expression of short hairpin RNAs targeting CdGAP were generated using the pSuper retro GFP/neo retroviral shRNA expression vector (Oligoengine, Seattle, WA, USA). Hairpins were derived from the template sequence: 5′-GGGACCAUCUGGUAUACAAtt-3′ (ref.12). HEK293 cells, human breast cancer cell lines were routinely culture as previously described. [12,18] CdGAP and Zeb2 constructs were as described previously. [12, 16, 37] pGL2Basic-EcadK1/EpaIMUT/EboxMUT/Ebox2MUT in the luciferase assay was a gift from Eric Fearon (Addgene plasmid # 19291). pGL2Basic-pEcad (−1008/+49) was a gift from Morag Park. HEK293 cells were transfected by the Polyethylenimine method, whereas Lipofectamine 2000 was used for transfection of ErbB2-expressing explants tumor cells, MCF7 and MDA-MB-231 cells. siRNAs for CdGAP knockdown in human breast cancer cell lines were described previously. [34]

Western blotting and Immunofluorescence

Cells were lysed in RIPA buffer (50mM Hepes pH 7.4, 0.1% SDS, 0.5% Deoxycholic acid, 1% Triton X-100,150mM NaCl, 10mM EDTA). Protein lysates were resolved by SDS/PAGE and transferred to nitrocellulose membranes. The membranes were visualized by ECL. Antibodies used: CdGAP, HPA036380, Sigma, Oakville, ON, Canada; Neu, sc-284-G, Santa Cruz Biotechnology, Dallas, TX, USA; α-Tubulin, T5168, Sigma; Lamin B1, ab16048, Abcam Toronto, ON, Canada; E-cadherin, 3195, Cell Signaling, Danvers, MA, UAS; Snail, 3895, Cell Signaling; Zeb2, sc-271984, Santa Cruz Biotechnology; Myc, 05-419, Millipore, Billerica, MA, USA; Flag, F-7425, Sigma; Rac1, 05-389, Millipore. Immunofluorescence was performed as described previously.[12] For GFP-CdGAP localization experiments, the images were imaged on a Zeiss LSM780 laser-scanning confocal microscope with a PLAN-Apochromat 63×/1.4 oil objective lens and analyzed with Zen2009 software (Carl Zeiss Microscopy, Toronto, ON, Canada) and at least 100 cells which express both GFP-CdGAP and Myc-Zeb2 or empty vector per condition were analyzed.

Rac1 activation assay

The Rac1 activation assay was performed as previously described. [37]

mRNA Isolation, RNA-sequencing and analysis

mRNA Isolation, RNA-sequencing were performed as previously described.[6] Briefly, total RNA from 3 independent samples of ErbB2-expressing control (shCON) or CdGAP-depleted (shCdGAP) explants cells was extracted using TRIZOL reagent (Invitrogen, Burlington, ON, Canada) and cleaned up using a RNeasy column. 10 μg of total RNA for each sample was used to generate expression libraries, cBot cluster. Deep sequencing were performed using Illumina TruSeq RNA sample Preparation kit, TruSeq SR cluster kit v2 and Illumina TruSeq SBS kit V2 for 50 cycles, respectively. Illumina HiSEq2000 platform was used to perform sequencing at the McGill University and Genome Quebec Innovation Center. After the bases were called with the Illumina CASAVA 1.8 pipeline, poor quality bases and adapter sequences were trimmed with Trimmomatic (v. 0.30),[38] the reads were aligned to the GRCm38 reference genome with TopHat (v. 2.0.14).[39] Differential gene expression was then quantified with Cuffdiff (v 2.2.1).[40] Both the reference genome and the annotations were downloaded from Ensembl. For the analysis, genes were considered significantly differentially expressed when the adjusted *P*-val < 0.01 and the fold change was either < 0.625 or > 1.6 and the average count > 40. For the DAVID analysis, only the genes found to be upregulated in the CdGAP-depleted cells were submitted (fold change > 1.6) and the Functional Annotation Clustering tool was used. The map of the TGFβ pathway was generated in Ingenuity Pathway Analysis and was annotated by hand and genes were considered differentially expressed when the adjusted *P*-val < 0.01 and the average count > 40.

Quantitative Real-Time PCR

Total RNA were extracted using RNeasy mini kit. mRNA was reverse-transcribed using enzymes from Invitrogen. Quantitative Real-Time PCR (Q-PCR) was performed as described previously[12] using primers listed in Supplementary Table 2. All Q-PCR were performed using Power SYBR® Green PCR Master Mix (Invitrogen). Gene expression was normalized to 18S ribosomal subunit. Each sample was analyzed in triplicates.

Luciferase assays

ErbB2-expressing breast cancer cells (shCON or shCdGAP) or HEK293 cells were co-transfected with human wild type E-Cadherin-luc (−1008/+49) or E-box mutant E-Cadherin-luc plasmids together with the CdGAP and/or Zeb2 constructs. The luciferase assays were performed 20 hours post-transfection as described previously. [16]

Subcellular fractionation

Nuclear and cytoplasmic fractions were prepared using the NE-PER nuclear and cytoplasmic extraction reagents kit (Thermo Fisher Pierce) according to the manufacturer's protocol.

Chromatin Immunoprecipitation (ChIP)

ChIP assays were conducted as described previously.[41] Soluble chromatin was immunoprecipitated using GFP (Invitrogen) or Myc antibodies (Millipore). A 251bp fragment in the proximal E-Cadherin promoter or a 207bp fragment in the upstream region of the *CDH1* gene as a negative control were amplified with primers described in Supplementary Table 2. Q-PCR was carried out according to the following program: 95°C for 3 min, 40 cycles at 95°C for 20s', 60°C for 30s' and 72°C for 30′. Each sample was analyzed in triplicates.

Tumorigenesis and Metastasis assay

To assess primary tumor growth of the pooled ErbB2-expressing breast cancer cells, 1×10^6 cells were injected into the number 4 mammary fat pad of female athymic mice. Tumor volumes were calculated using the formula: $V = \pi$ (length × width^2)/6. For tumor metastasis assay, 2×10^5 cells were directly injected into the lateral tail veins of athymic mice. Four step sections (40 μm/step) were collected from each set of lungs (all lobes) and stained with hematoxylin and eosin (H&E) to visualize metastatic lesions. From the scanned sections, the number of metastases per lung section was counted across all step sections. All animal protocols were approved by McGill University Animal Use and Care committee, in accordance with guidelines established by the Canadian Council on Animal Care.

Patients and histological samples

Archived formalin-fixed paraffin-embedded tumour blocks were retrieved from all cases of primary breast cancer accessioned in the pathology department at UBC Hospital from 1997 to 2002 (patients whose cancer were originally diagnosed between January 1, 1989 and Dec 31, 2002, and who were referred to the BCCA for treatment). Patients with a previously diagnosed breast cancer, or synchronous breast cancer (additional diagnosis of breast cancer within 6 months of diagnosis of the primary) were excluded. Tissue microarrays were constructed from all cases with sufficient tumour within the original blocks (i.e. cases with microinvasive primary tumours were excluded). There was sufficient tissue present to perform immunohistochemical staining in 355 cases.

TMA (Tissue microarray analysis) Construction

Each tissue block had a corresponding H&E-stained slide. Representative areas of invasive carcinoma were identified and marked on the H&E slide and its corresponding tissue block. The TMAs were assembled with a tissue-arraying instrument (Beecher Instruments, Silver Springs, MD), as described elsewhere.[42, 43] Briefly, the area of interest in the donor block was cored twice with a 0.6-mm diameter array needle and transferred to a recipient paraffin block for construction of a 2-fold redundant tissue microarray block. Three TMA blocks were designed.

Immunohistochemistry (IHC)

For tumorigenesis assays, primary mammary tumors were fixed in 4% paraformaldehyde overnight and paraffin embedded. IHC was performed with E-cadherin antibodies, or Ki67 antibodies (ab15580, Abcam, Toronto, ON, Canada) or polyclonal anti-CdGAP antibodies

obtained by immunization of rabbits with CdGAP C-terminal amino acids [aa] (1253-1425) fused to GST. Anti-CdGAP antibodies were affinity-purified on a CH-Sepharose column bound to His-tagged CdGAP (aa1253-1425). For TUNEL staining, an ApopTag® Peroxidase *In Situ* Apoptosis Detection Kit (s7100, Millipore) was using to detect apoptotic cells. All slides were counterstained using H&E. Slides were scanned using a Scanscope XT digital slide scanner (Aperio,) and analyzed with Imagescope software (Aperio). For human TMA experiments, 4μm thick sections were cut from the array blocks and transferred to adhesive-coated glass slides. CdGAP staining (using anti-CdGAP antibodies, HPA036380 (Sigma)) was performed using a Ventana Discovery XT automated immunohistochemical stainer (dilution: 1:20). Staining for CdGAP was scored for both nuclear and cytoplasmic staining. Cytoplasmic staining was scored based on intensity, using a 4 point scale: 0 - negative; 1 - weak; 2 - moderate; 3 - strong. Nuclear staining was scored based on the percentage of tumour cells demonstrating nuclear staining (any intensity). Percent nuclear staining was scored as a continuous variable. All scoring was performed by ZK (breast pathologist), blinded to outcome. All scores were entered into a standardized electronic spreadsheets and then processed by using the software TMA-Deconvoluter 1.10 (ref. 44). The processed score data were then analyzed with the SPSS for Windows statistical software package.

Compendium of microarray breast cancer datasets

We retrieved all clinical and gene expression data of previously reported microarray datasets as published in Haibe-Kains *et al*.[20] To ensure comparability of expression values across multiple data sets, ESR1, ERBB2, and AURKA gene expression module scores were rescaled prior to applying the subtype classifier as in Azim HA Jr. *et al*[45]. (we used SCMOD1 classifier which is referred to as SCM in the present study). Our rescaling approach is implemented and fully documented in our R/Bioconductor package genefu version 1.5.2. Differences in expression of CdGAP according to subtype was examined using the Kruskal-Wallis test.

Survival analysis

All survival analyses have been performed with R version 2.15 using the survival package version 2.36–14. Distant metastasis-free survival (DMFS) was the primary survival end-point, which is defined as the time elapsing between breast cancer diagnosis and date of local or systemic relapse, or death. When DMFS data were not reported, relapse-free survival (RFS) information was used if available. All survival data were censored at 10 years to ensure similar follow-up across the different datasets that are part of our compendium.[46] Survival plots according to the CdGAP tertiles were drawn using the Kaplan-Meier method, and the significance of the survival differences were evaluated using the log-rank *p*-test.

Statistical analysis

Statistical analysis was performed using a two-sample unequal-variance Student’s t test. Data are presented as the mean +/− SEM.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support: Canadian Institute for Health Research (CIHR) MOP-119544 to N.L.-V. and MOP- 144425 to J.F.C. J.F.C. is a recipient of a FRQS senior scholarship.

We thank Dr Patricia Tonin for helpful discussions and Dr John J. Bergeron for critically reading this manuscript. We are grateful to Dr Morag Park (McGill University) for providing the Zeb2 and E-Cadherin-luc constructs. We thank Dr. Torsten Nielsen for providing the human breast cancer TMA and the CdGAP IHC staining through the Genetic Pathology Evaluation Center (GPEC), ZF Dong for IHC staining of primary tumors, Dr Min Fu for assistance with confocal microscopy at the imaging platform of the RI-MUHC, and Vilayphone Luangrath for affinity purification of the CdGAP antibodies. N. L-V was a recipient of a FRSQ chercheur-national and a William Dawson scholar.

References

1. Gonzalez DM, Medici D. Signaling mechanisms of the epithelial-mesenchymal transition. Sci Signal. 2014; 7:re8. [PubMed: 25249658]
2. Wong TS, Gao W, Chan JY. Transcription regulation of E-cadherin by zinc finger E-box binding homeobox proteins in solid tumors. Biomed Res Int. 2014; 2014:921564. [PubMed: 25197668]
3. King CR, Kraus MH, Aaronson SA. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. Science. 1985; 229:974–976. [PubMed: 2992089]
4. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987; 235:177–182. [PubMed: 3798106]
5. Northey JJ, Chmielecki J, Ngan E, Russo C, Annis MG, Muller WJ, et al. Signaling through ShcA is required for TGFβ and Neu/ErbB-2 induced breast cancer cell motility and invasion. Mol Cell Biol. 2008; 28:3162–3176. [PubMed: 18332126]
6. Laurin M, Huber J, Pelletier A, Houalla T, Park M, Fukui Y, et al. Rac-specific guanine nucleotide exchange factor DOCK1 is a critical regulator of HER2-mediated breast cancer metastasis. Proc Natl Acad Sci U S A. 2013; 110:7434–7439. [PubMed: 23592719]
7. Lamarche-Vane N, Hall A. CdGAP, a novel proline-rich GTPase-activating protein for Cdc42 and Rac. J Biol Chem. 1998; 273:29172–29177. [PubMed: 9786927]
8. Tcherkezian J, Danek EI, Jenna S, Triki I, Lamarche-Vane N. Extracellular signal-regulated kinase 1 interacts with and phosphorylates CdGAP at an important regulatory site. Mol Cell Biol. 2005; 25:6314–6329. [PubMed: 16024771]
9. Danek EI, Tcherkezian J, Meriane M, Triki I, Lamarche-Vane N. Glycogen synthase kinase-3 phosphorylates CdGAP at a consensus ERK1 regulatory site. J Biol Chem. 2007; 282:3624–3631. [PubMed: 17158447]
10. Southgate L, Machado RD, Snape KM, Primeau M, Dafou D, Ruddy DM, et al. Gain-of-function mutations of ARHGAP31, a Cdc42/Rac1 GTPase regulator, cause syndromic cutis aplasia and limb anomalies. Am J Hum Genet. 2011; 88:574–585. [PubMed: 21565291]
11. Isrie M, Wuyts W, Van Esch H, Devriendt K. Isolated terminal limb reduction defects: extending the clinical spectrum of Adams-Oliver syndrome and ARHGAP31 mutations. Am J Med Genet A. 2014; 164A:1576–1579. [PubMed: 24668619]
12. He Y, Northey JJ, Primeau M, Machado RD, Trembath R, Siegel PM, et al. CdGAP is required for transforming growth factor β- and Neu/ErbB-2-induced breast cancer cell motility and invasion. Oncogene. 2011; 30:1032–1045. [PubMed: 21042277]
13. Hajra KM, Chen DY, Fearon ER. The SLUG zinc-finger protein represses E-cadherin in breast cancer. Cancer Res. 2002; 62:1613–1618. [PubMed: 11912130]

14. Elloul S, Elstrand MB, Nesland JM, Tropé CG, Kvalheim G, Goldberg I, et al. Snail, Slug, and Smad-interacting protein 1 as novel parameters of disease aggressiveness in metastatic ovarian and breast carcinoma. Cancer. 2005; 103:1631–1643. [PubMed: 15742334]
15. Wang CA, Drasin D, Pham C, Jedlicka P, Zaberezhnyy V, Guney M, et al. Homeoprotein Six2 promotes breast cancer metastasis via transcriptional and epigenetic control of E-cadherin expression. Cancer Res. 2014; 74:7357–7370. [PubMed: 25348955]
16. Long J, Zuo D, Park M. Pc2-mediated sumoylation of Smad-interacting protein 1 attenuates transcriptional repression of E-cadherin. J Biol Chem. 2005; 280:35477–35489. [PubMed: 16061479]
17. Sánchez-Tilló E, Siles L, de Barrios O, Cuatrecasas M, Vaquero EC, Castells A, et al. Expanding roles of ZEB factors in tumorigenesis and tumor progression. Am J Cancer Res. 2011; 1:897–912. [PubMed: 22016835]
18. Holliday DL, Speirs V. Choosing the right cell line for breast cancer research. Breast Cancer Res. 2011; 13:215. [PubMed: 21884641]
19. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002; 347:1999–2009. [PubMed: 12490681]
20. Haibe-Kains B, Desmedt C, Loi S, Culhane AC, Bontempi G, Quackenbush J, et al. A three-gene model to robustly identify breast cancer molecular subtypes. J Natl Cancer Inst. 2012; 104:311–325. [PubMed: 22262870]
21. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013; 6:pl1. [PubMed: 23550210]
22. Hu QP, Kuang JY, Yang QK, Bian XW, Yu SC. Beyond a tumor suppressor: soluble E-cadherin promotes the progression of cancer. Int J Cancer. 2015; doi: 10.1002/ijc.29982
23. Tcherkezian J, Lamarche-Vane N. Current knowledge of the large RhoGAP family of proteins. Biol Cell. 2007; 99:67–86. [PubMed: 17222083]
24. Kim TY, Vigil D, Der CJ, Juliano RL. Role of DLC-1, a tumor suppressor protein with RhoGAP activity, in regulation of the cytoskeleton and cell motility. Cancer and Metastasis Reviews. 2009; 28:77–83. [PubMed: 19221866]
25. Lahoz A, Hall A. DLC1: a significant GAP in the cancer genome. Genes Dev. 2008; 22:1724–1730. [PubMed: 18593873]
26. Barras D, Widmann C. GAP-independent functions of DLC1 in metastasis. Cancer Metastasis Rev. 2014; 33:87–100. [PubMed: 24338004]
27. Naji L, Pacholsky D, Aspenström P. ARHGAP30 is a Wrch-1-interacting protein involved in actin dynamics and cell adhesion. Biochem Biophys Res Commun. 2011; 409:96–102. [PubMed: 21565175]
28. Wang J, Qian J, Hu Y, Kong X, Chen H, Shi Q, et al. ArhGAP30 promotes p53 acetylation and function in colorectal cancer. Nat Commun. 2014; 5:4735.doi: 10.1038/ncomms5735 [PubMed: 25156493]
29. Meester JA, Southgate L, Stittrich AB, Venselaar H, Beekmans SJ, den Hollander N, et al. Heterozygous loss-of-function mutations in DLL4 causeAdams-Oliver syndrome. Am J Hum Genet. 2015; 97:475–482. [PubMed: 26299364]
30. Southgate L, Sukalo M, Karountzos AS, Taylor EJ, Collinson CS, Ruddy D, et al. Haploinsufficiency of the Notch1 receptor as a cause of Adams-Oliver syndrome with variable cardiac anomalies. Circ Cardiovasc Genet. 2015; 8:572–581. [PubMed: 25963545]
31. Shaheen R, Aglan M, Keppler-Noreuil K, Faqeih E, Ansari S, Horton K, et al. Mutations in EOGT confirm the genetic heterogeneity of autosomalrecessive Adams-Oliver syndrome. Am J Hum Genet. 2013; 92:598–604. [PubMed: 23522784]
32. Hassed SJ, Wiley GB, Wang S, Lee JY, Li S, Xu W, et al. RBPJ mutations identified in two families affected by Adams-Oliver syndrome. Am J Hum Genet. 2012; 91:391–395. [PubMed: 22883147]
33. Wu ZQ, Rowe RG, Lim KC, Lin Y, Willis A, Tang Y, et al. Snail1/Notch1 signalling axis controls embryonic vascular development. Nature Comm. 2013; 5:3998.

34. Caron C, DeGeer J, Fournier P, Duquette PM, Luangrath V, Ishii H, et al. The Cdc42/Rac1 GTPase regulator CdGAP is critical for vascular development and VEGF-mediated angiogenesis. Sci Rep. 2016; 6:27485. [PubMed: 27270835]

35. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, et al. Comprehensive molecular portraits of human breast tumours. Nature. 2012; 490:61–70. [PubMed: 23000897]

36. Metzger-Filho O, Sun Z, Viale G, Price KN, Crivellari D, Snyder RD, et al. Patterns of recurrence and outcome according to breast cancer subtypes in lymph node-negative disease: results from international breast cancer study group trials VIII and IX. J Clin Oncol. 2013; 31:3083–3090. [PubMed: 23897954]

37. Karimzadeh F, Primeau M, Mountassif D, Rouiller I, Lamarche-Vane N. A stretch of polybasic residues mediates Cdc42 GTPase-activating protein (CdGAP) binding to phosphatidylinositol 3,4,5-trisphosphate and regulates its GAP activity. J Biol Chem. 2012; 287:19610–19621. [PubMed: 22518840]

38. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30:2114–2120. [PubMed: 24695404]

39. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14:R36. [PubMed: 23618408]

40. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012; 7:562–578. [PubMed: 22383036]

41. Tabariès S, Annis MG, Hsu BE, Tam CE, Savage P, Park M, et al. Lyn modulates Claudin-2 expression and is a therapeutic target for breast cancer liver metastasis. Oncotarget. 2015; 6:9476–9487. [PubMed: 25823815]

42. Parker RL, Huntsman DG, Lesack DW, Cupples JB, Grant DR, Akbari M, et al. Assessment of interlaboratory variation in the immunohistochemical determination of estrogen receptor status using a breast cancer tissue microarray. Am J Clin Pathol. 2002; 117:723–728. [PubMed: 12090420]

43. Makretsov N, Gilks CB, Coldman AJ, Hayes M, Huntsman D. Tissue microarray analysis of neuroendocrine differentiation and its prognostic significance in breast cancer. Hum Pathol. 2003; 34:1001–1008. [PubMed: 14608533]

44. Liu CL, Prapong W, Natkunam Y, Alizadeh A, Montgomery K, Gilks CB, et al. Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays. Am J Pathol. 2002; 161:1557–1565. [PubMed: 12414504]

45. Azim HA Jr, Michiels S, Bedard PL, Singhal SK, Criscitiello C, Ignatiadis M, et al. Elucidating prognosis and biology of breast cancer arising in young women using gene expression profiling. Clin Cancer Res. 2012; 18:1341–1351. [PubMed: 22261811]

46. Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res. 2007; 13:3207–3214. [PubMed: 17545524]

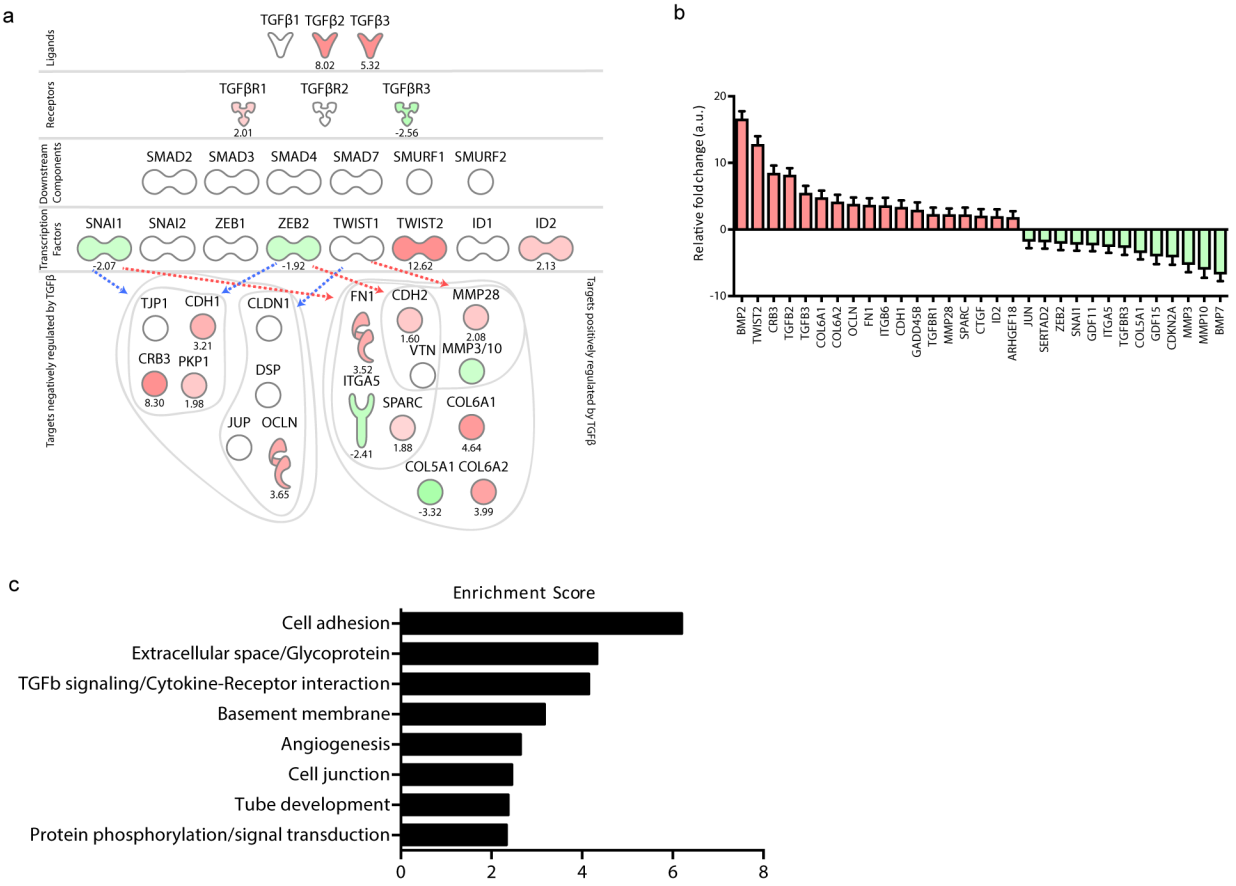

Figure 1.

CdGAP regulates the expression of genes involved in TGFβ signaling in breast cancer cells. (**a**) Map of the genes related to TGFβ signaling pathway differentially expressed between pooled ErbB2-expressing control (shCON) and CdGAP-depleted breast cancer cells (shCdGAP). Green: downregulated genes in shCdGAP, red: upregulated genes in shCdGAP, blue arrows: target genes downregulated, red arrows: target genes upregulated. The numbers shown represent the fold change shCdGAP/shCON (**b**) Expression level changes (shCdGAP/shCON) of epithelial-to-mesenchymal transition (EMT) related genes. $P < 0.01$. (**c**) Top 10 annotation clusters enriched in CdGAP-depleted cells. Annotation clusters enrichment was determined using DAVID and using genes upregulated in CdGAP-depleted cells.

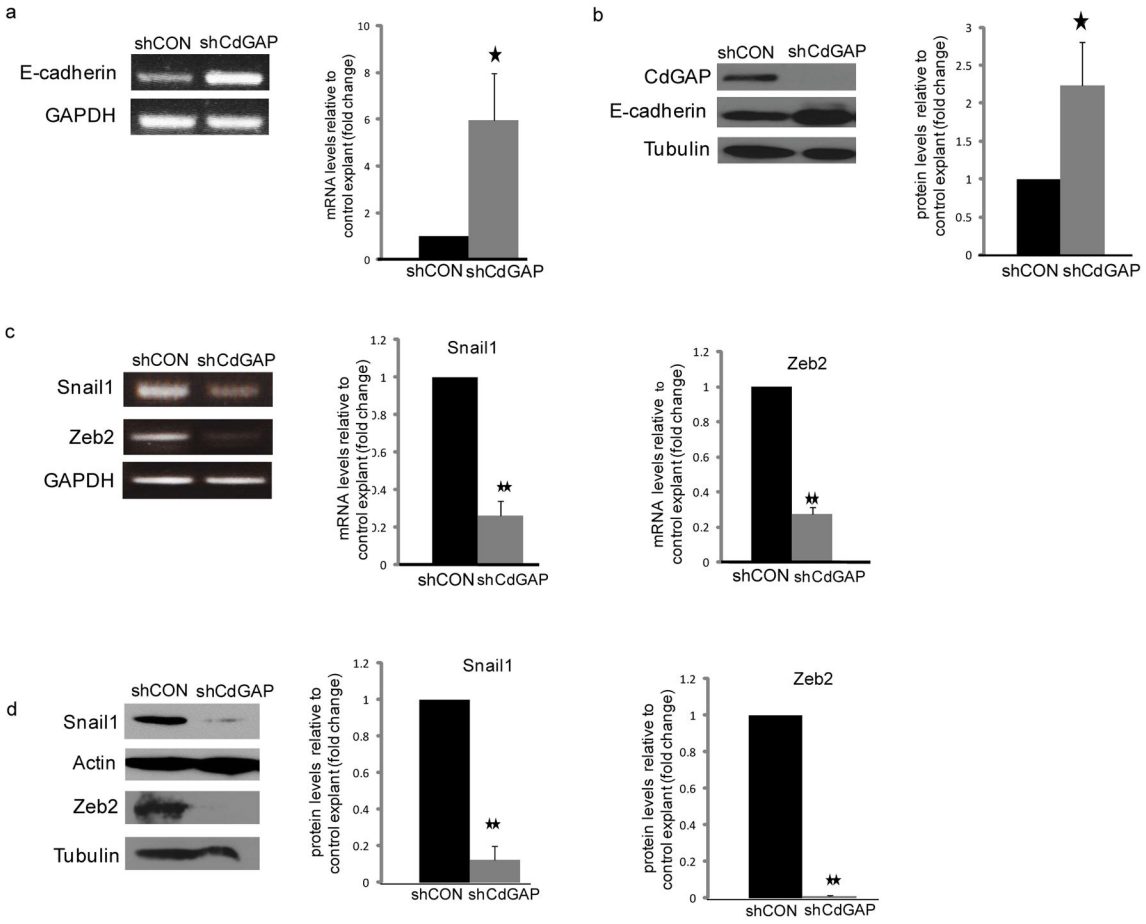

Figure 2.

The levels of E-cadherin, Snail1 and Zeb2 expression are altered in CdGAP-depleted ErbB2-expressing breast cancer cells. Q-PCR (**a** and **c**) of the indicated genes and immunoblot analysis (**b** and **d**) of the proteins from control (shCON) and CdGAP-deficient (shCdGAP) breast cancer cells. Error bars indicate SEM. n=3 **P*<0.05, ***P*<0.01.

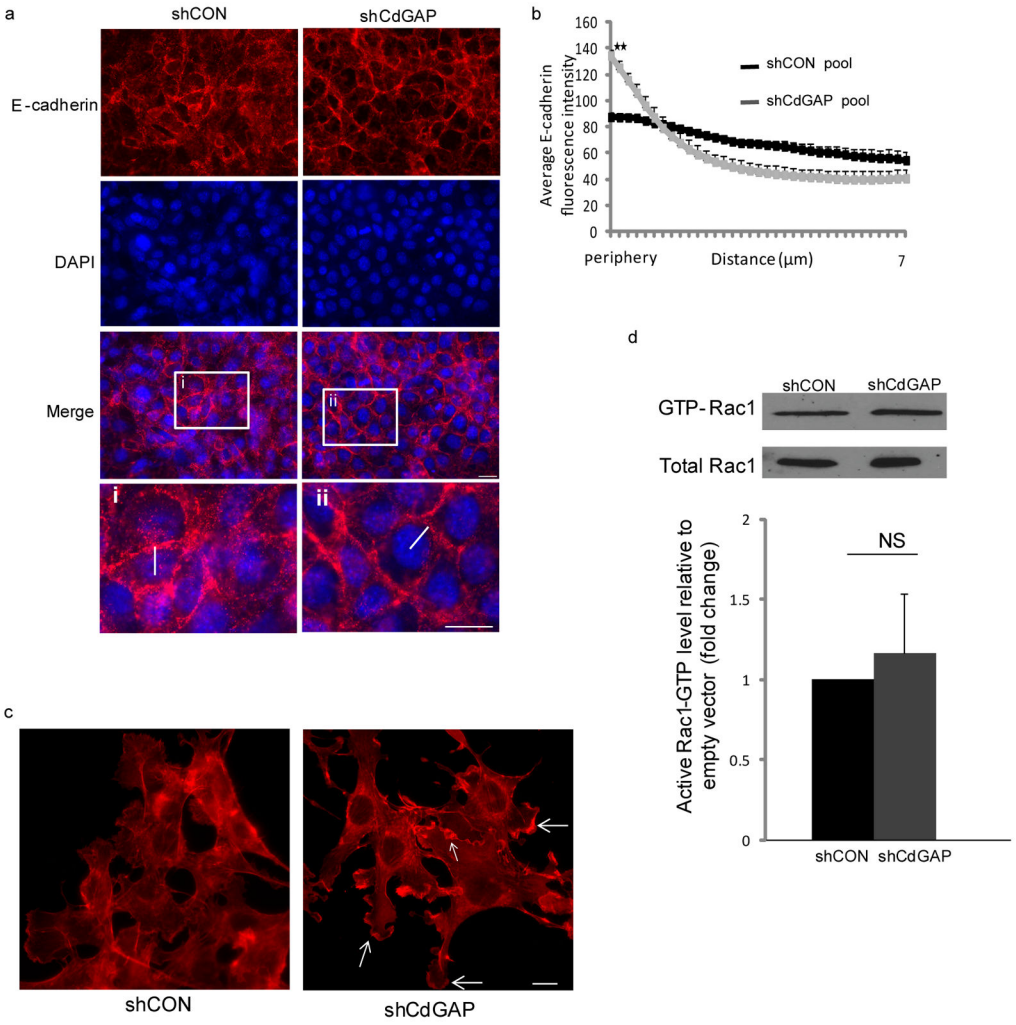

Figure 3.

Depletion of CdGAP restores E-cadherin at cell-cell junctions in ErbB2-expressing breast cancer cells. (**a** and **c**) ErbB2-expressing control (shCON) and CdGAP-depleted cells (shCdGAP) were fixed, stained for E-cadherin and with 4′, 6′-diamidino-2-phenylindole (DAPI) (**a**) or for F-actin with phalloidin (**c**). Arrows indicate the membrane ruffles. Scale bars, 10 μm. i and ii, insets (2.5 Fold). White lines represent the regions of interest used to measure E-cadherin fluorescence intensity in **b**. (**b**) The mean pixel intensity of E-cadherin fluorescence intensity along a 7-μm linescan from the cell periphery was calculated using Metamorph software. More than 75 cells were counted for each condition. n=3 $^{**}P<0.01$ versus shCON at cell periphery. (**d**) GTP-bound Rac1 was pulled down with GST-CRIB from control or CdGAP-depleted cell lysates. Densitometric ratio of GTP-bound Rac1/total Rac1 normalized to empty vector. n=7, NS, not significant. Error bars indicate SEM.

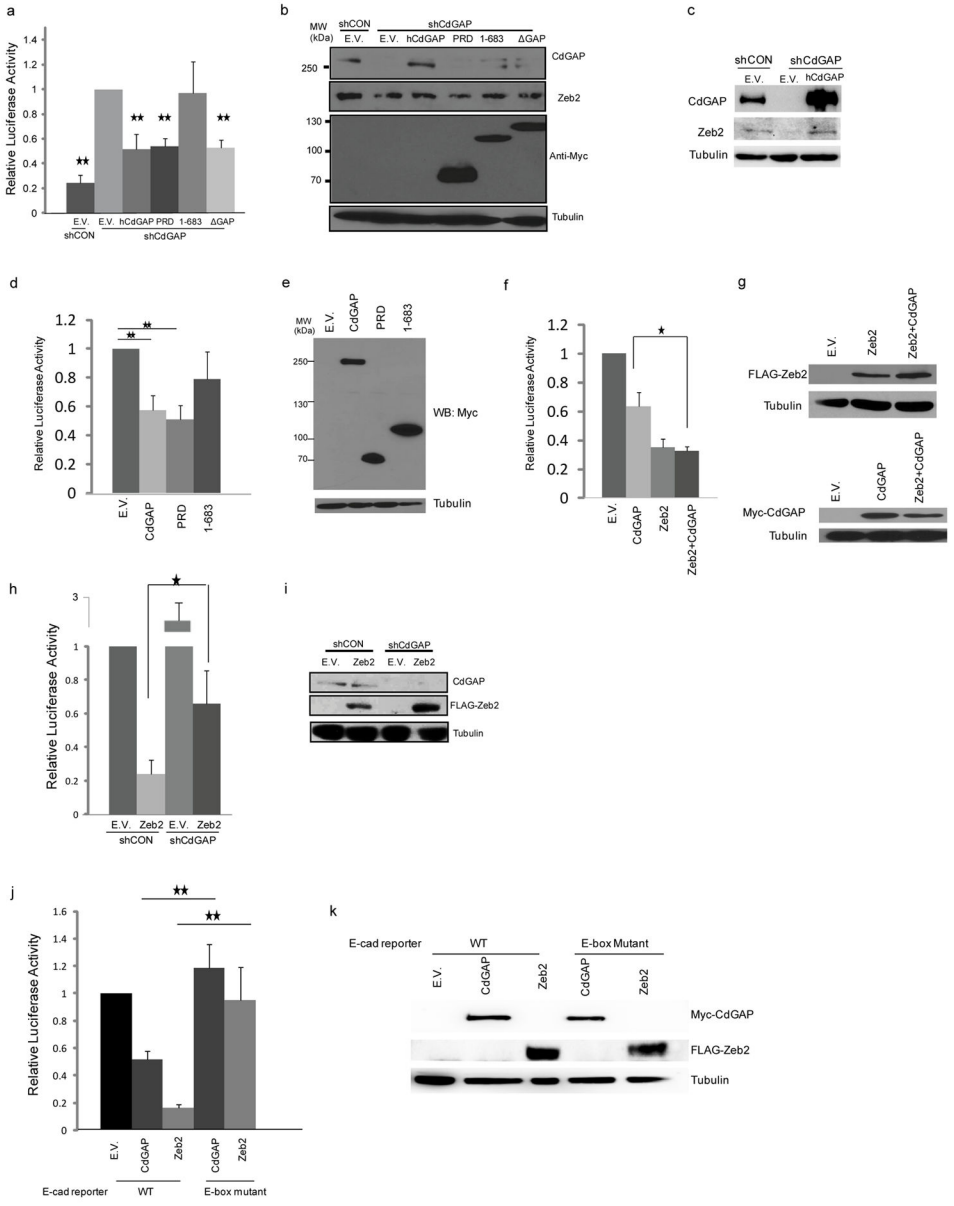

Figure 4.

CdGAP represses the E-cadherin promoter in a GAP-independent manner. (**a**) E-cadherin promoter luciferase assays were performed in ErbB2-expressing control (shCON) or CdGAP-depleted cells (shCdGAP) transfected with empty vector (E.V.), myc-tagged human CdGAP or the indicated CdGAP deletion mutants. Values are relative to that of CdGAP-depleted cells transfected with E.V. n=4. (**b**) Immunoblot analysis from (a) of the indicated proteins from control or CdGAP-depleted cells transfected with the indicated constructs. Protein cell lysates were collected 20 hours post-transfection. (**c**) Immunoblot analysis of the indicated proteins from control or CdGAP-depleted cells transfected with empty vector (E.V) or myc-tagged human CdGAP. Protein cell lysates were collected 48 hours post-transfection. (**d** and **f**) E-cadherin promoter luciferase assays were performed in HEK293 cells transfected with empty vector (E.V.), myc-tagged CdGAP, Flag-Zeb2 or the indicated

CdGAP deletion mutants. Values are relative to that of HEK293 cells transfected with empty vector. n=6. (**e** and **g**) Immunoblot analysis of the indicated proteins from HEK293 cell lysates. (**h**) E-cadherin promoter luciferase assays were performed in ErbB2-expressing control or CdGAP-depleted cells transfected with empty vector (E.V.) or Flag-Zeb2. Values are relative to that of ErbB2-expressing control cells. n=6. (**i**) Immunoblot analysis of the indicated proteins from control (shCON) or CdGAP-depleted cells transfected with empty vector (E.V.) or Flag-Zeb2. (**j**) HEK293 cells were co-transfected with luciferase constructs of wild-type (WT) or E-box mutant E-Cadherin promoter with CdGAP or Zeb2 constructs and the luciferase activity was measured 20 hours post-transfection. Values are relative to that of HEK293 cells transfected with empty vector and wild type E-cadherin promoter. n=3. (**k**) Immunoblot analysis of the indicated proteins from HEK293 lysates. Renilla luciferase control vector was co-transfected as the transfection control. Error bars indicate SEM. *$P<0.05$, **$P<0.01$.

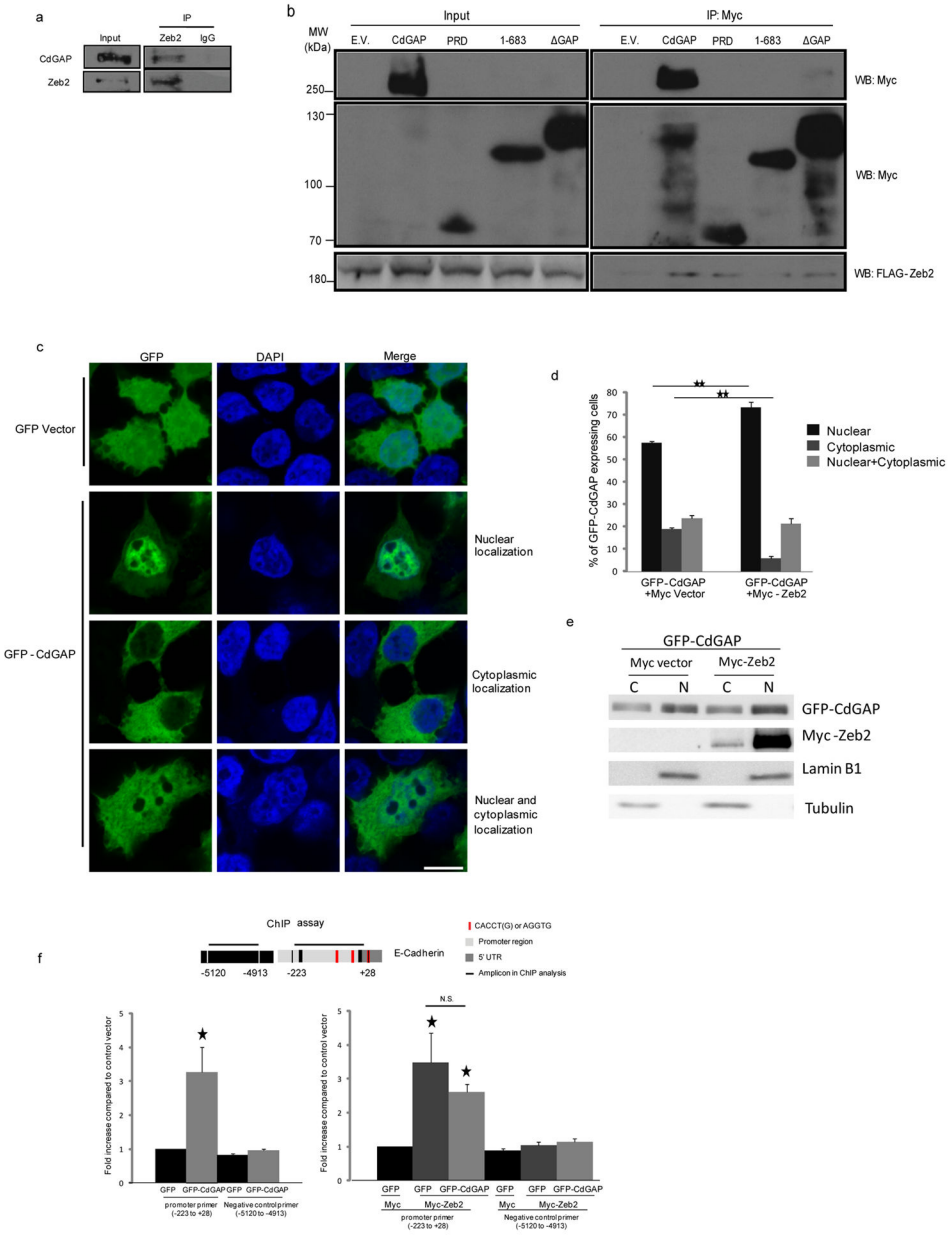

Figure 5.

CdGAP localizes to the nucleus with Zeb2 and interacts with the E-cadherin promoter. (**a**) Zeb2 was immunoprecipitated (IP) from lysates of ErbB2-expressing breast cancer cells with anti-Zeb2 antibodies or rabbit IgG as a control. IP proteins and total cell lysates (input) were immunoblotted with the indicated antibodies. (**b**) HEK293 cells were transfected with E.V., Flag-Zeb2 or myc-tagged CdGAP constructs followed by myc IP and immublotting with the indicated antibodies. Total cell lysates, input. (**c**) HEK293 cells were co-transfected with empty Myc vector and empty GFP vector or GFP-CdGAP. Fixed cells were stained with DAPI and GFP-CdGAP localization was assessed by confocal microscopy. Scale bar, 10 μm. (**d**) HEK293 cells were co-transfected with GFP-CdGAP and empty Myc vector or Myc-Zeb2. The percentage of GFP-CdGAP-expressing cells localizing to the nucleus, the

cytoplasm or both was calculated. More than 100 cells co-expressing GFP-CdGAP with Myc vector or Myc-Zeb2 were counted per condition. n=3. (**e**) Nuclear (N) and cytoplasmic (C) fractions were isolated from HEK293 cells co-transfected with GFP-CdGAP and empty Myc vector or Myc-Zeb2. Each fraction was immunoblotted with the indicated antibodies. Tubulin and Lamin B1 were used as specific markers of the cytoplasmic (C) and nuclear fractions (N), respectively. (**f**) Chromatin IP (ChIP) assay showing the ability of CdGAP and Zeb2 to bind the E-Cadherin promoter in HEK293 cells. n=3. Error bars indicate SEM.* $P<0.05$, ** $P<0.01$. N.S, not significant.

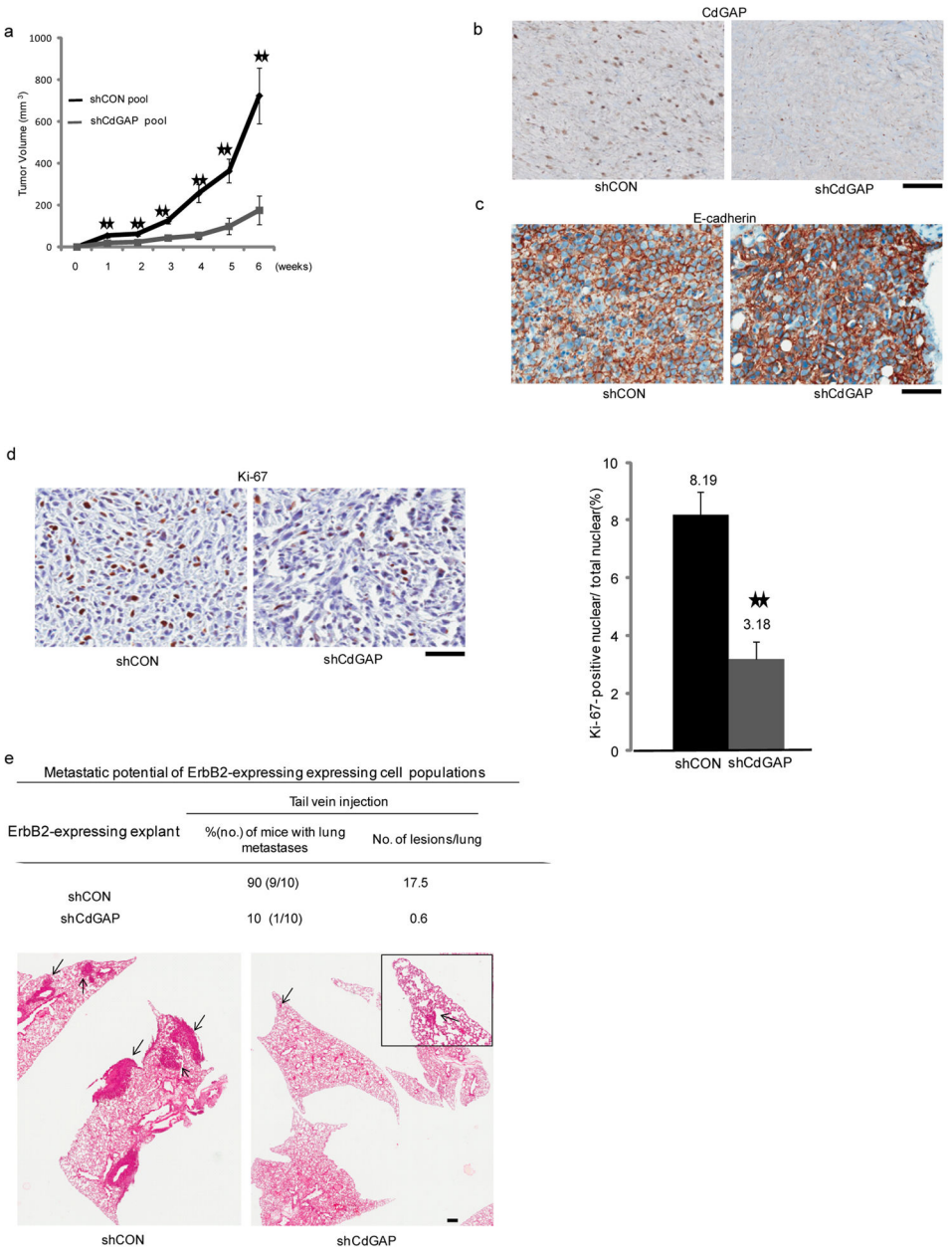

Figure 6.

Depletion of CdGAP inhibits primary tumor growth and lung metastasis *in vivo*. (**a**) ErbB2-expressing control vector (shCON) or CdGAP-depleted cells (shCdGAP) were injected into the mammary fat pads of athymic mice. Tumor growth was assessed by weekly caliper measurements and the average tumor volumes from independent control vector (n=9 mice) and shCdGAP tumors (n=8 mice) are plotted. Primary tumors were harvested and subjected to immunohistochemistry (IHC) staining for CdGAP expression (**b**), E-cadherin expression (**c**) and cell proliferation (**d**) by assessing the percentage of Ki67-positive nuclei. Error bars indicate SEM, $^{**}P<0.01$. Scale bars, 50 μm. (**e**) Control or CdGAP-depleted cells were injected into the lateral tail vein of athymic mice. The number of lung lesions was quantified

at necropsy (4 weeks) from 10 control or shCdGAP-injected mice. Representative images of H&E stained sections of the lungs collected 4 weeks after injection. Scale bar, 250 μm.

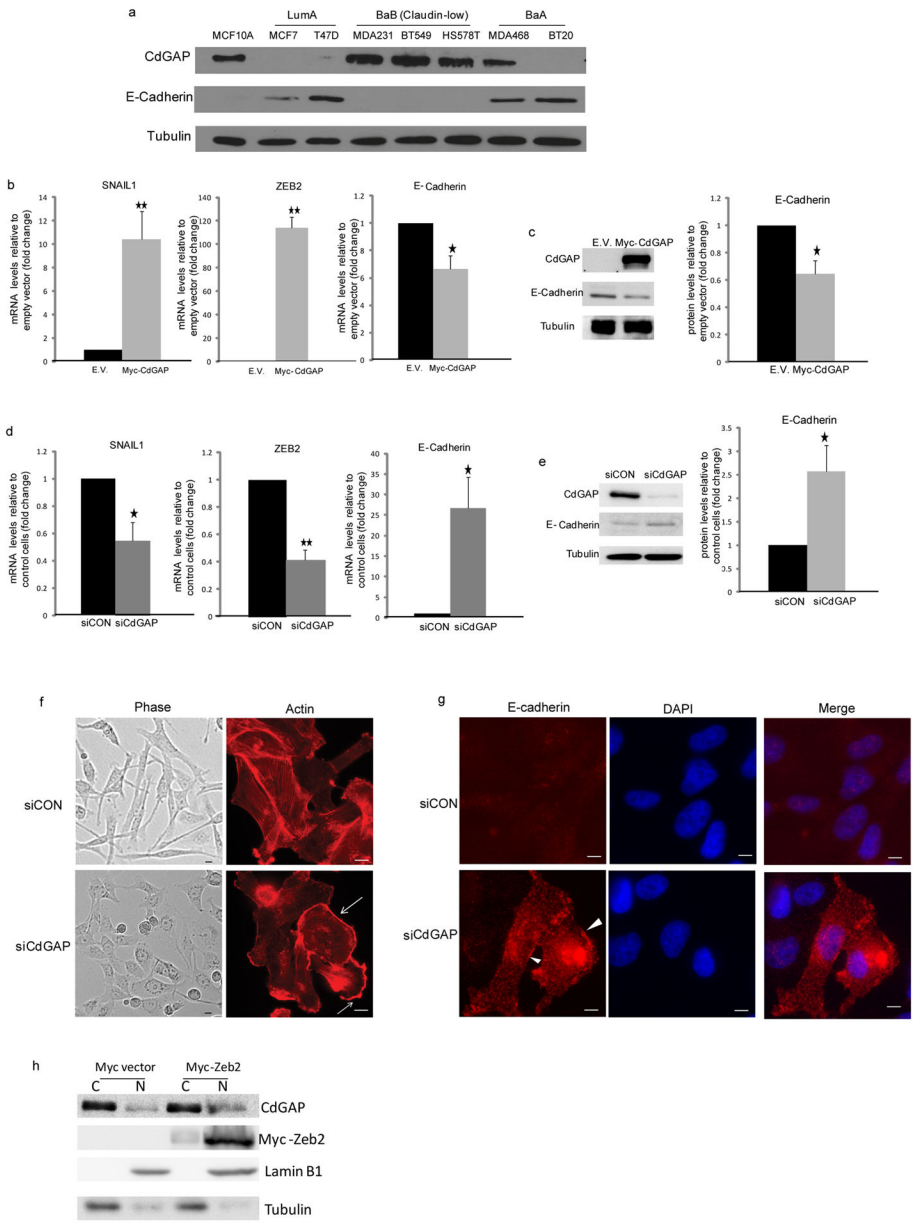

Figure 7.

CdGAP regulates the expression of SNAIL1, ZEB2, and E-Cadherin in human breast cancer cells and is highly expressed in basal subtype human breast cancer cell lines. (**a**) Immunoblot analysis of CdGAP and E-Cadherin levels in a panel of human normal mammary epithelial and breast cancer cell lines. Immortalized mammary epithelial cells (MCF10A), $ER^+/HER2^-$ low proliferation (luminal A) subtype cells (MCF7, T47D), $ER^-/HER2^-$ Claudin-low (Basal B) subtype cells (MDA-MB-231, BT549 and HS578T), $ER^-/HER2^-$ (Basal A) subtype cells (MDA-MB-468, BT20). Tubulin is used as a loading control. Q-PCR (**b**) of the indicated genes and immunoblot analysis (**c**) of the proteins from MCF7 cells transfected with empty vector (E.V.) or Myc-tagged human CdGAP. Values are relative to that of MCF7 cells transfected with empty vector. n=3. Q-PCR (**d**) of the indicated genes

and immunoblot analysis (**e**) of the proteins from MDA-MB-231 cells transfected with control siRNA (siCON) or siRNA targeting CdGAP (siCdGAP). Values are relative to that of MDA-MB-231 cells transfected with control siRNA. n=3. (**f** and **g**) MDA-MB-231 cells transfected with control siRNA (siCON) or siRNA targeting CdGAP (siCdGAP) were fixed, stained for F-actin with phalloidin (**f**) or for E-cadherin and nuclear staining with DAPI (**g**). Phase contrast images were captured to show cell morphological changes (**f**). Arrows indicate membrane ruffles. Arrowheads indicate cell membrane localization of E-cadherin (**g**). Scale bars, 10 μm. (**h**) Nuclear (N) and cytoplasmic (C) fractions were isolated from MDA-MB-231 cells transfected with empty Myc vector or Myc-Zeb2. Each fraction was immunoblotted with the indicated antibodies. Tubulin and Lamin B1 were used as specific markers of the cytoplasmic (C) and nuclear fractions (N), respectively. Error bars indicate SEM.* $P<0.05$, ** $P<0.01$

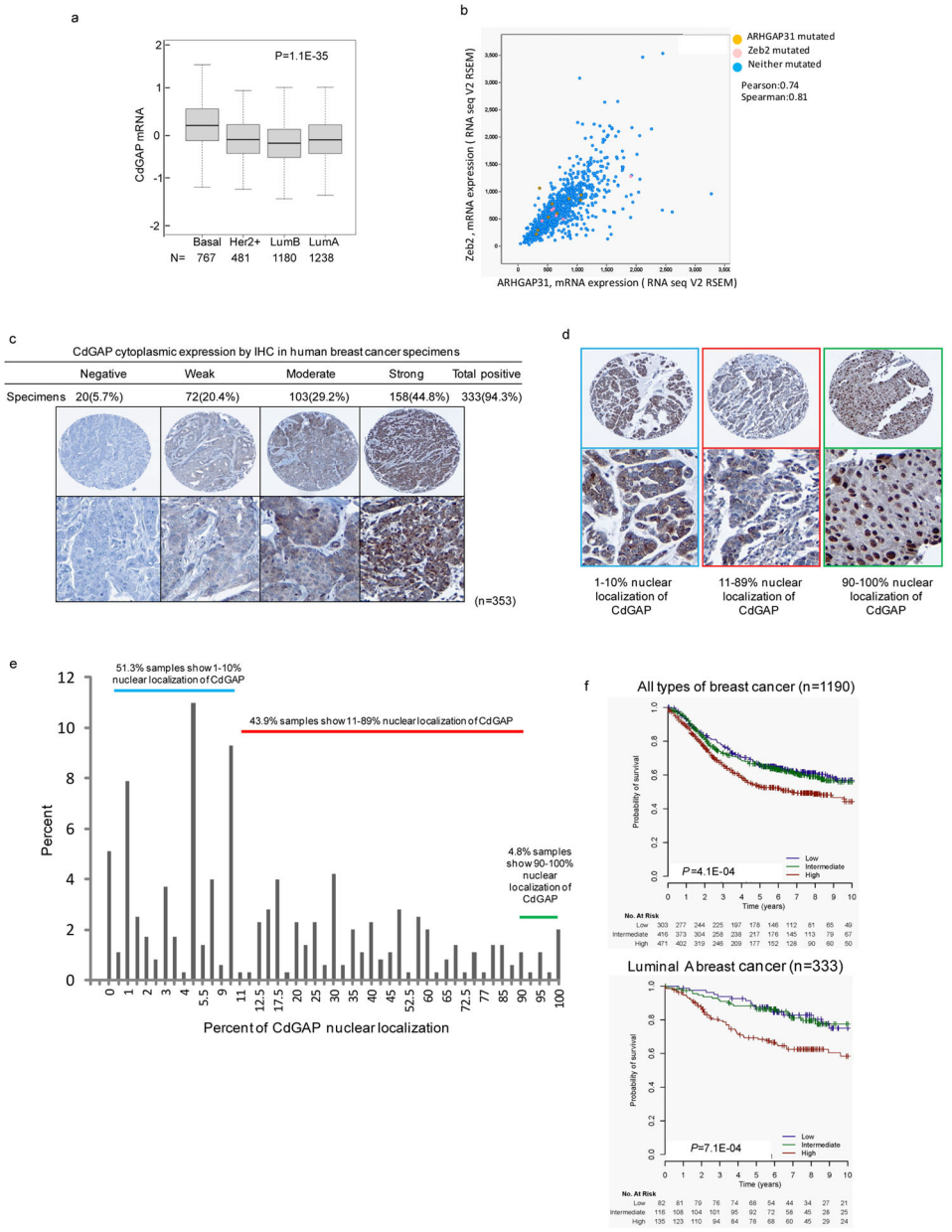

Figure 8.

CdGAP is a negative prognostic marker for human breast cancer survival. (**a**) Box-plot showing CdGAP transcript levels in human breast cancer subtypes using Kruskal-Wallis test. (**b**) Correlation analysis of CdGAP/ARHGAP31 and Zeb2 mRNA expression in a breast invasive carcinoma study (TCGA, provisional) from www.cbioportal.org. (**c**) IHC analysis of CdGAP cytoplasmic expression on a breast tumor microarray (TMA) containing 353 human breast cancer specimens. The percentage of specimens showing a negative, weak, moderate or strong CdGAP expression is represented in brackets. (**d** and **e**) The distribution of the specimens with CdGAP nuclear localization. (**f**) Kaplan-Meier curves representing the probability of survival of breast cancer patients (all types) or luminal A subtype based on relative levels of CdGAP expression.

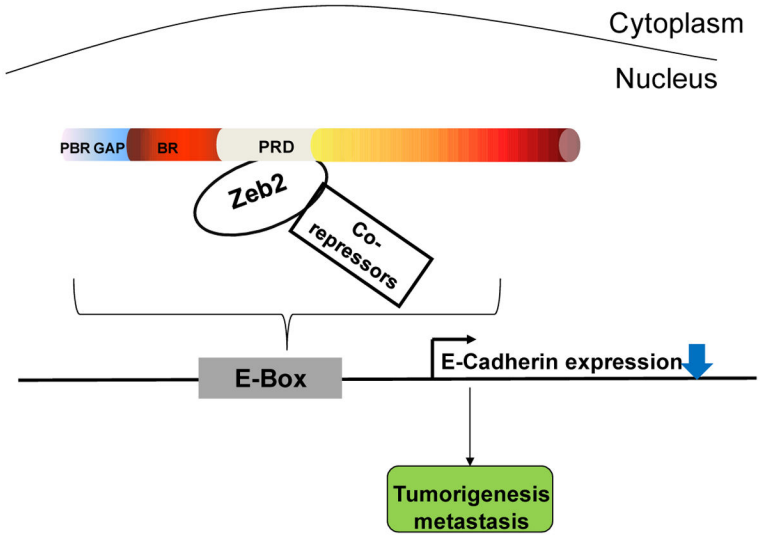

Figure 9.

Model of CdGAP as a novel transcriptional co-repressor of E-cadherin expression. CdGAP interacts with Zeb2 and acts as a negative transcriptional co-factor to repress E-cadherin expression, resulting in EMT, tumorigenesis and metastasis.