



## OPEN Tumor suppressor DEAR1 regulates mammary epithelial cell fate and predicts early onset and metastasis in triple negative breast cancer

Uyen Q. Le<sup>1</sup>, Nanyue Chen<sup>1</sup>, Seetharaman Balasenthil<sup>1</sup>, Eugene Lurie<sup>1</sup>, Fei Yang<sup>1</sup>, Suyu Liu<sup>2</sup>, Laura Rubin<sup>2</sup>, Luisa Maren Solis Soto<sup>1</sup>, Maria Gabriela Raso<sup>1</sup>, Harsh Batra<sup>1</sup>, Aysegul A. Sahin<sup>3</sup>, Ignacio I. Wistuba<sup>1</sup> & Ann McNeill Killary<sup>1</sup>✉

Triple negative breast cancer (TNBC) is a disease of poor prognosis, with the majority classified as the basal-like subtype associated with epithelial-mesenchymal transition and metastasis. Because basal breast cancers originate from proliferative luminal progenitor-like cells upon dysregulation of proper luminal differentiation, genes regulating luminal-basal transition are critical to elucidate novel therapeutic targets to improve TNBC outcomes. Herein we demonstrate that the tumor suppressor DEAR1/TRIM62 is a critical regulator of luminal cell fate. DEAR1 loss in human mammary epithelial cells results in significantly enhanced mammosphere formation that is accelerated in the presence of TGF- $\beta$ /SMAD3 signaling. Mammospheres formed following DEAR1 loss are enriched for ALDH1A1 and CK5 expression, EpCAM<sup>-</sup>/CD49f<sup>+</sup> and CD44<sup>high</sup>/24<sup>low</sup> basal-like epithelial cells, indicating that DEAR1 regulates stem/progenitor cell properties and luminal-basal progenitor transition. We show that DEAR1 maintains luminal differentiation as a novel ubiquitin ligase for SNAI2/SLUG, a master regulator driving stemness and generation of basal-like progenitor populations. We also identify a significant inverse correlation between DEAR1 and SNAI2 expression in a 103 TNBC case cohort and show that low DEAR1 expression significantly correlates with young age of onset and shorter time to metastasis, suggesting DEAR1 could serve as a biomarker to stratify early onset TNBCs for targeted stem cell therapies.

Triple negative breast cancer (TNBC) is a highly aggressive form of breast cancer lacking expression of the estrogen and progesterone receptors as well as the master oncogene HER2 (ER<sup>-</sup>, PR<sup>-</sup>, and HER2<sup>-</sup>). TNBC has a high risk of metastasis, a poor prognosis and lacks efficacy for conventional targeted therapies commonly utilized for treatment of ER<sup>+</sup>/PR<sup>+</sup> and HER2<sup>+</sup> breast cancer subtypes<sup>1</sup>. Once TNBC progresses to metastasis, the median survival is only 13 months; ultimately, metastatic TNBC patients die from their disease. TNBC also tends to occur in patients of younger age, making it a high priority for the field to identify predictive biomarkers to stratify TNBCs for targeted therapies as well as early detection strategies to prevent metastasis from this devastating disease<sup>2,3</sup>.

A contributing factor to the dismal prognosis associated with TNBC is its intratumor heterogeneity, derived from both differentiation-state heterogeneity and cellular-state plasticity, which hinders the effectiveness of therapeutic treatments<sup>4</sup>. The normal adult mammary gland epithelium controls the differentiation and organization of a variety of epithelial cell states therein such as basal, mature luminal, and luminal progenitor cell populations. While the exact differentiation hierarchy of these cell types is still unclear, studies suggest that they share common progenitors and can transition from one to another<sup>5</sup>. Genetic and microenvironmental perturbations alter the normal mammary epithelial differentiation cascade, leading to changes in cellular reprogramming, cellular plasticity, transitions, and shifts in occupancies of various cellular states within the differentiation landscape<sup>6–10</sup>.

<sup>1</sup>Department of Translational Molecular Pathology, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA. <sup>2</sup>Department of Biostatistics, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA. <sup>3</sup>Department of Pathology, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA. ✉email: akillary@mdanderson.org

As a result, a single tumor contains mixtures of phenotypically diverse cell subpopulations that transition along various positions in a multidimensional scale along luminal to basal, differentiated to stem-like, and epithelial to mesenchymal axes<sup>11,12</sup>. Where a cell falls within this space, the dynamics of which are shaped by the particular genetic or microenvironmental influences, is thought to determine its distinct role (i.e. in proliferation, invasion, etc.) in tumorigenesis<sup>13</sup>, as well as its sensitivity to therapies<sup>14</sup>. The co-occurrence of cells across various cell states within a given tumor population has been observed in breast cancer, made possible by the resolution afforded by in vivo mouse models and single cell experiments, suggesting a much more plastic tumor cell state occupying partial or intermediate/hybrid epithelial-mesenchymal transition (EMT) states along the EMT spectrum and could better explain the interplay between stemness and EMT<sup>15,16</sup>.

The vast majority of TNBCs are categorized as overlapping with the basal-like breast cancer (BLBC) subtype, which is characterized by the expression of basal/myoepithelial signatures and mesenchymal markers suggestive of EMT occurrence<sup>10,17–19</sup>. The BLBC subtype is thought to originate from proliferative luminal progenitor-like cells, which harbor driver mutations (e.g. *BRCA1*) causing their dedifferentiation and transition towards a basal state<sup>10,20</sup>. Luminal progenitor-like epithelial subpopulations in normal breast cells can spontaneously generate mesenchymal-like cells with partial myoepithelial traits, and invasive behavior, through EMT and inhibition of luminal differentiation by *SNAI2* and *ZEB1*<sup>21</sup>. The inherent phenotypic plasticity in normal progenitor cells, maintained through the inhibition of luminal cell differentiation by factors such as *SNAI2*<sup>6</sup>, is exploited in a cancerous state for the generation of a functionally diverse set of cell states required for the various stages of tumorigenesis (e.g. invasion and metastasis)<sup>6,21</sup>.

TRIM family member proteins have been shown to play a role in regulating acquisition and maintenance of stem-like states, including in cancer cells, through their regulation of core transcription factors<sup>22–25</sup>. We previously identified the E3 ubiquitin ligase *DEAR1/TRIM62* as a tumor suppressor and master regulator of acinar morphogenesis and cell polarity, as well as a negative regulator of TGF- $\beta$ -mediated EMT and migration in immortalized, non-tumorigenic HMECs<sup>26,27</sup>. *DEAR1* negatively regulates the expression of genes related to EMT and mammary epithelial cell fate, such as *SNAI1/2*, *ZEB1/2*, and *TWIST*, via ubiquitination of their upstream activator *SMAD3*<sup>26,28,29</sup>. Interestingly, *DEAR1* expression by immunohistochemistry in an early onset, node negative breast cancer cohort (n = 123) correlated with 95% local recurrence free survival over 15 years post-surgery as well as loss of *DEAR1* expression significantly correlated with the TNBC phenotype (52/123 cases)<sup>27</sup>. Altogether, our previous work led us to hypothesize that *DEAR1* loss is associated with improper luminal cell fate specification.

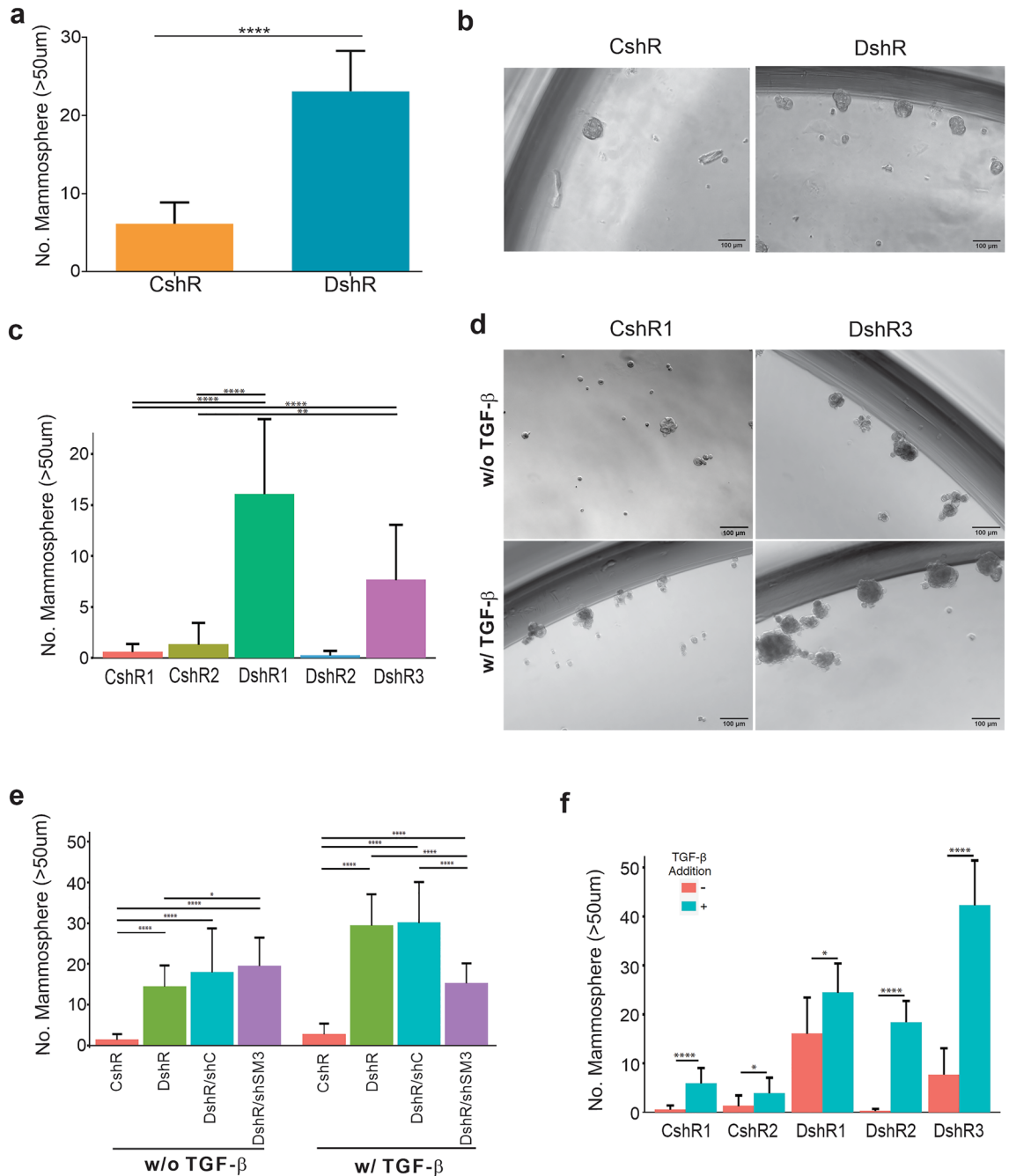
In the present study, cumulative results document compelling evidence that *DEAR1* is an important, novel regulator of luminal cell fate and also provide evidence for its potential clinical utility as a biomarker to stratify TNBC patients for risk of metastasis or targeted therapies aimed at the pathways and cell plasticity regulated by *DEAR1*.

## Results

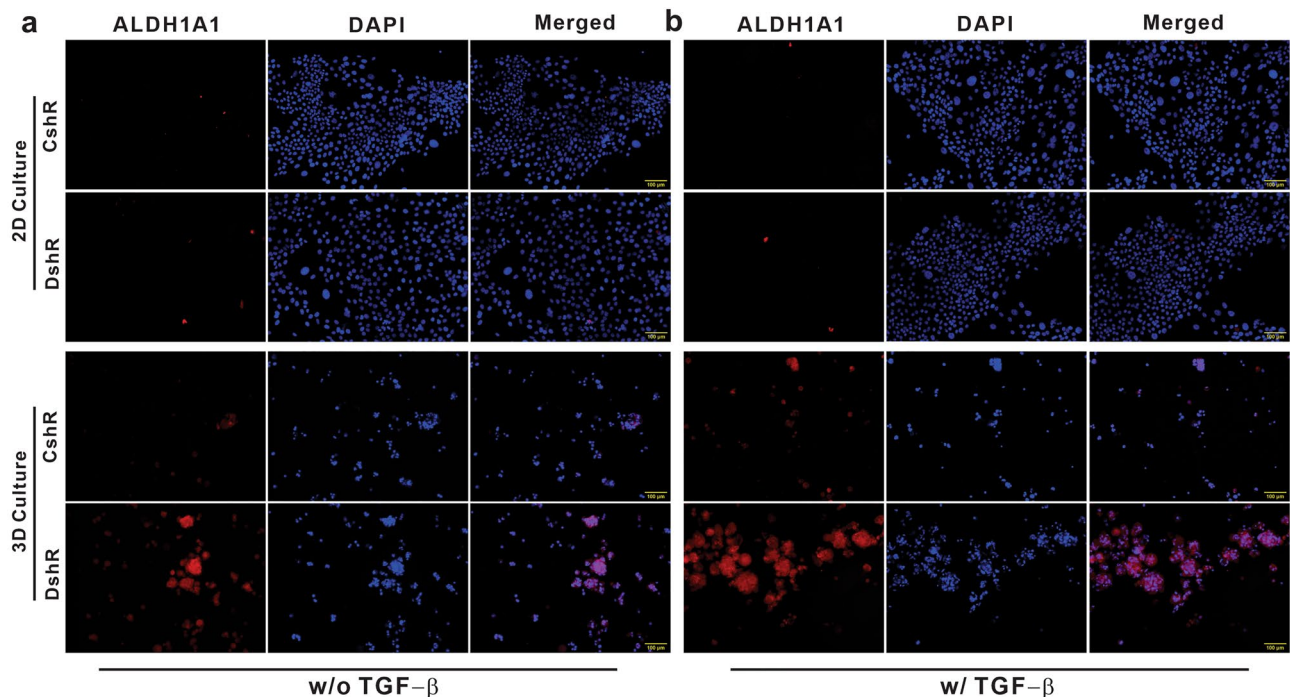
**Loss of *DEAR1* expression in HMECs enhances primary mammosphere formation.** Previous studies have provided strong evidence that inhibition of luminal differentiation by *SNAI2* and *ZEB1* leads to maintenance of luminal-progenitor-like states in HMECs<sup>21</sup>. Because our previous studies suggested that loss of *DEAR1* resulted in increased *SMAD3* and concomitant upregulation of master EMT transcriptional activators<sup>26</sup>, we hypothesized that the tumor suppressor *DEAR1* plays a role in maintaining the luminal differentiation state in HMECs. To address this hypothesis, we utilized the mammosphere assay to examine *DEAR1*'s role in regulating populations with stem or progenitor potential. Stable lentivirus-mediated *DEAR1* knockdown (KD) clones in MCF10A (MCF10A-DshR)<sup>26</sup> were propagated under mammosphere culture conditions with each cell line grown in at least 8 wells of 96-well plates for 17 days for each individual trial. Results (Fig. 1a,b) document mean counts from three separate trials/cell lines and indicate that controls with wild type *DEAR1* expression formed few mammospheres; however, a significant increase in mammosphere formation was observed in MCF10A-DshR (average of 3.8 fold increase) compared to control vector cells (MCF10A-CshR) ( $23.11 \pm 5.18$  compared to  $6.13 \pm 2.75$ , p-value < 0.0001) (Fig. 1a,b). We conducted a similar experiment using 3 *DEAR1* KD 76N-E6 human mammary epithelial cell (HMEC) (E6-DshR) clones<sup>26</sup> in which 2 of 3 *DEAR1* KD clones (DshR 1 and DshR 3) generated a significantly higher number of mammospheres compared to both control vector clones (p < 0.05, Fig. 1c,d). Results indicate that *DEAR1* expressing E6-CshR HMECs maintain the luminal differentiation state; however, loss of *DEAR1* expression drives significant stem/progenitor cell activity in two different HMEC cell lines.

We have previously established that *DEAR1* blocks TGF- $\beta$ -induced EMT by binding specifically to *SMAD3*, resulting in its polyubiquitination, and that loss of *DEAR1* unbridles the TGF- $\beta$ -*SMAD3* axis to elicit *SMAD3*-directed EMT<sup>26</sup>. To determine if the mammosphere formation is dependent on *SMAD3* signaling, we utilized *DEAR1-SMAD3* double KD HMECs (E6-DshR/shSM3) that we previously reported<sup>26</sup> to perform mammosphere assays. We then analyzed the mean number of mammospheres formed from CshR, DshR, DshR/shC, and DshR/shSM3 cells without TGF- $\beta$  treatment across three trials (Fig. 1e left panel). Results indicated that loss of *DEAR1* (DshR) significantly enhanced mammosphere formation, but in double knockdowns, loss of *SMAD3* did not rescue mammosphere formation to control levels (p < 0.0001, Fig. 1e left panel). There was no significant difference in the number of mammospheres formed from E6-DshR/shSM3 cells compared to *DEAR1* KD *SMAD3* shRNA controls (E6-DshR/shC) ( $19.46 \pm 6.95$  vs.  $17.92 \pm 10.75$ , p-value = 0.345, Fig. 1e left panel). Thus, downregulation of the *DEAR1/SMAD3* axis did not fully rescue mammosphere formation, indicating that *DEAR1* functions in part independently of the TGF- $\beta$ -*SMAD3* axis to negatively regulate stem or progenitor functions in HMECs.

**TGF- $\beta$  accelerates the stem/progenitor phenotype in *DEAR1* KD HMECs.** Because published evidence supports a connection between EMT-inducers and progenitor states<sup>21</sup>, we examined the effect of TGF- $\beta$



**Figure 1.** Stable *DEAR1*-KD in HMECs results in increased numbers of mammosphere independent of SMAD3 expression. **(a)** Representative bar plots showing mean  $\pm$  SD counts of primary mammospheres (>50µm) from control vector (CshR) and *DEAR1*-KD (DshR) MCF10A cells across three trials. Each cell line was grown in at least 8 wells of 96-well plate in mammosphere growing conditions for 17 days for each individual trial. P-values were calculated using a two-tailed Student t-test between groups. \*\*\*\*  $p < 0.0001$ . **(b)** Representative images of mammospheres from **(a)** taken using a Zeiss microscope at  $\times 10$  magnification (Scale bars equal 100 µm). **(c)** Bar graphs representing the mean  $\pm$  SD number of mammospheres (> 50 µm) formed from two control vector clones (CshR) and three *DEAR1*-KD clones (DshR) from 76N-E6 cells grown in mammosphere-growing conditions for 14 days across three trials. P-values were calculated using Wilcoxon Rank Sum tests. \*\* $p < 0.01$  and \*\*\*\* $p < 0.0001$ . **(d)** Representative images at  $\times 10$  magnification of CshR1 (left) and DshR3 (right) mammospheres quantified in panel **(c)** in the absence (top) and presence (bottom) of TGF-β (Scale bars equal 100 µm). **(e)** Bar plots showing the mean  $\pm$  SD number of mammospheres formed from CshR, DshR, DshR/shC, and DshR/shSM3 cells without (left) and with (right) TGF-β treatment across 3 trials. p-values calculated using Wilcoxon Rank Sum tests. \* $p < 0.05$ –0.1 and \*\*\*\* $p < 0.0001$ . **(f)** Bar graphs represent the mean  $\pm$  SD number of mammospheres (> 50 µm) formed from two control vector clones (CshR) and three *DEAR1*-KD clones (DshR) from 76N-E6 cells treated with TGF-β and grown in mammosphere growing conditions for 14 days across three trials. P-values were calculated using Wilcoxon Rank Sum tests. \* $p < 0.05$ –0.1 and \*\*\*\* $p < 0.0001$ . The barplots in **(c)**, **(e)** and **(f)** were created using RStudio (Version 1.0.143), <https://www.rstudio.com>.



**Figure 2.** The effect of *DEAR1*-KD on ALDH1A1 expression in 2D culture and 3D mammosphere conditions. **(a)** ALDH1A1 (red) and DAPI (blue) staining of 76N-E6 control (CshR) and *DEAR1*-KD clones (DshR) grown in 2D (top) or 3D (bottom) mammosphere culture in the absence of TGF- $\beta$ . **(b)** ALDH1A1 staining of CshR and DshR clones grown in 2D (top) or 3D (bottom) mammosphere culture in the presence of TGF- $\beta$ . Fluorescence images were taken using OLYMPUS cellSens Software. Scale bars equal 100  $\mu$ m.

on mammosphere number in the same experiment using *DEAR1* KD E6-DshR HMECs and controls with and without prior culture with TGF- $\beta$  to determine if loss of *DEAR1* in the presence of TGF- $\beta$  would further increase mammosphere formation. Results indicated that treatment of 76N-E6 *DEAR1* wild type control cells with 4 ng/mL of TGF- $\beta$  for 48 h prior to plating significantly increased mammosphere numbers compared to control HMECs without TGF- $\beta$  treatment, although the total number of mammospheres formed remained low ( $p < 0.0001$  and  $p < 0.05$ , Fig. 1f). Results demonstrate that *DEAR1* maintains HMEC luminal state in control clones in which most cells failed to form mammospheres, although an increase in number was observed following supplementation of cultures with TGF- $\beta$ . *DEAR1* KD in the presence of TGF- $\beta$  resulted in significantly increased mammosphere numbers compared to *DEAR1* KD cells without exposure to TGF- $\beta$  in 3/3 clones ( $p < 0.05$ ,  $p < 0.0001$ , and  $p < 0.0001$ , respectively Fig. 1f), indicating that loss of *DEAR1* enhances the stem cell/progenitor phenotype which is further amplified in the presence of TGF- $\beta$ .

To determine whether the increase in mammosphere number in response to TGF- $\beta$  was directed through SMAD3, mammosphere growth in TGF- $\beta$ -treated *DEAR1*-SMAD3 double knockdown cells (E6-DshR/shSM3) and *DEAR1* KD SMAD3 shRNA controls (E6-DshR/shC) was also analyzed. In the presence of TGF- $\beta$ , E6-DshR/shSM3 cells formed significantly fewer primary mammospheres compared to E6-DshR/shC vector controls ( $15.25 \pm 4.91$  v.  $30.13 \pm 10.01$ ,  $p < 0.0001$ , Fig. 1e right panel); however, mammosphere formation was not completely rescued to HMEC control levels ( $p < 0.0001$ , Fig. 1e, right panel). This partial rescue of mammosphere formation in double knockdown cells provides evidence that, when *DEAR1* expression is lost, the TGF- $\beta$ -SMAD3 pathway contributes to the stem/progenitor phenotype, but it is not the sole pathway leading to a stem/progenitor cell state. Thus, *DEAR1* plays a role in regulating luminal differentiation both through its regulation of the TGF- $\beta$  pathway and also independently of this axis.

**Loss of *DEAR1* generates mammospheres enriched for progenitor cell activity.** Increased mammosphere formation results from stem/progenitor cell activity and the expansion of cells with mesenchymal/myoepithelial features<sup>30</sup>. To determine if *DEAR1* plays a role in the regulation of stem/progenitor activity, we examined *DEAR1* E6-CshR and E6-DshR HMEC clones for expression of aldehyde dehydrogenase (ALDH1A1), an established marker of stem/progenitor cells that is overexpressed in both normal and cancer stem cells<sup>31</sup>. Results indicated that, in the absence of TGF- $\beta$ , ALDH1A1 expression was barely detectable in both E6-CshR and E6-DshR clones grown in 2D culture conditions (Fig. 2a, in 2D culture panels). However, E6-DshR *DEAR1* KD clones in 3D mammosphere conditions showed a clear increase in ALDH1A1 signal (Fig. 2a in 3D culture panels). Furthermore, TGF- $\beta$  treatment resulted in further elevation of ALDH1A1 expression in 3D culture (Fig. 2b, right panel). These data indicate that *DEAR1* deficiency plays a more important role in inducing stem/progenitor properties in the more intrinsic 3D culture conditions rather than in 2D conditions in TGF- $\beta$ -independent and dependent ways. Interestingly, a small fraction of control HMECs E6-CshR also stained positive for ALDH1A1 in both 2D and 3D culture (Fig. 2a,b), as has been reported<sup>21</sup>.

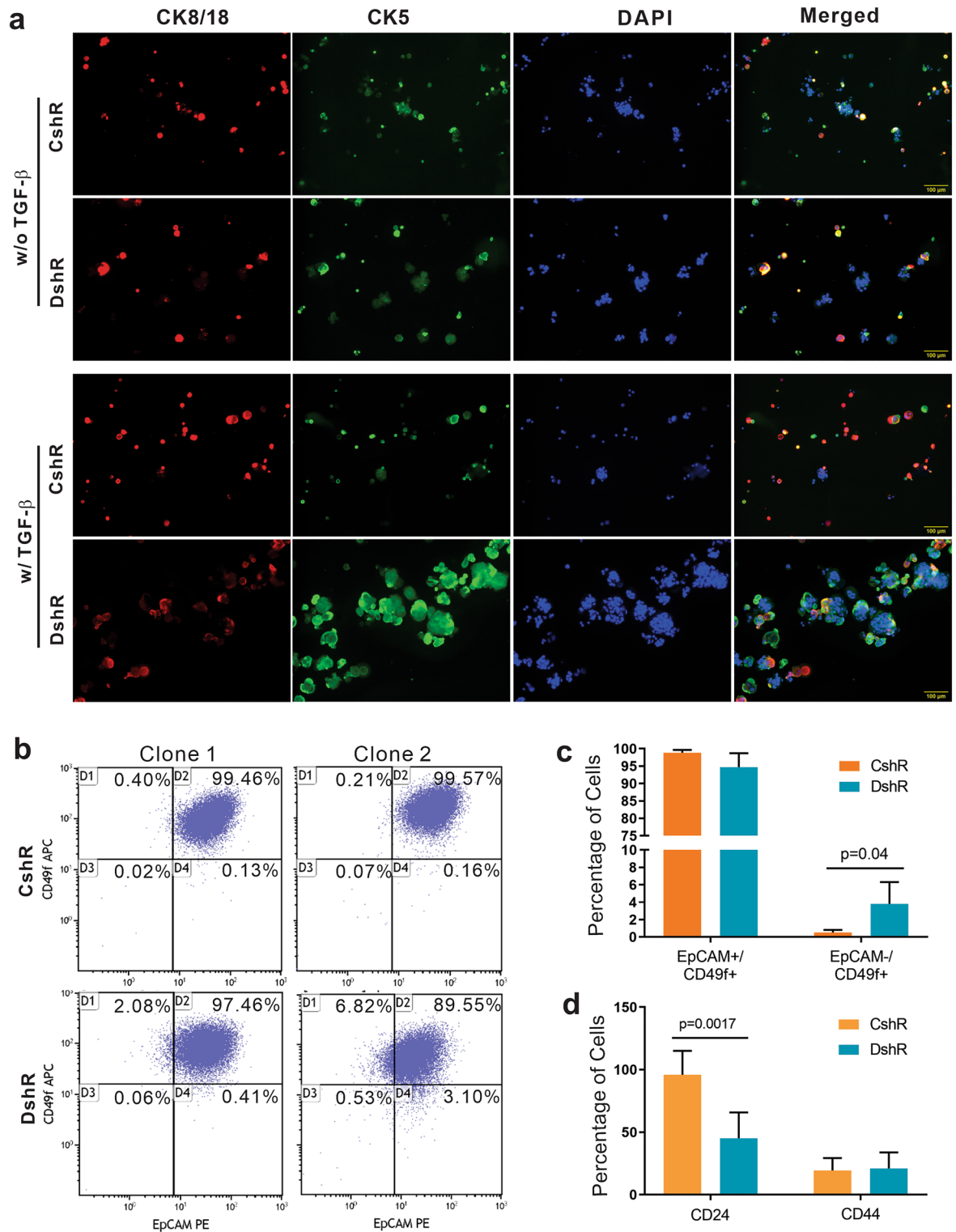
To determine if loss of DEAR1 contributes to self-renewal, primary mammospheres were passaged through dissemination of spheres into single cell suspension and re-plated at the same density to form secondary and tertiary mammospheres. Mammosphere forming efficiency (MFE) was calculated as a percentage of the number of mammospheres larger than 50  $\mu\text{M}$ /total number of single cells seeded per well. Results indicated that in the absence or presence of TGF- $\beta$ , loss of DEAR1 expression in HMECs did not result in an increase of the MFE with increasing passages from primary mammospheres (Supplementary Table S1). Thus, our data indicates that loss of DEAR1 in HMECs does not result in stem cell self-renewal, a major stem cell phenotype; rather, DEAR1 plays a role in regulating progenitor cell properties or state occupancy as *DEAR1*-KD resulted in many more mammospheres expressing ALDH1A1.

**Loss of DEAR1 expands subpopulations of cells expressing basal markers.** Epithelial progenitor populations in human breast tissue primarily consist of luminal, basal and bipotent progenitor cells<sup>32</sup>. Because epithelial ALDH-expressing proliferative cells are known to display multi-lineage differentiation capacity<sup>13</sup>, we next assayed the expression of markers of various cell lineages to examine which cell lineages are upregulated by loss of DEAR1. Mammospheres were stained for cytokeratin (CK) 8/18, a luminal-specific cell surface marker, and CK5, a myoepithelial/basal-specific cell surface marker<sup>33</sup>. Results indicated that mammospheres contained a small population of bipotent cells in both DEAR1 control and KD clones, in the absence and presence of TGF- $\beta$  (Fig. 3a, yellow in the merged figures). However, in the presence of TGF- $\beta$ , *DEAR1*-KD cells showed intense staining with myoepithelial marker CK5, which was not observed for CK8/18, indicating that TGF- $\beta$  promotes myoepithelial/basal-like progenitor features in *DEAR1*-KD cells (Fig. 3a, TGF- $\beta$  panel). Staining for epithelial cell adhesion molecule (EpCAM) and  $\alpha 6$  integrin (CD49f) identifies luminal and basal/myoepithelial cells from breast tissues. Mature luminal cells are characterized by EpCAM<sup>+</sup>/CD49f<sup>-</sup> staining, EpCAM<sup>+</sup>/CD49f<sup>+</sup> identifies luminal progenitor cells, while EpCAM<sup>-</sup>/CD49f<sup>+</sup> identify basal progenitor cells<sup>32</sup>. FACS analysis using EpCAM and CD49f revealed an increased proportion of the EpCAM<sup>-</sup>/CD49f<sup>+</sup> cell population in two *DEAR1* knockdown clones compared to control vector clones (3.82%  $\pm$  2.5 v. 0.54%  $\pm$  0.28,  $p = 0.04$ , Fig. 3b,c, and Supplementary Fig. S1). This result indicates that in *DEAR1* KD cells there is a shift away from differentiated luminal cells toward more basal progenitor and/or bipotent cell types<sup>19</sup>. To further explore this shift we performed FACS analysis using CD44 and CD24 surface markers to examine whether a shift from luminal to basal-like cells as characterized by lower CD24 expression and a higher CD44:CD24 ratio, a cancer stem cell index<sup>34</sup>, was evident. Results revealed that *DEAR1*-KD clones have a higher CD44:CD24 ratio compared to control vector clones due to a significant loss of CD24 expression ( $p = 0.0017$  vs E6-CshR, Fig. 3d), providing further evidence of a shift of cells toward a more basal/myoepithelial profile<sup>35</sup>. Cumulatively, data indicate that *DEAR1* KD results in an incomplete but significant shift from luminal to more basal-like cell types, consistent with loss of DEAR1 maintenance of the luminal cell lineage.

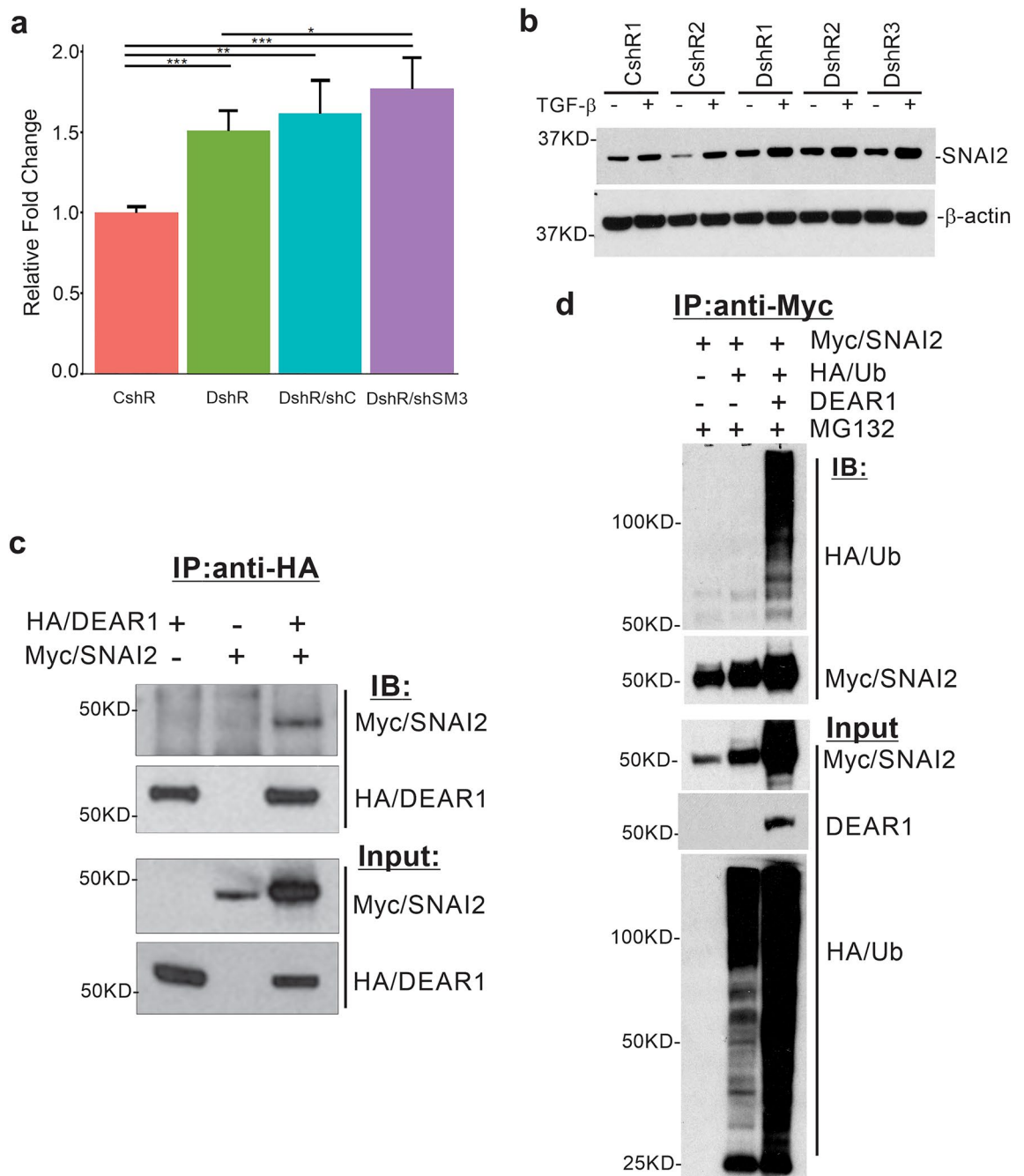
**DEAR1 negatively regulates the master EMT/stemness transcription factor SNAI2 independently of the TGF- $\beta$ -SMAD3 axis.** To investigate the molecular mechanisms underlying the shift towards expansion of basal-like cells and increased mammosphere formation upon DEAR1 loss, we focused on the master regulator of EMT and stemness *SNAI2*. *SNAI2* has been shown to maintain luminal progenitor and mesenchymal phenotypes through inhibition of luminal cell differentiation and we have previously shown it to be upregulated at the mRNA level in HMECs upon knockdown of DEAR1<sup>21,26</sup>. Since we found DEAR1 loss of expression results in mammosphere formation independently of the TGF- $\beta$ -SMAD3 axis (Fig. 1e), we then examined if DEAR1 regulates *SNAI2* independently of the TGF- $\beta$ -SMAD3 axis. Results indicated *SNAI2* mRNA expression levels in *DEAR1*-*SMAD3* double KD (DshR/shSM3) cells remained elevated relative to control (adjusted  $p < 0.001$ , Fig. 4a), similar to the *DEAR1* knockdown alone (DshR), and had no significant difference with DshR SM3-control (DshR/shC) ( $p = 0.31$ , Fig. 4a). These results demonstrate that loss of DEAR1 can confer upregulation of *SNAI2* mRNA levels independently of SMAD3 signaling. At the protein level, *SNAI2* expression in *DEAR1*-KD HMECs was elevated compared to control vector clones, with enhanced expression in the presence of TGF- $\beta$  (Fig. 4b and Supplementary Fig. S2). Cumulative results indicate that DEAR1 regulates *SNAI2* expression both independently and through the TGF- $\beta$ -SMAD3 axis.

**E3 Ligase DEAR1 binds to and ubiquitinates SNAI2.** We further investigated if DEAR1 directly regulates *SNAI2* at the protein level. Co-immunoprecipitation in HEK293T cells using Myc-*SNAI2* and HA-*DEAR1* expression constructs demonstrated that DEAR1 interacts with *SNAI2* (Fig. 4c and Supplementary Fig. S3). We then co-transfected Myc-*SNAI2*, HA-ubiquitin, and pcDNA-*DEAR1* into HEK293T cells to investigate whether DEAR1 promotes the ubiquitination of *SNAI2*. Myc-*SNAI2* pulldown by anti-Myc was observed as well as polyubiquitination of *SNAI2* was detected using anti-HA tag antibody. Results indicated that cotransfection of DEAR1 promoted *SNAI2* polyubiquitination (Fig. 4d and Supplementary Fig. S4). These results demonstrate that DEAR1 functions independently of the canonical TGF- $\beta$  pathway to regulate *SNAI2* expression and activity by promoting *SNAI2* ubiquitination, thereby also regulating *SNAI2* expression at the protein level in addition to the transcript level.

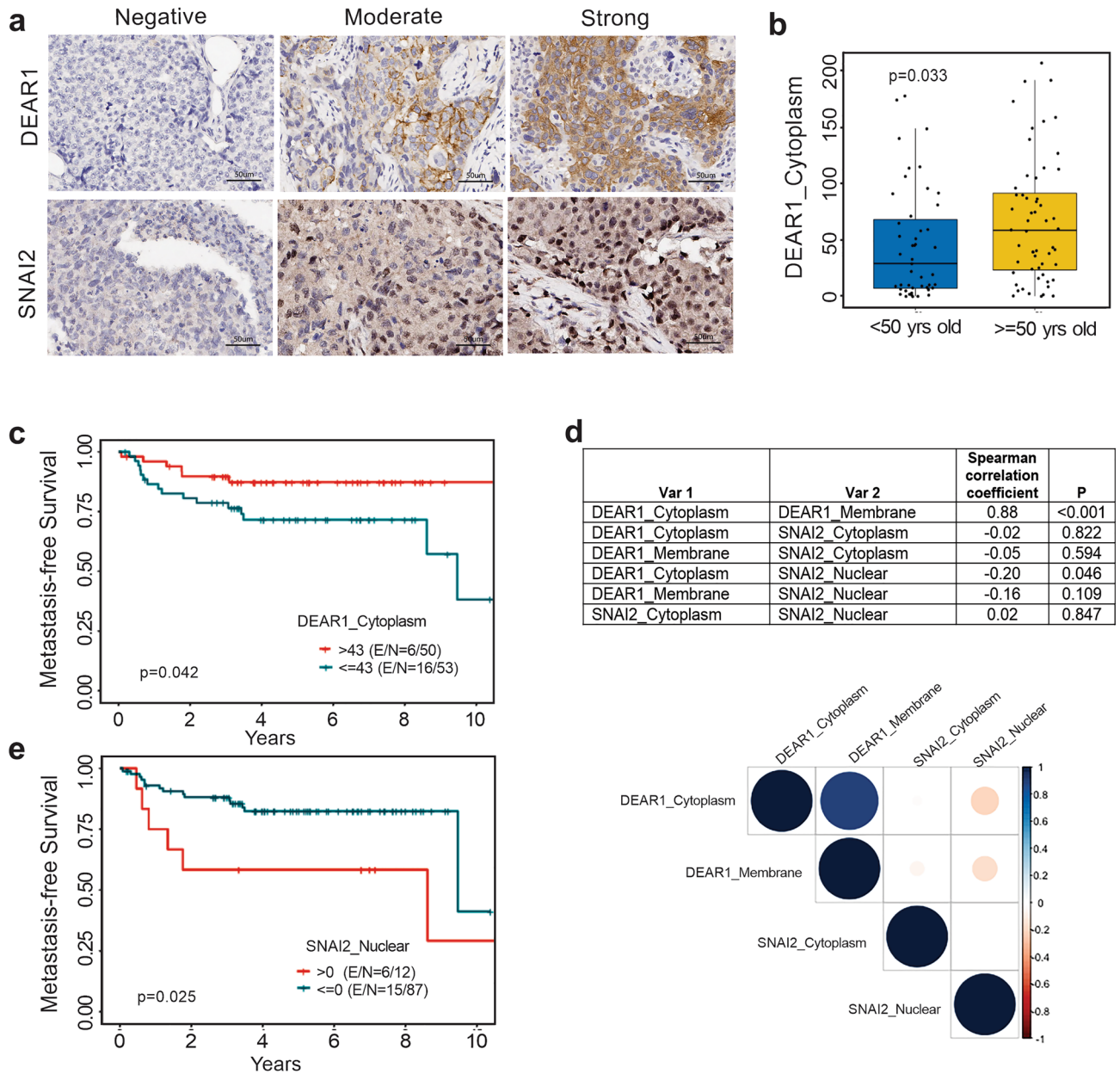
**Loss of DEAR1 expression predicts early age of onset of TNBC and shorter time to metastasis.** Our previous work studied a cohort of 123 early onset breast cancers, which identified that DEAR1 was a strong predictor of local recurrence-free progression, as well as correlated with the triple negative phenotype<sup>27</sup>. As many of these tumors fall under the basal subtype, coupled with our results showing the expansion of progenitor and basal-like cells due to lack of inhibition of *SNAI2* upon DEAR1 loss, we interrogated several publicly available datasets of TNBC and basal-like breast cancers (BLBC) for DEAR1 expression trends. Utilizing



**Figure 3.** The effect of *DEARI*-KD on luminal and myoepithelial/basal surface marker expression in HMEC cells. **(a)** Images at  $\times 10$  magnification of staining of 76N-E6 control (CshR) and *DEARI*-KD (DshR) clones grown in 3D mammosphere culture in the absence (top) or presence (bottom) of TGF- $\beta$  with luminal and myoepithelial/basal cyokeratin surface markers, CK 8/18 (red) and CK 5 (green), respectively. Fluorescence images were taken using OLYMPUS cellSens Software and scale bars equal 100  $\mu$ m. **(b)** FACS data showing proportion of cells that express EpCAM and CD49f from 2 CshR clones (top) and 2 DshR clones (bottom). **(c)** Bar plots showing the mean  $\pm$  SD percentage of cells displaying luminal (EpCAM<sup>+</sup>/CD49f<sup>+</sup>) and basal (EpCAM<sup>-</sup>/CD49f<sup>+</sup>) surface markers from CshR (orange) and DshR (teal) clones. p-values were calculated using a two-tailed Student t-test between groups. **(d)** Bar plots quantifying FACS data assaying CD24 and CD44 protein expression from CshR (orange) and DshR (teal) clones. p-values were calculated using a two-tailed Student t-test between groups.



**Figure 4.** The regulation of SNAI2 by DEAR1. **(a)** Bar plots of qPCR results (mean  $\pm$  SD) of SNAI2 mRNA levels in control (CshR), DEAR1-KD (DshR), DEAR1-KD/shRct clone (DshR/shC), and DEAR1-SMAD3 double KD (DshR-shSM3) across 2 runs. P-values were calculated using Wilcoxon Rank Sum tests. \* $p < 0.05$ –0.1, \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . The barplots were created using RStudio (Version 1.0.143), <https://www.rstudio.com>. **(b)** Western blot showing SNAI2 protein expression in control (CshR) and DEAR1-KD clones (DshR) with and without TGF- $\beta$  treatment. Full length gels for Western blots are shown in Supplementary Fig. S2. **(c)** Co-immunoprecipitation (Co-IP) to determine possible interaction between DEAR1 and SNAI2. HEK293T cells were co-transfected with an HA-DEAR1-expressing plasmid and/or a Myc-SNAI2 expression plasmid. After 24 h post-transfection, cell lysates were immunoprecipitated with anti-HA (rabbit) antibody. SNAI2 proteins were blotted with anti-Myc (mouse) antibody. Full length gels for Western blots are shown in Supplementary Fig. S3. **(d)** Co-IP assay to detect ubiquitination of SNAI2 by DEAR1. HEK293T cells were co-transfected with Myc-SNAI2, HA-ubiquitin, and DEAR1. Cells were lysed 48 h post-transfection. Lysates were denatured and immunoprecipitated with anti-Myc (mouse) antibody. Ubiquitin was blotted with anti-HA (Rat) antibody. Full length gels for Western blots are shown in Supplementary Fig. S4.

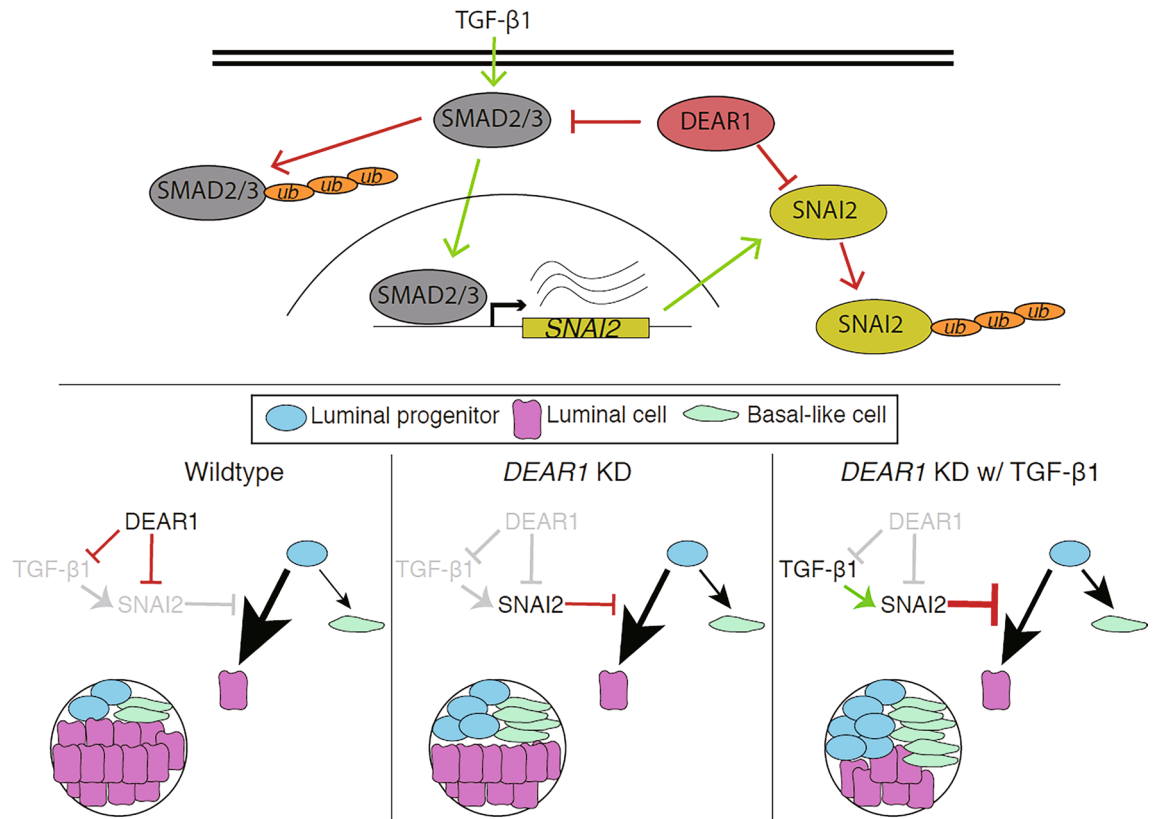


**Figure 5.** DEAR1 and SNAI2 expression analysis in a TNBC TMA cohort. **(a)** Representative images of DEAR1 and SNAI2 IHC in TNBC TMA analysis (Scale bars equal 50  $\mu$ m). **(b)** Boxplot of cytoplasmic DEAR1 expression in groups of patients under (blue) or over (yellow) the age of 50 years. **(c)** Kaplan–Meier analysis of metastasis-free survival in DEAR1 groups dichotomized by the median of cytoplasmic DEAR1 expression (High expression = red, low = blue). **(d)** Spearman’s correlations of DEAR1 and SNAI2 expression across different cellular fractions. Spearman’s correlation was performed using R Statistical Software (version 3.4.2; R Foundation for Statistical Computing, Vienna, Austria). **(e)** Kaplan–Meier analysis of metastasis-free survival in SNAI2 groups dichotomized by presence (red) or absence (teal) of nuclear SNAI2 expression. Both **(c)** and **(e)** were analyzed by the log-rank test.

bc-GenExMiner v4.0<sup>36,37</sup>, we found a significant trend in *DEAR1* mRNA downregulation in basal-like tumors compared to other subtypes ( $p < 0.0001$ , Supplementary Fig. S5a), as well as in the TNBC subgroup ( $p < 0.0001$ , Supplementary Fig. S5b).

Since we observed a significant reduction of *DEAR1* expression in BLBCs and TNBCs, we further investigated the potential of utilizing DEAR1 expression to predict TNBC patient outcomes. In a tissue microarray containing 103 TNBC samples (for patient characteristics, see Supplementary Table S2), we stained for DEAR1 and SNAI2. As illustrated in Fig. 5a, DEAR1 is mainly expressed in the cytoplasm but can also be detected in the membrane, while SNAI2 expression is located in the cytoplasm and nucleus. A statistically significant association between age at diagnosis and DEAR1 cytoplasmic expression was detected. When using an age of 50 as the cutoff, older patients had higher DEAR1 cytoplasm expression compared to younger patients ( $p = 0.033$ , Fig. 5b). Low DEAR1





**Figure 6.** Simplified graphical model of DEAR1 regulation of luminal differentiation via TGF- $\beta$  dependent and independent mechanisms of SNAI2 inhibition. Top: DEAR1 inhibits SNAI2 (red pathway) by direct binding and ubiquitination, as well as by blocking SNAI2 transcription by binding and ubiquitinating upstream activator SMAD3 that is activated by TGF- $\beta$  (green pathway). Bottom: Shifts in luminal differentiation dynamics or state occupancies (as depicted by arrow sizes, where larger size indicates higher likelihood) of luminal progenitors (blue), toward incomplete basal-like (green) or mature luminal cells (pink), depending on the presence (black) or absence (gray) of DEAR1, SNAI2, and TGF- $\beta$ . The figure was created using Adobe Illustrator 2020 (Version 24.2.3), <https://www.adobe.com/products/illustrator.html>.

cytoplasmic expression, when dichotomized by the median value, was also associated with a shorter time to metastasis ( $p=0.042$ , Fig. 5c) in the TNBC patients (see Supplementary Table S3). These data indicate that DEAR1 expression could be useful in stratifying TNBC for risk of developing early-onset TNBC and aggressive metastasis. In line with our results showing DEAR1-mediated SNAI2 ubiquitination (Fig. 4d), we found that DEAR1 cytoplasmic and SNAI2 nuclear levels were negatively correlated in the TMA (Fig. 5d) and the presence of SNAI2 nuclear expression was significantly associated with a shorter time to metastasis ( $p=0.025$ , Fig. 5e).

## Discussion

Our work highlights the role of DEAR1 in maintaining proper luminal differentiation of human mammary epithelial cells, complementing our previous work showing failure of lumen formation and acinar morphogenesis upon knockdown of *DEAR1*<sup>27</sup>. We extend these findings to document herein that loss of DEAR1 significantly upregulates mammosphere formation through enhancing progenitor cell activity, reprogramming the basal/myoepithelial profile, and upregulating the master EMT and stem cell transcription factor SNAI2. SNAI2 has been shown to maintain a luminal progenitor fate by inhibiting luminal differentiation in normal HMECs<sup>21</sup>. Results demonstrate that DEAR1 acts as an upstream inhibitor of this process by directly binding to and ubiquitinating SNAI2. The prolonged time spent in a plastic luminal progenitor-like state (i.e. ALDH-expressing cells) in the absence of DEAR1 allows increased opportunity for the spontaneous generation of a mesenchymal-like CD44<sup>high</sup>/CD24<sup>low</sup> state, which reflects an altered or incomplete myoepithelial differentiation<sup>21</sup>, consistent with our results (Fig. 6). Importantly, this subpopulation is associated with increased mammosphere formation due to its survival ability in anchorage-independent conditions, invasive properties at the tumor front, and is enriched in the BLBC subtype<sup>13,38</sup>. Therefore, loss of DEAR1 disrupts normal mammary epithelial homeostasis, leading to the generation of greater quantities of invasive mesenchymal-like cells with basal signatures<sup>39</sup>. This may be an early event in the tumorigenesis of basal-like breast cancers and provides an explanation for our previous observation of lower DEAR1 expression correlating with the triple-negative phenotype<sup>27</sup>, as well as our current association between lower DEAR1 expression with earlier age of onset and time to metastasis in a large TNBC cohort.

Our current results also indicate that the generation of basal/myoepithelial-like cells from luminal progenitor-like precursors, which occurs through EMT effectors such as SNAI2<sup>21</sup>, is further amplified upon exposure to

TGF- $\beta$  in the absence of DEAR1. This is consistent with our previous work showing DEAR1 as a negative regulator of TGF- $\beta$ -driven EMT and migration in HMEC cells, and by inhibiting SMAD3 and its downstream effectors such as SNAI2<sup>26</sup>. Our results herein document compelling evidence that DEAR1 controls luminal differentiation through a TGF- $\beta$  independent axis with DEAR1 functioning as an E3 ligase to bind to and ubiquitinate SNAI2 independent of TGF- $\beta$ . In addition to repressing the luminal cell state, SNAI2 has been shown to induce mammosphere formation and an expansion of the basal progenitor state<sup>28,40–43</sup>. Thus, DEAR1 regulates luminal cell fate and mammosphere formation through binding to and ubiquitinating SNAI2. Taken together, these data and our previous work indicate DEAR1 functions to inhibit EMT and stemness through multiple mechanisms which now includes limiting SNAI2 expression at both the transcriptional and posttranslational levels (Fig. 6).

Herein we document that DEAR1 governs epithelial plasticity by maintaining luminal differentiation as an ubiquitin ligase for SNAI2, even in the presence of exogenous stimulus from TGF- $\beta$ . Loss of DEAR1 significantly upregulates mammosphere formation and increases the progenitor cell activity and the basal/myoepithelial state. In addition, microenvironmental sources of TGF- $\beta$  contribute to the aggressive and metastatic nature of TNBC by enhancing SNAI2's activity to further increase the plasticity of progenitors and subsequent numbers of invasive mesenchymal-like cancer cells (Fig. 6). In addition to the SMAD3 pathway, TGF- $\beta$  also induces SNAI2 via non-SMAD pathways to drive metastasis via major oncogenic signaling such as PI3K/Akt/mTOR and MAPK (ERK, p38, or JNK) pathways<sup>44–47</sup>. Interestingly, DEAR1/TRIM62 has been reported to suppress tumor metastasis through inhibiting MAPK/JNK-induced SNAI2 expression<sup>48</sup>. This report then, in concert with our published and data reported herein, indicate that DEAR1 is a critical regulator of a major cytokine pathway, blocking TGF- $\beta$  signaling and preventing the induction of SNAI2-driven increase basal-mesenchymal-cell populations, as well as binds directly to SNAI2 to block oncogenic signaling pathways, including reported inhibition of metastatic signaling cascade through SNAI2-MAPK/JNK pathways.

Additionally, results suggest possible important translational potential. We have shown a significant negative correlation between DEAR1 and SNAI2 in a large TNBC cohort. Furthermore, a significant association was observed between lower DEAR1 expression and earlier age of onset and time to metastasis. Hence, DEAR1 may serve as a valuable prognostic and predictive biomarker for highly aggressive forms of TNBC. Our results furthermore suggest impact for therapeutic targeting of TNBC patients. The expanded population of CD44<sup>high</sup>/CD24<sup>low</sup> cells we observed upon DEAR1 loss has been shown to contribute to resistance to multiple chemotherapeutic drugs<sup>14,49</sup>. Moreover, the cell-state plasticity and associated cell-state heterogeneity that contributes to drug resistance in TNBC and BLBC, appears to be targetable by co-treatment with the PI3K/mTOR inhibitor BEZ235 and the BET inhibitor JQ1 in model systems<sup>4</sup>. Therefore, DEAR1 expression could be a valuable biomarker for stratifying patients for such therapies. Future work elucidating the pathways upstream, as well as others downstream of DEAR1 will aid in identifying additional therapeutic targets aimed at cell plasticity and controlled by DEAR1.

## Materials and methods

**Cell culture.** 76N-E6 HMEC clones were grown in D-medium as previously described<sup>26</sup>. HEK 293 T cells purchased from the MD Anderson characterized cell line core (CCLC) were grown in DMEM (Corning, MT10017CV) and 5% FBS (Sigma, F8192). MCF10A cells were cultured in medium as previously described<sup>26</sup>. Cell lines were authenticated by STR profiling and all cell lines are routinely checked for mycoplasma contamination using the Lonza MycoAlert kit (LT07-118).

**Western blot analysis.** To obtain whole cell lysates, cells were harvested using 1× SDS sample buffer with 50 mM DTT. Equal amount of protein were loaded into a 4–12% SDS-PAGE gradient gel (ThermoFisher Scientific, NW04120BOX), and transferred onto a nitrocellulose membrane (BioRad, 1620094). SNAI2/SLUG (C19G7, #9585) was purchased from Cell Signaling (Danvers, MA, United States). Westerns were normalized using  $\beta$ -actin (A5441) purchased from Sigma (Saint Louis, MO, United States).

**Plasmids.** The pcDNA-DEAR1 and pCMV-HA-DEAR1 constructs were generated as previously described<sup>26</sup>. SNAI2 cDNA was amplified from the pEGFP-C2-SNAI2 vector (kindly provided by Dr. Kaname Kawajiri, Saitama Cancer Center, Saitama, Japan) using PCR and ligated into a pcDNA-6Myc vector digested with EcoRI and XhoI. pMT123-8x-HA-Ubiquitin was kindly provided by Dr. Dirk Bohmann (Department of Biomedical Genetics, University of Rochester Medical Center, Rochester, New York).

**Mammosphere assay.** HMECs (MCF10A or 76N-E6) were grown in 2D culture in normal growing conditions. The mammosphere assay was followed as previously described with some modifications. The mammosphere medium contained Mammary Epithelial Growth Medium Complete (LONZA, CC-3051, without BPE), 20 ng/mL bFGF (Corning, 47743-574), 10 ng/mL EGF (Sigma, E9644), 4  $\mu$ g/mL heparin (Sigma, H3149-50KU), and 1% methylcellulose (Sigma, M0430-100G). When required, cells were pre-treated with 4 ng/mL TGF- $\beta$ 1 for 48 h prior to seeding cells in mammosphere growing conditions. Cells were seeded into 4–8 wells at  $1 \times 10^3$  cells per well in flat bottom ultra-low attachment 96-well plates (Corning, 3474) and propagated according to published protocol<sup>50</sup>. Mammospheres larger than 50  $\mu$ m were counted and imaged using a Zeiss light microscope at 10× magnification on Day 14. Secondary mammospheres generated by harvesting primary mammospheres, were chemically dissociated with 1× TrypLE Express, and mechanically dissociated with a pipette to achieve a single cell suspension. Cells were re-counted, seeded at 500 cells/well and monitored for growth. For these experiments, 2 stable control vector (CshR) clones and 3 DEAR1-KD (DshR) stable clones were used and experiments were performed in triplicate.

**Fluorescence activated cell sorting (FACS).** Control clones and *DEAR1*-KD 76N-E6 clones were grown in 2D culture and stained for FACS using EpCAM-PE (eBioscience, 12-9326-41), CD49f-APC (eBioscience, 17-0495-80), CD24-PE (BD Biosciences, 555428), and CD44-FITC (BD Biosciences, 555478) according to the manufacturer's Staining Cell Surface Antigens for Flow Cytometry protocol. Plated cells were trypsinized, counted, and resuspended in Flow Cytometry Staining Buffer for a final concentration of  $1 \times 10^7$  cells/mL. 50  $\mu$ L of the cell suspension was incubated with primary antibody so that the final volume was 100  $\mu$ L per sample. The cells incubated for 30 min in the dark at 4 °C, washed, and analyzed. For these experiments, 2 stable control vector (CshR) clones and 2 *DEAR1*-KD (DshR) stable clones were used experiments were performed in triplicate.

**Immunofluorescence.** Whole mammospheres were collected and resuspended in 100  $\mu$ L of  $1 \times$  PBS for cytospin centrifugation onto slides using Shandon Cytospin 4 (Thermo Scientific, 1000 rpm, 8 min). Slides were stained using standard immunofluorescence staining protocol using 0.3% Triton X-100 in PBS for permeabilization and blocked with anti-Goat serum for one hour at 4 °C. Anti-CK8/18 (rabbit, Abcam, cat#: ab53280, 1:100), anti-CK5 (mouse, Pierce, cat#: MA5-17057, 1:100), and ALDH1A1 (rabbit, Abcam, cat#: ab52492, 1:100) were used to examine co-expression of luminal and basal cytokeratins and a marker of stem/progenitor cells. Secondary antibodies goat anti-mouse (Alexa 488) and goat anti-rabbit (Alexa 555) were used. Nuclei were stained with DAPI (Sigma, 236276). Slides were treated with ProLong Gold anti-fade fluorescent mounting agent (ThermoFisher Scientific, P10144). Fluorescent images were taken using OLYMPUS cellSens Software.

**Primers.** To amplify *SNAI2* from the pEGFP-C2 backbone via PCR, forward and reverse primers were designed to include EcoRI and XhoI restriction enzyme cut sites, respectively, for ligation into the pcDNA-6Myc backbone (FW primer, 5'-TAACGAGAATTCATGCCGCGCTCCTTC-3'; RV primer, 5'-TCGTTACTCGAGTCAGTGTGCTACACAGCA-3').

**Co-immunoprecipitation and ubiquitination assays.** For co-immunoprecipitation (Co-IP) assays, HEK293T cells were transfected with Myc-SNAI2 and/or HA-DEAR1 plasmids in a 6 cm plate using Mirus TransIT-LT1 transfection reagent (Mirus, MIR 2300) for 24 h. Cells were treated with MG-132 proteasome inhibitor (Calbiochem, 133407-82-6) for 2 h, before harvest. Cells were lysed with M-PER lysis buffer (ThermoFisher Scientific, 78501) and incubated with pull-down antibody overnight at 4 °C. Then, lysates were incubated with protein A/G agarose beads (Santa Cruz, sc-2003) for 2 h at 4 °C followed by three 5-min washes with RIPA lysis buffer. Proteins were eluted from beads in 30  $\mu$ L of  $2 \times$  SDS sample buffer for western blot analysis. Rabbit polyclonal anti-HA (H6908) or anti-Myc (M4439) from Sigma were used for pull-down and Mouse monoclonal anti-HA (H3663) from Sigma or anti-Myc (#562) from MBL International Corporation were used for immunoblotting. For in vivo ubiquitination assays, HEK293T cells were transfected with Myc-SNAI2, 8x-HA-Ub, and/or DEAR1 plasmids in a 10 cm plate. At 20 h after transfection, cells were treated with MG132 20  $\mu$ mol/L for 4 h and harvested. Cells were lysed with  $1 \times$  SDS-lysis buffer, denatured by heating at 95 °C for 10 min and then diluted with 0.5% NP-40 buffer and used for further analysis.

**Quantitative real-time PCR (qRT-PCR).** Cells were treated with 4 ng/mL TGF- $\beta$ 1 (EMD Millipore, 616450-1UG) for 3 and 40 h, RNA extracted and reverse transcribed using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, 4368814). Real time quantitative PCR was performed using TaqMan<sup>®</sup> Universal Master Mix II, no UNG (Applied Biosystems, 4440047) and TaqMan<sup>®</sup> probes for TRIM62 (Hs00217089\_m1), SNAI2 (Hs00950344\_m1), ZEB1 (Hs00232783\_m1), ZEB2 (Hs00207691\_m1), and GAPDH (Hs02786624\_g1) in triplicate and in MicroAmp optical 384-well plates (Applied Biosystems, 4309849).

**Breast cancer tissue microarray (TMA) construction and patient cohort.** The study was approved by the institutional review board of MD Anderson Cancer Center (IRB protocol number: PA14-0318), and all methods were carried out in accordance with approved guidelines. Informed consent was obtained from all patients to use residual tissue for research. Archival, formalin-fixed and paraffin-embedded (FFPE) material from surgically resected breast cancer specimens were obtained from the Breast Tumor Bank at M. D. Anderson Cancer Center from 2001 to 2013. Tumor tissues were histologically examined, classified using the World Health Organization (WHO) classification of Breast Tumors and selected for TMA construction, after which TMAs were prepared using triplicate 1-mm-diameter cores per tumor.

**Patient characteristics.** Detailed clinical and pathologic information, including demographic, pathologic TNM staging, overall survival, and time of recurrence was collected according to Institutional Review Board guidelines with informed consent. A total of 103 TNBC patients were included in the analysis, and the average age was 54 (range 26, 80).

**Immunohistochemistry (IHC) for breast cancer tissue microarray (TMA).** DEAR1 staining was performed as previously described<sup>26</sup>. SNAI2 staining utilized an antibody purchased from Santa Cruz, Biotechnology (A-7, SC-166476, 1:50 dilution) on Leica BOND-MAX. DEAR1 and SNAI2 expression were evaluated using Aperio to determine H-scores. Both DEAR1 and SNAI2 expression in the cytoplasm and nucleus of breast tumor cells on the TMA were evaluated separately using a standard microscope approach. Expression was scored by staining intensity 0 (no staining), 1+ (weak staining), 2+ (moderate staining), or 3+ (strong staining) and the percentage of cells with positive IHC expression as well as an H Score was obtained (0–300).

**Statistics.** All values for in vitro analyses are expressed as mean  $\pm$  standard deviation (SD) and represent three independent trials. Unless otherwise specified, comparisons between groups were performed using a two-tailed Student t-test. In the case of multiple comparisons, p-values were corrected using the Benjamini and Hochberg method. For analyses involving the TMA, Wilcoxon or Kruskal–Wallis rank sum tests were used to assess the difference of biomarkers between groups. Overall survival and time to metastasis were defined as the time between date of surgery and date of death or date of metastasis, respectively. Their survival functions were estimated using the Kaplan–Meier method and log-rank test was used to compare the survival between groups. All tests were 2-sided and p-values < 0.05 was considered statistically significant. R version 3.4.2 was used to perform statistical analyses. Asterisks represent: \*p < 0.05–0.1, \*\*p < 0.01, \*\*\*p < 0.001, and \*\*\*\*p < 0.0001.

Received: 8 September 2021; Accepted: 14 October 2022

Published online: 14 November 2022

## References

- Dent, R. *et al.* Pattern of metastatic spread in triple-negative breast cancer. *Breast Cancer Res. Treat.* **115**, 423–428. <https://doi.org/10.1007/s10549-008-0086-2> (2009).
- Kassam, F. *et al.* Survival outcomes for patients with metastatic triple-negative breast cancer: Implications for clinical practice and trial design. *Clin. Breast Cancer* **9**, 29–33. <https://doi.org/10.3816/CBC.2009.n.005> (2009).
- Bonotto, M. *et al.* Measures of outcome in metastatic breast cancer: Insights from a real-world scenario. *Oncologist* **19**, 608–615. <https://doi.org/10.1634/theoncologist.2014-0002> (2014).
- Risom, T. *et al.* Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer. *Nat. Commun.* **9**, 3815. <https://doi.org/10.1038/s41467-018-05729-w> (2018).
- Anstine, L. J. & Keri, R. A new view of the mammary epithelial hierarchy and its implications for breast cancer initiation and metastasis. *J. Cancer Metastasis Treat.* **5**, 50. <https://doi.org/10.20517/2394-4722.2019.24> (2019).
- Phillips, S. *et al.* Cell-state transitions regulated by SLUG are critical for tissue regeneration and tumor initiation. *Stem Cell Rep.* **2**, 633–647. <https://doi.org/10.1016/j.stemcr.2014.03.008> (2014).
- Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196. <https://doi.org/10.1038/nrm3758> (2014).
- Klevebring, D. *et al.* Sequencing of breast cancer stem cell populations indicates a dynamic conversion between differentiation states in vivo. *Breast Cancer Res. BCR* **16**, R72. <https://doi.org/10.1186/bcr3687> (2014).
- Gupta, P. B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644. <https://doi.org/10.1016/j.cell.2011.07.026> (2011).
- Karaayvaz, M. *et al.* Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nat. Commun.* **9**, 3588. <https://doi.org/10.1038/s41467-018-06052-0> (2018).
- Granit, R. Z. *et al.* EZH2 promotes a bi-lineage identity in basal-like breast cancer cells. *Oncogene* **32**, 3886–3895. <https://doi.org/10.1038/onc.2012.390> (2013).
- Granit, R. Z., Slyper, M. & Ben-Porath, I. Axes of differentiation in breast cancer: Untangling stemness, lineage identity, and the epithelial to mesenchymal transition. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **6**, 93–106. <https://doi.org/10.1002/wsbm.1252> (2014).
- Liu, S. *et al.* Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Rep.* **2**, 78–91. <https://doi.org/10.1016/j.stemcr.2013.11.009> (2014).
- Creighton, C. J. *et al.* Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13820–13825. <https://doi.org/10.1073/pnas.0905718106> (2009).
- van Dijk, D. *et al.* Recovering gene interactions from single-cell data using data diffusion. *Cell* **174**, 716–729. <https://doi.org/10.1016/j.cell.2018.05.061> (2018).
- Jolly, M. K. *et al.* Coupling the modules of EMT and stemness: A tunable “stemness window” model. *Oncotarget* **6**, 25161–25174. <https://doi.org/10.18632/oncotarget.4629> (2015).
- Shah, S. P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395–399. <https://doi.org/10.1038/nature10933> (2012).
- Burstein, M. D. *et al.* Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. *Clin. Cancer Res.* **21**, 1688–1698. <https://doi.org/10.1158/1078-0432.Ccr-14-0432> (2015).
- Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res. BCR* **12**, R68. <https://doi.org/10.1186/bcr2635> (2010).
- Lim, E. *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat. Med.* **15**, 907–913. <https://doi.org/10.1038/nm.2000> (2009).
- Sarrio, D., Franklin, C. K., Mackay, A., Reis-Filho, J. S. & Isacke, C. M. Epithelial and mesenchymal subpopulations within normal basal breast cell lines exhibit distinct stem cell/progenitor properties. *Stem Cells (Dayton, Ohio)* **30**, 292–303. <https://doi.org/10.1002/stem.791> (2012).
- Wang, X., Guo, H., Yao, B. & Helms, J. miR-15b inhibits cancer-initiating cell phenotypes and chemoresistance of cisplatin by targeting TRIM14 in oral tongue squamous cell cancer. *Oncol. Rep.* **37**, 2720–2726. <https://doi.org/10.3892/or.2017.5532> (2017).
- Wang, F., Ruan, L., Yang, J., Zhao, Q. & Wei, W. TRIM14 promotes the migration and invasion of gastric cancer by regulating epithelial-to-mesenchymal transition via activation of AKT signaling regulated by miR-195-5p. *Oncol. Rep.* **40**, 3273–3284. <https://doi.org/10.3892/or.2018.6750> (2018).
- Sato, T., Okumura, F., Ariga, T. & Hatakeyama, S. TRIM6 interacts with Myc and maintains the pluripotency of mouse embryonic stem cells. *J. Cell Sci.* **125**, 1544–1555. <https://doi.org/10.1242/jcs.095273> (2012).
- Nicklas, S. *et al.* A complex of the ubiquitin ligase TRIM32 and the deubiquitinase USP7 balances the level of c-Myc ubiquitination and thereby determines neural stem cell fate specification. *Cell Death Differ.* **26**, 728–740. <https://doi.org/10.1038/s41418-018-0144-1> (2019).
- Chen, N. *et al.* DEAR1 Is a chromosome 1p35 tumor suppressor and master regulator of TGF- $\beta$  driven epithelial–mesenchymal transition. *Cancer Discov.* **3**, 1172–1189 (2013).
- Lott, S. T. *et al.* DEAR1 is a dominant regulator of acinar morphogenesis and an independent predictor of local recurrence-free survival in early-onset breast cancer. *PLoS Med.* **6**, e1000068 (2009).
- Phillips, S. & Kuperwasser, C. SLUG: Critical regulator of epithelial cell identity in breast development and cancer. *Cell Adhes. Migr.* **8**, 578–587. <https://doi.org/10.4161/19336918.2014.972740> (2014).
- Chen, N., Balasenthil, S., Reuther, J. & Killary, A. M. DEAR1, a novel tumor suppressor that regulates cell polarity and epithelial plasticity. *Cancer Res.* **74**, 5683–5689 (2014).

30. Borgna, S., Armellin, M., di Gennaro, A., Maestro, R. & Santarosa, M. Mesenchymal traits are selected along with stem features in breast cancer cells grown as mammospheres. *Cell Cycle (Georgetown, Tex.)* **11**, 4242–4251. <https://doi.org/10.4161/cc.22543> (2012).
31. Christophe Ginestier, M. H. H. *et al.* ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* **1**, 555–567. <https://doi.org/10.1016/j.stem.2007.08.014> (2007).
32. Stingl, J., Eaves, C. J., Zandieh, I. & Emerman, J. T. Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. *Breast Cancer Res. Treat.* **67**, 93–109. <https://doi.org/10.1023/a:1010615124301> (2001).
33. Keller, P. J. *et al.* Mapping the cellular and molecular heterogeneity of normal and malignant breast tissues and cultured cell lines. *Breast Cancer Res. BCR* **12**, R87. <https://doi.org/10.1186/bcr2755> (2010).
34. Li, W. *et al.* Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. *Sci. Rep.* **7**, 13856. <https://doi.org/10.1038/s41598-017-14364-2> (2017).
35. Sheridan, C. *et al.* CD44+/CD24–breast cancer cells exhibit enhanced invasive properties: An early step necessary for metastasis. *Breast Cancer Res. BCR* **8**, R59. <https://doi.org/10.1186/bcr1610> (2006).
36. Jezequel, P. *et al.* bc-GenExMiner: An easy-to-use online platform for gene prognostic analyses in breast cancer. *Breast Cancer Res. Treat.* **131**, 765–775. <https://doi.org/10.1007/s10549-011-1457-7> (2012).
37. Jezequel, P. *et al.* bc-GenExMiner 4.5: New mining module computes breast cancer differential gene expression analyses. *Database (Oxford)*. <https://doi.org/10.1093/database/baab007> (2021).
38. Ricardo, S. *et al.* Breast cancer stem cell markers CD44, CD24 and ALDH1: Expression distribution within intrinsic molecular subtype. *J. Clin. Pathol.* **64**, 937–946. <https://doi.org/10.1136/jcp.2011.090456> (2011).
39. Knudsen, E. S. *et al.* Progression of ductal carcinoma in situ to invasive breast cancer is associated with gene expression programs of EMT and myoepithelia. *Breast Cancer Res. Treat.* **133**, 1009–1024. <https://doi.org/10.1007/s10549-011-1894-3> (2012).
40. Storci, G. *et al.* TNFalpha up-regulates SLUG via the NF-kappaB/HIF1alpha axis, which imparts breast cancer cells with a stem cell-like phenotype. *J. Cell. Physiol.* **225**, 682–691. <https://doi.org/10.1002/jcp.22264> (2010).
41. Chimgo, N. O. *et al.* Regulation of breast cancer metastasis by Runx2 and estrogen signaling: The role of SNAI2. *Breast Cancer Res.* **13**, R127. <https://doi.org/10.1186/bcr3073> (2011).
42. Mendoza-Villanueva, D. *et al.* The C/EBPδ protein is stabilized by estrogen receptor α activity, inhibits SNAI2 expression and associates with good prognosis in breast cancer. *Oncogene* **35**, 6166–6176. <https://doi.org/10.1038/ncr.2016.156> (2016).
43. Manne, R. K. *et al.* A microRNA/ubiquitin ligase feedback loop regulates slug-mediated invasion in breast cancer. *Neoplasia* **19**, 483–495. <https://doi.org/10.1016/j.neo.2017.02.013> (2017).
44. Chen, H. *et al.* Extracellular signal-regulated kinase signaling pathway regulates breast cancer cell migration by maintaining slug expression. *Cancer Res.* **69**, 9228–9235. <https://doi.org/10.1158/0008-5472.Can-09-1950> (2009).
45. Tan, E. J., Olsson, A.-K. & Moustakas, A. Reprogramming during epithelial to mesenchymal transition under the control of TGFβ. *Cell Adhes. Migr.* **9**, 233–246. <https://doi.org/10.4161/19336918.2014.983794> (2015).
46. Aomatsu, K. *et al.* TGF-β induces sustained upregulation of SNAI1 and SNAI2 through Smad and non-Smad pathways in a human corneal epithelial cell line. *Investig. Ophthalmol. Vis. Sci.* **52**, 2437–2443. <https://doi.org/10.1167/iovs.10-5635> (2011).
47. Hao, Y., Baker, D. & Ten Dijke, P. TGF-beta-mediated epithelial–mesenchymal transition and cancer metastasis. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms20112767> (2019).
48. Liu, T. Y. *et al.* Tripartite motif containing 62 is a novel prognostic marker and suppresses tumor metastasis via c-Jun/Slug signaling-mediated epithelial–mesenchymal transition in cervical cancer. *J. Exp. Clin. Cancer Res.* **35**, 170. <https://doi.org/10.1186/s13046-016-0445-5> (2016).
49. Ji, P. *et al.* CD44hiCD24lo mammosphere-forming cells from primary breast cancer display resistance to multiple chemotherapeutic drugs. *Oncol. Rep.* **35**, 3293–3302. <https://doi.org/10.3892/or.2016.4739> (2016).
50. Mani, S. A. *et al.* The epithelial–mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715. <https://doi.org/10.1016/j.cell.2008.03.027> (2008).

## Acknowledgements

We thank Dr. Sendurai Mani and lab members for providing assistance with the mammosphere protocol. We thank Dr. Kaname Kawajiri at the Research Institute for Clinical Oncology, Saitama Cancer Center, Japan for the generous gift of the pEGFP-C2-SLUG vector. We thank Dr. Dirk Bohmann from University of Rochester Medical Center, Rochester, New York for the generous gift of the pMT123-8x-HA-Ubiquitin. We also thank Dr. Subrata Sen, Dr. Ralf Krahe, Dr. Sadhan Majumder, and Dr. Kenneth Tsai for their time and valuable input on this project. We also thank Barbara Mino, Jianling Zhou and Wei Lu in our Immunoprofiling Laboratory for IHC technique support.

## Author contributions

U.Q.L. conceived of, designed and performed experiments, analyzed data, and prepared the manuscript. N.C. performed experiments and contributed to manuscript preparation. N.C. and S.B. contributed to experimental design and interpretation of results. E.L. contributed to manuscript preparation, and statistical analyses. F.Y. and A.A.S. collected patient samples for IHC for the TNBC TMA. L.M.S.S., M.G.R. and H.B. carried out IHC and scored the slides. I.I.W. provided support and oversight to patient sample collections. S.L. and L.R. performed statistical analyses. A.M.K. is the corresponding author and conceived with UQL of the overall project, interpreted results, and edited the manuscript. The corresponding author had full access to the data and final responsibility for the decision to submit for publication. All authors approved the final version of the manuscript.

## Funding

This research was supported by the Cancer Prevention Research Institute of Texas Grant Number RP150403. E.L. was supported by the NIH T32 CA217789 to A.M.K. and I.I.W.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-22417-4>.

**Correspondence** and requests for materials should be addressed to A.M.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022