

ARTICLE

<https://doi.org/10.1038/s41467-020-14556-x>

OPEN

# C/EBP $\alpha$ is crucial determinant of epithelial maintenance by preventing epithelial-to-mesenchymal transition

Ana Rita Lourenço<sup>1,2</sup>, M. Guy Roukens<sup>1,2,4</sup>, Danielle Seinstra<sup>3,4</sup>, Cynthia L. Frederiks<sup>1,2,4</sup>, Cornelieke E. Pals<sup>1,2</sup>, Stephin J. Vervoort<sup>1,2</sup>, Andreia S. Margarido<sup>3</sup>, Jacco van Rheenen<sup>3</sup> & Paul J. Coffe<sup>1,2\*</sup>

Extracellular signals such as TGF- $\beta$  can induce epithelial-to-mesenchymal transition (EMT) in cancers of epithelial origin, promoting molecular and phenotypical changes resulting in pro-metastatic characteristics. We identified C/EBP $\alpha$  as one of the most TGF- $\beta$ -mediated downregulated transcription factors in human mammary epithelial cells. C/EBP $\alpha$  expression prevents TGF- $\beta$ -driven EMT by inhibiting expression of known EMT factors. Depletion of C/EBP $\alpha$  is sufficient to induce mesenchymal-like morphology and molecular features, while cells that had undergone TGF- $\beta$ -induced EMT reverted to an epithelial-like state upon C/EBP $\alpha$  re-expression. In vivo, mice injected with C/EBP $\alpha$ -expressing breast tumor organoids display a dramatic reduction of metastatic lesions. Collectively, our results show that C/EBP $\alpha$  is required for maintaining epithelial homeostasis by repressing the expression of key mesenchymal markers, thereby preventing EMT-mediated tumorigenesis. These data suggest that C/EBP $\alpha$  is a master epithelial “gatekeeper” whose expression is required to prevent unwarranted mesenchymal transition, supporting an important role for EMT in mediating breast cancer metastasis.

<sup>1</sup>Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands. <sup>2</sup>Regenerative Medicine Center, Uppsalalaan 8, Utrecht, The Netherlands. <sup>3</sup>Department of Molecular Pathology, Oncode Institute, Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam, The Netherlands. <sup>4</sup>These authors contributed equally: M. Guy Roukens, Danielle Seinstra, Cynthia L. Frederiks. \*email: [pcoffer@umcutrecht.nl](mailto:pcoffer@umcutrecht.nl)

The epithelial-to-mesenchymal transition (EMT) is a highly dynamic and reversible process important during embryonic development, tissue repair, cancer progression, and chemotherapy resistance<sup>1–3</sup>. Although the pathological contribution of EMT during metastasis is currently under debate, recent studies have shown that, by a complex multistep molecular program, EMT clearly contributes to the shift of a polarized epithelial phenotype toward a mesenchymal state, where cells acquire migratory and invasive properties<sup>4,5</sup>. Several key signaling pathways, including TGF- $\beta$ , Wnt, or Notch, have been shown to induce EMT<sup>6</sup>. Binding of TGF $\beta$  to its cognate receptor results in a multistep phosphorylation cascade of both receptors and key players, the Smad2 and Smad3 transcription factors (TFs), where they regulate expression of crucial genes involved in the initiation of the EMT program, including *SNAIL*, *ZEB1*, and *TWIST1*<sup>7</sup>. These transcription factors, known as EMT-TFs, are able to regulate cellular processes, including extracellular matrix remodeling, cell adhesion and migration, and angiogenesis<sup>8</sup>. The majority of these EMT-effectors are responsible for transcriptional repression of *CDH1* (E-cadherin) and induction of N-cadherin expression, which contributes to weaker cell–cell interactions facilitating cell motility and invasion<sup>9</sup>. In addition, mesenchymal markers such as fibronectin and matrix metalloproteinases (MMPs) are upregulated during EMT. Fibronectin is important for transmitting signals from the extracellular matrix (ECM) to the cell cytoskeleton, and MMPs contribute to the remodeling of the ECM by degrading components within the ECM and cell–ECM adhesion proteins<sup>9,10</sup>. Hence, TGF- $\beta$  signaling can contribute to EMT by inducing gene expression of transcription factors involved in epithelial integrity disruption, and acquisition of mesenchymal traits contributing to cytoskeleton reorganization, ECM remodeling, and cell motility.

The transcription factor CCAAT/enhancer-binding protein alpha (C/EBP $\alpha$ ) is the founding member of a subfamily of the basic region leucine zipper (bZIP) transcription factors<sup>11</sup>. It plays a pivotal role in the regulation of cell cycle and in the expression of several lineage-specific genes<sup>12</sup>. Mutations in the *CEBPA* gene have been described in around 10% of acute myeloid leukemia (AML), establishing a tumor suppressor role of C/EBP $\alpha$  in cancer development<sup>13,14</sup>. In addition, extensive sequencing studies revealed that *CEBPA* is one of the 125 genes that when altered by intragenic mutations can contribute to tumorigenesis<sup>15</sup>. Deregulated C/EBP $\alpha$  expression has been observed in a variety of solid tumors such as liver, breast, or lung cancer; however, the relevance of this for tumorigenesis remains largely unclear<sup>16</sup>.

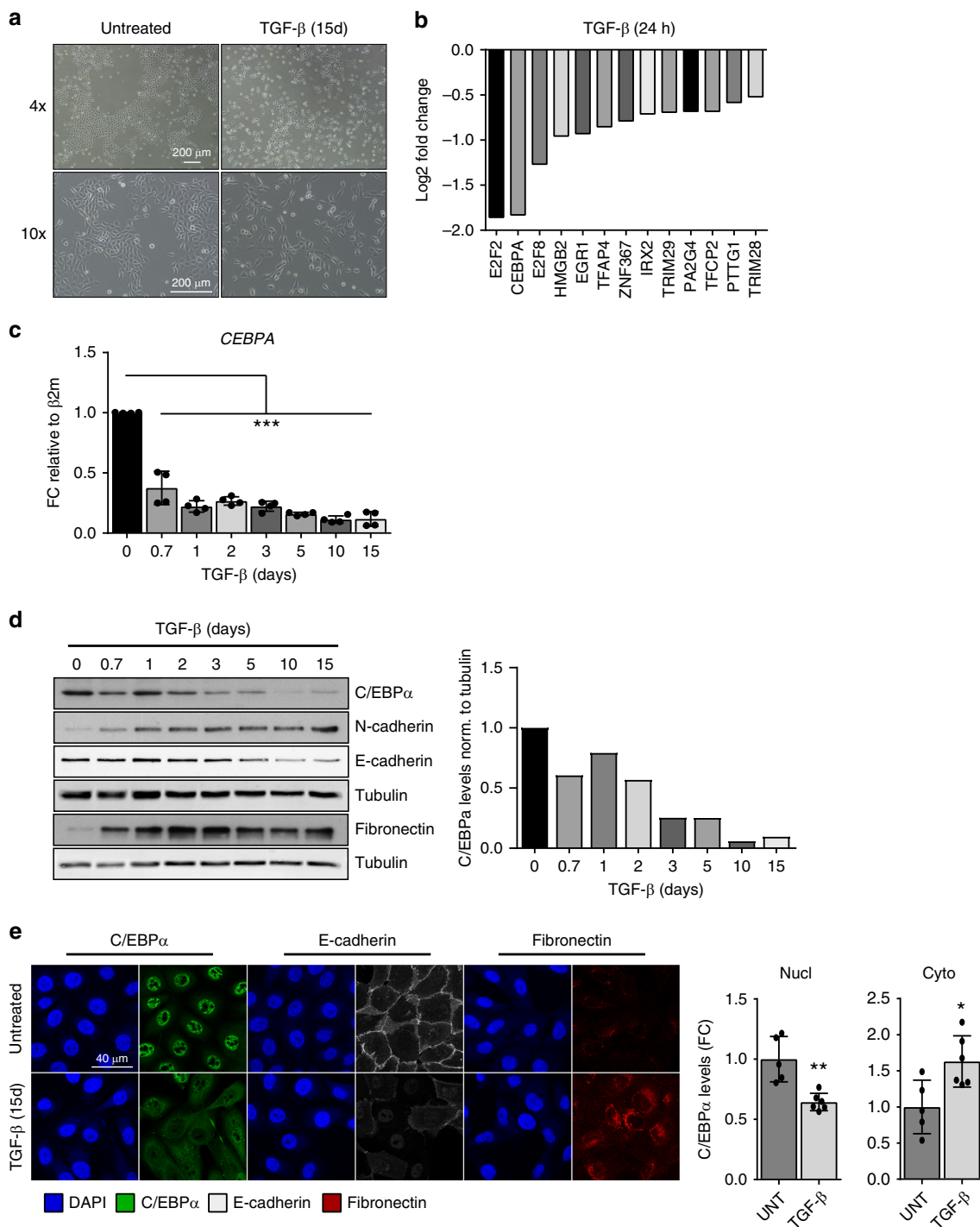
Analysis of C/EBP $\alpha$  expression in healthy breast tissue and breast carcinomas has revealed that C/EBP $\alpha$  levels are reduced in the majority of breast cancer specimens<sup>17</sup>. The functional consequences of these observations remain, however, unknown. Here, we identify C/EBP $\alpha$  as a crucial transcription factor in maintaining epithelial architecture of human mammary cells, preventing epithelial-to-mesenchymal transition and thereby acting as a repressor of breast cancer progression in vivo.

## Results

***CEBPA* is a SMAD3-repressed target during TGF- $\beta$ -mediated EMT.** To identify novel transcription factors with a potential epithelial-gatekeeper function, we performed global RNA-sequencing analysis of TGF- $\beta$ -treated or untreated human epithelial mammary (HMLE) cells. HMLE cells have been extensively used to study EMT, as they can undergo molecular and phenotypical changes upon TGF- $\beta$  treatment resulting in the acquisition of mesenchymal features (Fig. 1a, Supplementary Fig. 1a)<sup>18,19</sup>. To specifically identify transcription factors whose expression is repressed by TGF- $\beta$ , genes displaying significantly

reduced expression were analyzed by Gene-Ontology (GO)-term analysis using DAVID. Further analysis of “transcription factor activity” category revealed that E2F2, C/EBP $\alpha$ , and E2F8 comprise the three transcription factors that are most strongly affected by TGF- $\beta$  treatment (Fig. 1b). Considering that E2F2 and E2F8 are widely expressed cell cycle regulators, we choose to further explore the role of C/EBP $\alpha$  as it has been shown to play a role in cell differentiation, especially during myelopoiesis, adipogenesis, and lung maturation<sup>16,20</sup>. Analysis of RNA-seq profiles of the *CEBPA* locus show a strong decrease in *CEBPA* mRNA levels upon TGF- $\beta$  stimulation (Supplementary Fig. 1b). Furthermore, analysis of the microarray data obtained from Taube et al.<sup>21</sup> also showed decreased *CEBPA* expression in TGF- $\beta$ 1-overexpressing HMLE cells compared with control cells (Supplementary Fig. 1c). To evaluate potential redundancy between the closely related C/EBP $\alpha$  and C/EBP $\beta$ , we analyzed the RNA-seq profile of *CEBPB* locus in TGF- $\beta$ -treated and untreated HMLE cells (Supplementary Fig. 1d). *CEBPB* expression was unaltered upon TGF- $\beta$  stimulation (Supplementary Fig. 1d). Likewise, *CEBPB* mRNA levels were barely changed in HMLE cells treated with TGF- $\beta$  for 24 h (Supplementary Fig. 1e). To determine the effect of TGF- $\beta$  on C/EBP $\alpha$  expression during EMT, HMLE cells were left untreated or treated with TGF- $\beta$  for 15 days, and mRNA and protein were isolated at the indicated time points (Fig. 1c, d). The EMT program was effectively induced by TGF- $\beta$  as illustrated by the increase of mesenchymal markers *CDH2* (N-cadherin), *FNI* (Fibronectin), *VIM* (Vimentin), *SNAIL* (Snail), and *ZEB1*, and decrease of epithelial marker *CDH1* (E-cadherin) (Fig. 1d, Supplementary Fig. 1a). *CEBPA* mRNA expression was rapidly reduced upon TGF- $\beta$  treatment, and reduced levels of *CEBPA* are maintained throughout the 15 days (Fig. 1c; Supplementary Fig. 1f). Furthermore, on the protein level, the expression of C/EBP $\alpha$  full-length (p42) was also found decreased upon TGF- $\beta$ -mediated EMT (Fig. 1d; Supplementary Fig. 1g). Endogenous expression of C/EBP $\alpha$  p30 isoform in the cell lines utilized was undetectable. Unlike C/EBP $\alpha$  p42 isoform, shorter C/EBP $\alpha$  isoform lacks two critical transactivation domains (TAD1 and TAD2) responsible for interaction with elements of the RNA polymerase II basal transcriptional machinery, and therefore is unable to initiate transcription. As the decrease in C/EBP $\alpha$  protein levels upon TGF- $\beta$  treatment was somewhat delayed compared with mRNA, we treated HMLE cells with TGF- $\beta$  for shorter time points and assessed C/EBP $\alpha$  expression and cellular localization by immunofluorescence analysis. Immunostaining results show a gradual increase in cytoplasmic relocalization of C/EBP $\alpha$  after 8–72 h incubation of TGF- $\beta$  (Supplementary Fig. 1h). This suggests that in addition to a rapid transcriptional downregulation of C/EBP $\alpha$  by TGF- $\beta$ , C/EBP $\alpha$  can also be regulated at the posttranslational level leading to its gradual nuclear exclusion. Immunofluorescence analysis demonstrated that 15 days of TGF- $\beta$ -induced EMT results in loss of C/EBP $\alpha$  and E-cadherin expression and increased expression levels of Fibronectin (Fig. 1e). In line with our previous observations, TGF- $\beta$  treatment results in a decrease of nuclear C/EBP $\alpha$  levels and in an increase of C/EBP $\alpha$  protein localized in the cytoplasm (Fig. 1e; Supplementary Fig. 1i), suggesting that C/EBP $\alpha$  activity is regulated posttranscriptionally during TGF- $\beta$ -mediated EMT.

Similar to HMLE cells, MCF10A cells are human non-transformed epithelial mammary cells responsive to TGF- $\beta$ -stimulation<sup>22,23</sup>. Here, TGF- $\beta$  treatment also increased the expression of mesenchymal markers including *CDH2*, *FNI* and *ZEB1*, and decreased expression of the epithelial marker *CDH1* (Supplementary Fig. 1j). *CEBPA* expression was also found reduced after 16 h of TGF- $\beta$  treatment, and decreased expression was maintained during the 15 days (Supplementary Fig. 1j).



**Fig. 1 C/EBPα is a repressed target by TGF-β.** **a** HMLE cells were left untreated or treated with 2.5 ng/ml of TGF-β for 15 days. Bright-field microscopy images showing phenotypical changes in the presence of TGF-β signaling. Scale bar: 200 μm. Data are representative of three independent experiments. **b** HMLE cells were left untreated or treated with 2.5 ng/ml of TGF-β for 24 h. RNA samples were isolated, and purified RNA was sequenced as described in “Methods”. Gene-ontology (GO) analysis was performed on the significant downregulated genes in HMLE treated cells. Transcription factor activity category and its log2 fold change is represented. **c** The results from qRT-PCR showing CEBPA mRNA expression in HMLE cells treated with 2.5 ng/ml of TGF-β as indicated. Data are represented as mean ± SD of four independent experiments. *p*-values were calculated using unpaired two-tailed Student’s *t* test. \*\*\**p* < 0.001. **d** The results from immunoblotting analysis showing the effect of TGF-β treatment on C/EBPα protein levels and established EMT markers in HMLE cells. Quantification of C/EBPα protein levels normalized to tubulin are shown. Data are representative of three independent experiments. **e** Immunofluorescence staining of C/EBPα and well-characterized EMT markers upon 15 days of TGF-β treatment in HMLE cells. Quantification of C/EBPα nuclear (nucl) and cytoplasmic (cyto) expression levels in untreated and TGF-β-treated cells is shown. Data represented as mean ± SD of five (untreated) and six (TGF-β) independent experiments. *p*-values were calculated using unpaired two-tailed Student’s *t* test. \**p* < 0.05, \*\**p* < 0.01. d days, h hours, FC fold change, β2m beta-2-microglobulin, UNT untreated.

These data demonstrate that *C/EBPα* expression is rapidly and chronically downregulated by TGF- $\beta$  signaling.

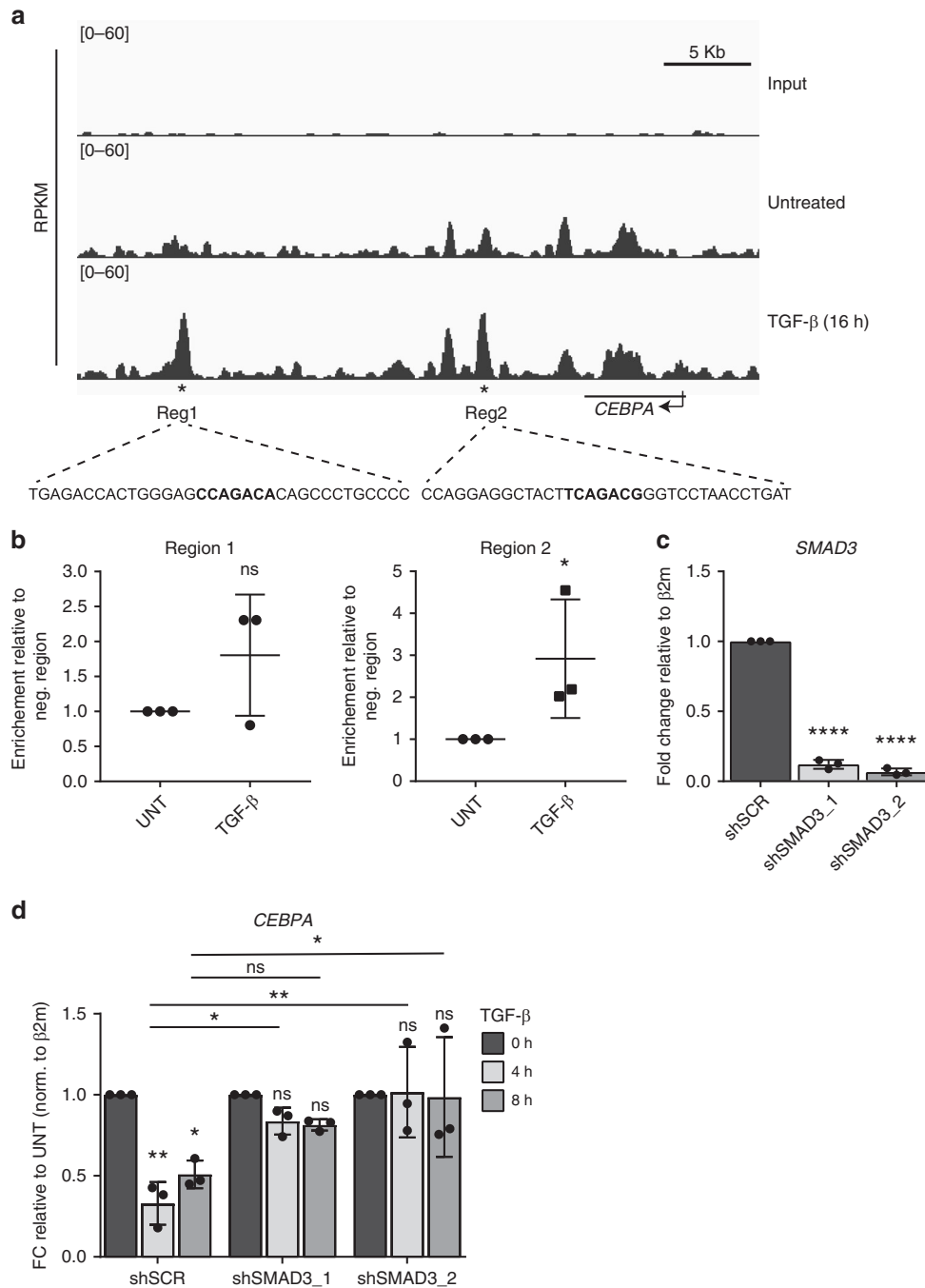
To determine the molecular mechanism underlying TGF- $\beta$ -mediated repression of *C/EBPα*, we first assessed whether the binding of the TGF- $\beta$  pathway effector transcription factor SMAD3 to the *CEBPA* locus was enriched upon TGF- $\beta$  stimulation utilizing chromatin-immunoprecipitation sequencing (ChIP-seq). Unlike SMAD3, SMAD2 does not directly bind to DNA and therefore lacks transcriptional activity<sup>24</sup>, providing a rationale to explore the role of SMAD3 in the repression of *C/EBPα* expression. ChIP-seq profile analysis of *CEBPA* locus revealed that binding of SMAD3 is increased after treatment with TGF- $\beta$  (Fig. 2a). SMAD3 binding to these regions was validated using chromatin immunoprecipitation followed by qRT-PCR (ChIP-qRT-PCR). Although we observed a degree of enrichment in both regions, only SMAD3 binding on region 2 was significantly enriched in the presence of TGF- $\beta$  (Fig. 2b). In order to confirm that TGF- $\beta$ -mediated repression of *C/EBPα* is dependent on SMAD3, we generated HMLE cells expressing two independent shRNAs targeting human *SMAD3* (shSMAD3\_1 and shSMAD3\_2) (Fig. 2c). *C/EBPα* expression was effectively reduced upon TGF- $\beta$  treatment in control HMLE cells (SCR), however in *SMAD3* knockdown cells, TGF- $\beta$ -mediated repression of *CEBPA* was significantly impaired (Fig. 2d). Taken together, these data indicate that TGF- $\beta$ -SMAD3 signaling regulates *C/EBPα* expression in human epithelial mammary cells and suggests that TGF- $\beta$ -mediated repression of *C/EBPα* may be crucial for the progression of EMT.

***C/EBPα* repression is crucial for TGF- $\beta$ -induced EMT.** To further define the functional role of *C/EBPα*, we globally assessed whether constitutive *C/EBPα* expression could interfere with TGF- $\beta$ -regulated gene expression. HMLE cells constitutively expressing *C/EBPα* and corresponding control cells were treated with TGF- $\beta$  for 24 h, and the global transcription response was evaluated by RNA sequencing (RNA-seq). Analysis of RNA-seq data revealed 2588 differentially regulated genes after TGF- $\beta$  treatment, of which 355 were also regulated by *C/EBPα* (Fig. 3a; Supplementary Fig. 2a). GO-term analysis of these 355 genes revealed significant enrichment for biological processes, such as cell adhesion, extracellular matrix, angiogenesis, and cell migration, all indicating a strong association with an EMT program (Fig. 3b). To confirm whether these processes were enriched in EMT, GO-term analysis was also performed using Taube et al.<sup>21</sup> publically available gene expression signature (known as EMT core genes). Significant enrichment of the extracellular matrix, cell adhesion, and angiogenesis was also observed by analyzing these EMT core genes (Supplementary Fig. 2b). Furthermore, to identify common transcriptional targets among the EMT core genes and *C/EBPα*-TGF- $\beta$  core, genes were compared between both sets (Fig. 3c). This revealed a core of 37 genes included in EMT signature that can be modulated by *C/EBPα* expression (Fig. 3c, d). Closer examination of these 37 genes showed two main clusters: genes that are upregulated by TGF- $\beta$  in control cells but repressed in the presence of *C/EBPα* overexpression, and genes which expression is increased in *C/EBPα*-overexpressing cells in both untreated and TGF- $\beta$ -treated conditions. GO-Term enrichment analysis revealed that the first cluster is enriched for processes involved in cell-matrix adhesion and organization whereas second cluster is related to processes involved in inflammatory response (Supplementary Fig. 2c). We focused on genes comprising the first cluster as initial steps of EMT-induced tumor progression requires extensive cell-cell and cell-matrix reorganization<sup>25</sup>. To validate these targets, qRT-PCR was performed using HMLE cells expressing *C/EBPα* or empty vector

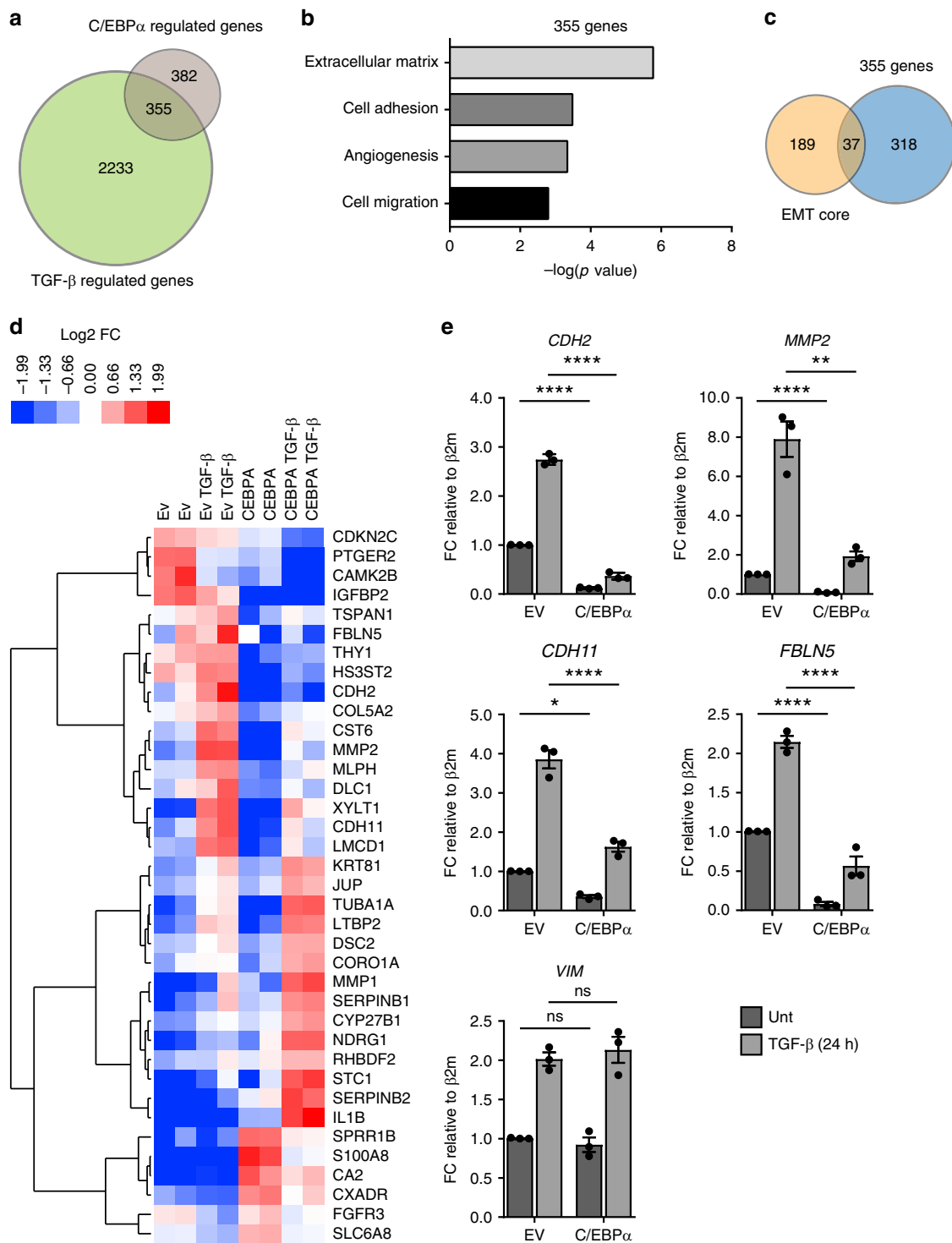
(EV) in the absence or presence of TGF- $\beta$ . For the selected genes, suppression of TGF- $\beta$ -dependent targets was observed in the presence of *C/EBPα* expression (Fig. 3e). In addition, vimentin (*VIM*) was utilized as a negative control since it was unaffected by *C/EBPα* expression (Fig. 3e).

To further evaluate the effect of *C/EBPα* expression during TGF- $\beta$ -mediated EMT, HMLE cells overexpressing *C/EBPα* were treated with TGF- $\beta$  for 15 days, and both mRNA and protein samples were harvested at the indicated time points. In addition, bright-field microscopy pictures were taken at the same time points. Analysis of cell morphology revealed that unlike control cells, HMLE cells overexpressing *C/EBPα* maintained their epithelial phenotype during TGF- $\beta$ -induced EMT (Fig. 4a, arrowheads) and displayed fewer mesenchymal traits. No difference in cell growth was observed upon *C/EBPα* overexpression (Supplementary Fig. 3a). qRT-PCR results demonstrated that *C/EBPα* expression impairs TGF- $\beta$ -mediated induction of mesenchymal markers including *CDH2* (N-cadherin), *MMP2*, or *FBLN5* (Fibulin-5) and repression of epithelial markers including *CDH1* (E-cadherin), *EPCAM*, and *OCN* (occludin) supporting the phenotypical features of these cells (Fig. 4b; Supplementary Fig. 3b). Consistent with the qRT-PCR results, immunoblotting analysis showed that *C/EBPα* expression also reduces the TGF- $\beta$ -dependent increase of mesenchymal markers such as N-cadherin and Fibronectin, and contributes to the maintenance of higher levels of E-cadherin (Fig. 4c). We further confirmed that the effect of *C/EBPα* on TGF- $\beta$ -mediated EMT was not due to a defective TGF- $\beta$  signaling pathway, since SMAD3 is equally phosphorylated upon TGF- $\beta$  treatment, confirming that cells are able to signal properly upon *C/EBPα* overexpression (Supplementary Fig. 3c). Since *C/EBPα* modulates genes associated with cell migration (Fig. 3b), we asked whether *C/EBPα* affects the pro-migratory role of TGF- $\beta$  in HMLE cells. To this end, we performed wound-healing migration assays using HMLE cells ectopically expressing *C/EBPα* or control empty vector in untreated or TGF- $\beta$ -treated conditions. In order to mimic physiological conditions, migratory capacity of HMLE cells were evaluated on three different components of the extracellular matrix, collagen I, collagen IV, and fibronectin. In all ECM coatings, *C/EBPα* overexpression was sufficient to decrease wound closure of TGF- $\beta$ -treated HMLE cells (Fig. 4d; Supplementary Fig. 3d). Taken together, these data strongly suggest that *C/EBPα* contributes to the inhibition of genes involved in epithelial-to-mesenchymal switch, and repression of *C/EBPα* is a requirement for initiating the EMT program.

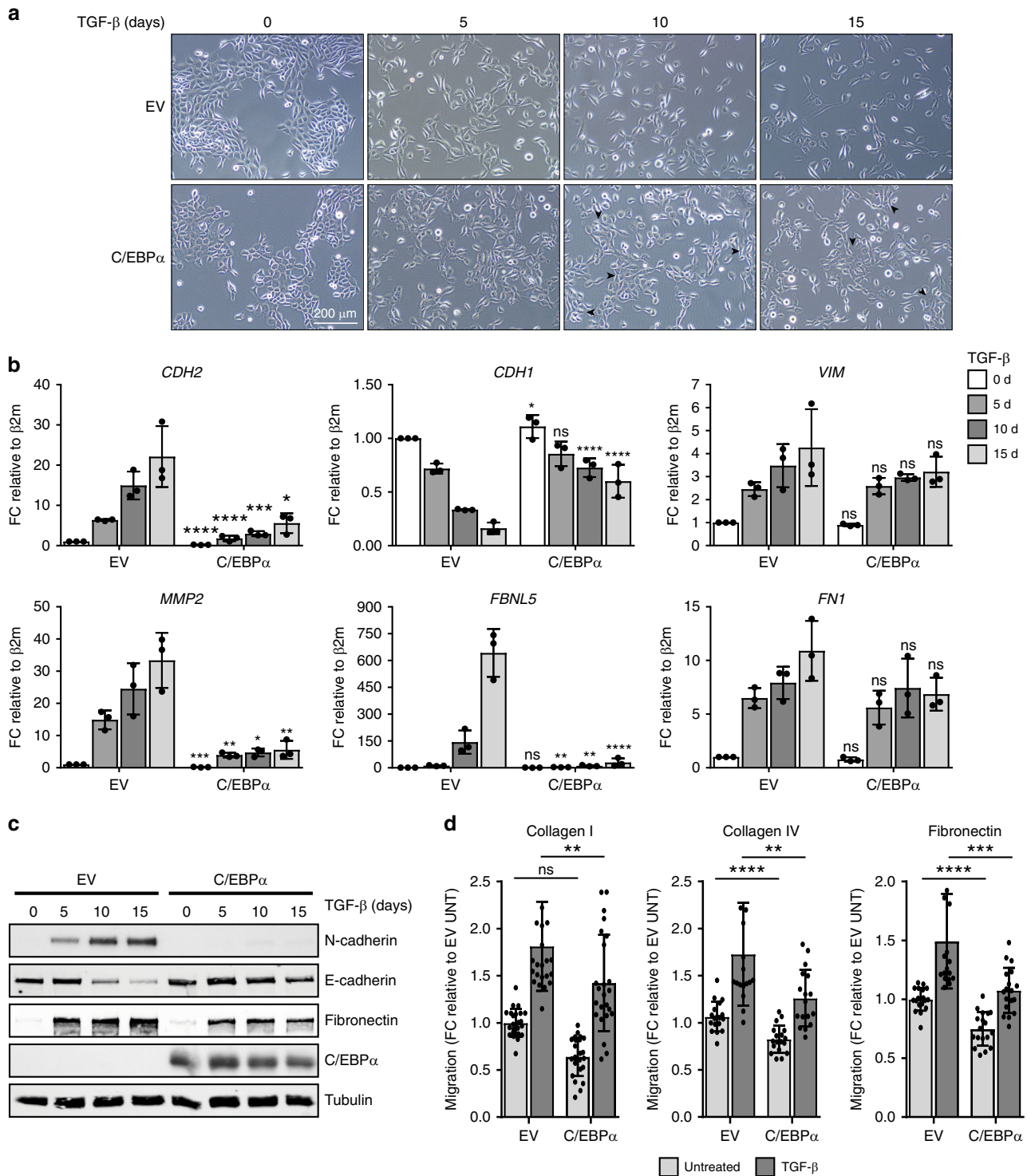
***CEBPA* knockdown impairs epithelial homeostasis.** To evaluate whether *C/EBPα* is a crucial determinant in maintaining epithelial integrity and thereby preventing spontaneous EMT, *CEBPA*-depleted HMLE cells were generated using two independent shRNA constructs. qRT-PCR analysis of *CEBPA* expression showed a strong reduction in *CEBPA* mRNA levels in only one of the shRNAs used (Fig. 5a), therefore HMLE cells expressing shCEBPA\_1 were used for further experiments. Whereas HMLE cells expressing shRNA control (SCR) showed no phenotypical alterations, *CEBPA* knockdown cells displayed elongated shape and reduced cell-cell interactions (Fig. 5b). To assess whether morphological changes induced by *CEBPA* knockdown were associated with EMT, mRNA and protein levels of several well-established EMT markers were analyzed. qRT-PCR results revealed increased expression of mesenchymal markers, including *CDH2* and *FNI*, and decreased expression of epithelial markers in cells expressing shRNA *CEBPA* or transfected with a siRNA pool targeting *CEBPA* (Fig. 5c; Supplementary Fig. 4a). Likewise, immunoblotting analysis showed increased levels of N-cadherin and fibronectin, and reduced



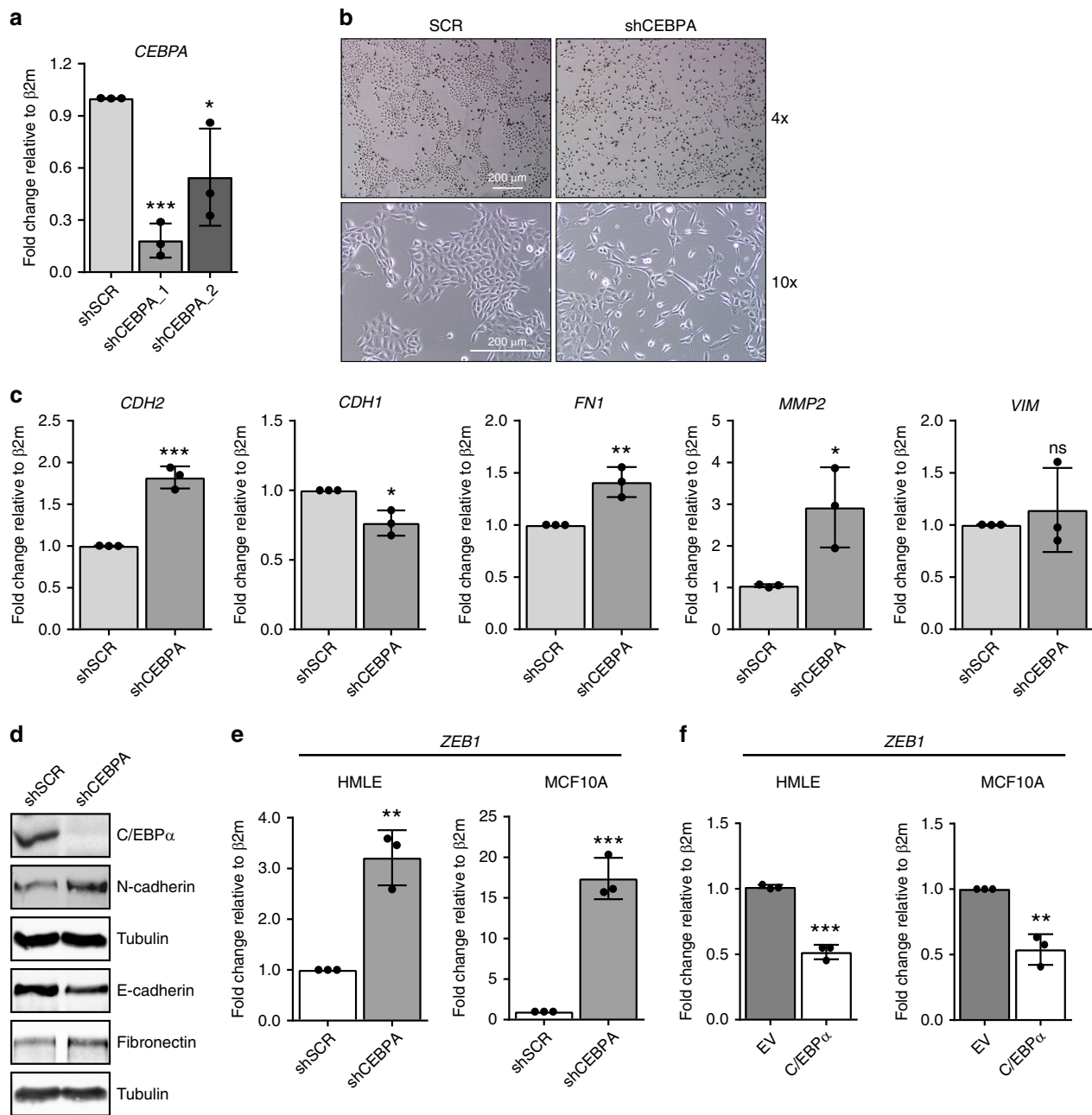
**Fig. 2** TGF- $\beta$ -induced downregulation of C/EBP $\alpha$  is mediated by the transcription factor Smad3. **a** Visualization of SMAD3 CHIP-seq profile within the genomic region surrounding the *CEBPA* locus. Stars indicate enriched picks upon 16 h of TGF- $\beta$  stimulation. Part of the SMAD3-binding regions sequence is displayed, showing in bold SMAD3 known primary binding motif (C/T/C/T/A/C/GACA/G). **b** Chromatin immunoprecipitation was performed in TGF- $\beta$  treated or untreated HMLE cells using Smad3 antibody and according to “Methods”. qRT-PCR was performed using specific primers targeting enriched areas (regions 1 and 2) on *CEBPA* locus. Data represented as mean  $\pm$  SD of three independent experiments, normalized for a negative region. *p*-values were calculated using unpaired one-tailed Student’s *t* test. \**p* < 0.05. **c** The results from qRT-PCR showing *SMAD3* mRNA levels in HMLE cells expressing shRNA control (SCR) or two independent shRNAs targeting human SMAD3 (shSMAD3\_1 and shSMAD3\_2). Data represented as mean  $\pm$  SD of three independent experiments. *p*-values were calculated using unpaired two-tailed Student’s *t* test. \*\*\*\**p* < 0.0001. **d** HMLE cells expressing shSCR or shSMAD3\_2 were left untreated or treated with TGF- $\beta$  for 4 or 8 h. qRT-PCR analysis displaying the effect of TGF- $\beta$  stimulation on *CEBPA* mRNA levels in the presence of SMAD3 knockdown. Data represented as mean  $\pm$  SD of three independent experiments. *p*-values were calculated using two-way ANOVA with Tukey’s multiple comparisons test. neg negated, h hours, FC fold change,  $\beta$ 2m beta-2-microglobulin, UNT untreated.



**Fig. 3 C/EBP $\alpha$  is important for TGF- $\beta$ -induced EMT.** **a** HMLE cells expressing empty vector (EV) or constitutive *CEBPA* were either left untreated or treated with 2.5 ng/ml of TGF- $\beta$  for 24 h. RNA samples were isolated, and purified RNA was sequenced as described in “Methods”. Venn diagram showing the overlap of genes regulated by TGF- $\beta$  with genes regulated by constitutive expression of C/EBP $\alpha$ . **b** Gene-ontology (GO) analysis on the 355 TGF- $\beta$ - and C/EBP $\alpha$ -regulated genes. **c** Venn diagram showing the overlap of 355 TGF- $\beta$ - and C/EBP $\alpha$ -regulated genes with public available EMT core genes. **d** Heatmap displaying gene expression of the 37 genes overlapped between EMT core and TGF- $\beta$ -C/EBP $\alpha$  core. **e** qRT-PCR analysis validating RNA-sequencing data. HMLE cells expressing empty vector (EV) or constitutive *CEBPA* were left untreated or treated with 2.5 ng/ml of TGF- $\beta$  for 24 h. Vimentin (*VIM*) was taken along as negative control. Data represented as mean  $\pm$  SD of three independent experiments.  $p$ -values were calculated using unpaired two-tailed Student’s  $t$  test. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . ns not significant, FC, fold change,  $\beta$ 2m beta-2-microglobulin.



**Fig. 4** Constitutive expression of C/EBP $\alpha$  impairs TGF- $\beta$ -induced EMT. **a** HMLE cells expressing empty vector (EV) or constitutive CEBPA were either left untreated or treated with 2.5 ng/ml of TGF- $\beta$  for 0, 5, 10, or 15 days. Bright-field microscopy images showing cell morphology during treatment. Data are representative of three independent experiments. Scale bar: 200  $\mu$ m. **b** qRT-PCR results displaying the expression levels of well-established EMT markers during TGF- $\beta$  stimulation in the presence of CEBPA constitutive expression. Data represented as mean  $\pm$  SD of three independent experiments. *p*-values were calculated using unpaired two-tailed Student's *t* test. \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. ns not significant. **c** Immunoblotting results showing the protein levels of well-characterized epithelial and mesenchymal markers during TGF- $\beta$  signaling in the presence of CEBPA constitutive expression. Data are representative of three independent experiments. **d** Quantification of wound closure in HMLE cells expressing EV or C/EBP $\alpha$  in untreated and TGF- $\beta$  treatment conditions using 24-wells coated with the indicated substrates. Data represented as mean  $\pm$  SD of *n* = 18–24 independent biological replicates. *p*-values were calculated using unpaired two-tailed Student's *t* test. \**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. ns not significant, d days, FC fold change,  $\beta$ 2m beta-2-microglobulin, UNT untreated.



**Fig. 5** Knockdown of C/EBP $\alpha$  impairs epithelial homeostasis. **a** The results from qRT-PCR analysis showing *CEBPA* mRNA levels in HMLE cells expressing shRNA control (SCR) or HMLE cells expressing two independent shRNAs targeting *CEBPA* (shCEBPA\_1 and shCEBPA\_2). Data represented as mean  $\pm$  SD of three independent experiments. *p*-values were calculated using unpaired two-tailed Student's *t* test. \**p* < 0.05, \*\*\**p* < 0.001. **b** Bright-field microscopy images of HMLE cells expressing shRNA control (SCR) or shRNA targeting *CEBPA* in two different magnifications. Data are representative of three independent experiments. Scale bar: 200  $\mu m$ . **c** qRT-PCR data showing the expression levels of known epithelial and mesenchymal markers upon *CEBPA* knockdown in HMLE cells. Data represented as mean  $\pm$  SD of three independent experiments. *p*-values were calculated using unpaired two-tailed Student's *t* test. \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001. ns not significant. **d** The results from immunoblotting analysis displaying protein levels of well-characterized EMT markers upon *CEBPA* knockdown in HMLE cells. Data are representative of three independent experiments. **e** The results from qRT-PCR analysis showing the mRNA levels of transcription factor *ZEB1* upon *CEBPA* knockdown in HMLE and MCF10A cells. Data represented as mean  $\pm$  SD of three independent experiments. *p*-values were calculated using unpaired two-tailed Student's *t* test. \*\**p* < 0.01, \*\*\**p* < 0.001. **f** *ZEB1* mRNA levels in HMLE or MCF10A cells expressing EV or constitutive *CEBPA*. Data represented as mean  $\pm$  SD of three independent experiments. *p*-values were calculated using unpaired two-tailed Student's *t* test. \*\**p* < 0.01, \*\*\**p* < 0.001.  $\beta 2m$  beta-2-microglobulin.

expression of E-cadherin in the absence of C/EBP $\alpha$  expression (Fig. 5d). Furthermore, we observed similar phenotypical and molecular changes in MCF10A cells upon shRNA-mediated depletion of *CEBPA* (Supplementary Fig. 4b–e). In addition to EMT-associated morphological and molecular changes, we also

observed that depletion of *CEBPA* is sufficient to increase migration of MCF10A cells in a transwell assay system (Supplementary Fig. 4f). This is in accordance with our RNA-seq results that show that genes regulated by C/EBP $\alpha$  were associated with cell migration (Fig. 3b). To determine whether



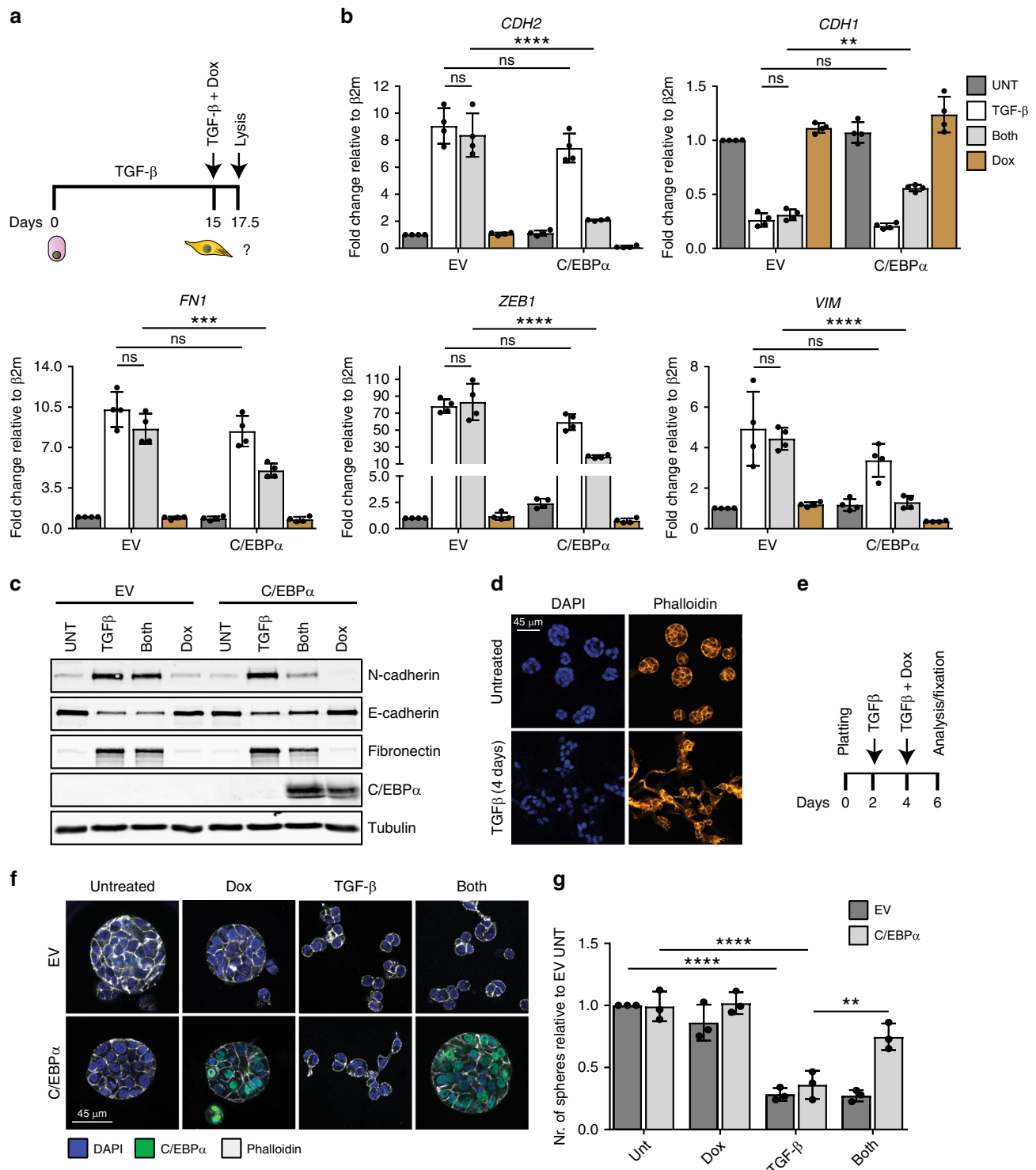
reduced *CEBPA* expression was associated with increased expression of EMT-inducing transcription factors, expression levels of well-characterized EMT master regulators were evaluated. *CEBPA* knockdown resulted in robust induction of *ZEB1* expression in both HMLE and MCF10A cells (Fig. 5e), while no notable changes were observed in *SOX4*, *TWIST1/2*, or *SNAI2* (slug) expression (Supplementary Fig. 5a). Since decreased C/EBP $\alpha$  expression results in increased *ZEB1* levels, we assessed whether ectopic expression of C/EBP $\alpha$  conversely leads to a reduction in *ZEB1* levels. qRT-PCR analysis of control and C/EBP $\alpha$  expressing HMLE or MCF10A cells revealed robust inhibition of *ZEB1* expression in the presence of C/EBP $\alpha$  (Fig. 5f). Bioinformatic analysis using the ContraV2 software revealed numerous highly conserved C/EBP $\alpha$ -binding motifs in the promoter region of *ZEB1* gene (Supplementary Fig. 5b), suggesting a repressor function for C/EBP $\alpha$  in the regulation of *ZEB1*. Taken together, these findings suggest that C/EBP $\alpha$  is indispensable for epithelial cells to retain their epithelial traits, and this may be, at least in part, due to repression of *ZEB1* expression.

**Restoration of C/EBP $\alpha$  levels induces EMT.** Since constitutive C/EBP $\alpha$  expression was sufficient to impair TGF- $\beta$ -induced EMT, we evaluated whether restoration of C/EBP $\alpha$  expression in mesenchymal cells was able to drive mesenchymal-to-epithelial transition (MET). To this end, doxycycline (Dox)-inducible CEBPA-HMLE cells were generated, allowing conditional expression of C/EBP $\alpha$ . In order to induce mesenchymal features, control and CEBPA-inducible HMLE cells were first treated with TGF- $\beta$  for 15 days in the absence of doxycycline. Subsequently, TGF- $\beta$  and doxycycline were added for 60 h after which RNA and protein samples were isolated (Fig. 6a). As expected, doxycycline treatment showed a potent induction of C/EBP $\alpha$  expression in CEBPA-HMLE cells (Supplementary Fig. 6a). In addition, *CDH2*, *VIM*, *FN1*, and *ZEB1* mRNA expression were induced, and *CDH1* expression was repressed upon TGF- $\beta$  treatment (Fig. 6b). The combination of TGF- $\beta$  and doxycycline treatments in the control cells showed no differences in the regulation of gene expression compared with TGF- $\beta$ -treated cells (Fig. 6b). However, TGF- $\beta$ - and Dox-treated CEBPA-HMLE cells displayed strong reduction of the expression of mesenchymal markers and increased expression of the epithelial markers (Fig. 6b, Supplementary Fig. 6a). In accordance with qRT-PCR results, C/EBP $\alpha$  activation resulted in the inhibition of N-cadherin and fibronectin expression, and increased E-cadherin protein levels (Fig. 6c), supporting that conditional expression of C/EBP $\alpha$  in TGF- $\beta$ -induced mesenchymal cells is itself sufficient to rescue their epithelial features. In order to assess whether C/EBP $\alpha$  was able to restore epithelial traits in a more complex epithelial cell culture system, we investigated the role of C/EBP $\alpha$  in MCF10A-epithelial cells cultured as three-dimensional (3D) spheroids. Culture of MCF10A cells on a reconstituted basement membrane, such as matrigel, results in the formation of well-organized epithelial spheroids comprising features found in human mammary gland in vivo<sup>26</sup>. TGF- $\beta$  treatment has been shown to interfere with the architecture of MCF10A-epithelial spheroids, contributing to an active and invasive phenotype<sup>27</sup>. To confirm the effect of TGF- $\beta$  signaling on the MCF10A mammary acini, cells were treated with TGF- $\beta$  2 days after plating. Four days of TGF- $\beta$  stimulation resulted in the reduction of spheroid numbers and impairment of three-dimensional architecture (Fig. 6d). To assess whether expression of C/EBP $\alpha$  was sufficient to recover the formation of MCF10A-epithelial spheroids, we generated GFP-positive MCF10A cells expressing doxycycline-mediated

conditional activation of *CEBPA*. Prior to *CEBPA* induction, cells were treated with TGF- $\beta$  for 48 h, in order to induce mesenchymal features (Fig. 6e). Subsequently, TGF- $\beta$  was refreshed and doxycycline added for 48 h, after which cells were analyzed by fluorescence microscopy (Supplementary Fig. 6b). Spheroids were fixed and stained for C/EBP $\alpha$  and phalloidin (Fig. 6f). As expected, TGF- $\beta$  treatment resulted in the disruption of the spheroids in both control and CEBPA-MCF10A cells. Likewise, the combination of TGF- $\beta$  and doxycycline in empty-vector (EV) MCF10A cells resulted in the abrogation of epithelial spheroids. However, analysis of TGF- $\beta$  and doxycycline treated CEBPA-MCF10A cells revealed the presence of well-structured spheroids, suggesting that activation of C/EBP $\alpha$  expression is sufficient to drive the reorganization of spherical architecture (Fig. 6f, g; Supplementary Fig. 6b). Immunoblotting analysis of control and CEBPA-MCF10A cells cultured in standard two dimensional system revealed that activation of C/EBP $\alpha$  contributes to decreased levels of N-cadherin and fibronectin whereas E-cadherin levels are maintained high (Supplementary Fig. 6c). Taken together, these data strongly suggest that restoration of C/EBP $\alpha$  expression in mesenchymal populations is sufficient to revert these cells toward an epithelial state, and this can be observed even under conditions of 3D epithelial organization.

#### C/EBP $\alpha$ suppresses lung metastasis in a PyMT-mouse model.

To further study the potential for C/EBP $\alpha$  to modulate EMT-induced cancer progression in vivo, we ectopically expressed C/EBP $\alpha$  and empty-vector control in tumor organoids derived from YFP-positive PyMT breast tumors. This MMTV-PyMT; MMTV-Cre;R26R-YFP;E-cad-mCFP model of breast cancer metastasis has been previously shown to spontaneously develop ductal mammary tumors that metastasize primarily to the lungs with resemblances of EMT-mediated tumorigenesis<sup>4,28</sup>. Unlike breast cancer cell lines, tumor organoids display cellular heterogeneity, contributing to a more relevant model for development of breast cancer. Furthermore, while our previous analyses have focused on TGF- $\beta$ -driven EMT, this PyMT model allows us to evaluate the impact of C/EBP $\alpha$  during EMT in an alternative system. As determined by qRT-PCR, *CEBPA* expression was effectively increased in C/EBP $\alpha$ -tumor organoids compared with control organoids (Supplementary Fig. 7a). To test whether ectopic expression of C/EBP $\alpha$  affects primary tumor growth, identical number of organoids expressing empty-vector (EV) control or C/EBP $\alpha$  were injected in five mice per group, and tumor size of both experimental groups was monitored during 5 weeks after injection. Analysis of primary tumor growth showed similar growth rate for both groups (Fig. 7a). Mammary tumors from this model, and similar to human ductal carcinomas, consist of a very small percentage of E-cadherin<sup>low</sup> cells which makes detecting changes in E-cadherin<sup>low</sup> cell number challenging<sup>4</sup>. However, we analyzed the percentage of E-cadherin<sup>low</sup> and E-cadherin<sup>high</sup> cells comprising the primary tumors of mice injected with control tumor organoids or tumor organoids overexpressing C/EBP $\alpha$ . Although not statistically significant, we observed that primary tumors derived from C/EBP $\alpha$ -expressing tumor organoids show a trend toward lower percentage of E-cadherin<sup>low</sup> cells compared with control group (Fig. 7b). Characterization of E-cadherin<sup>low</sup> cells demonstrated that these cells display EMT features, including increased expression of N-cadherin and *ZEB1*, as well as increased motility traits<sup>4</sup>. In order to assess the effect of C/EBP $\alpha$  on formation of metastasis, primary tumors of both groups were resected and mice were maintained for an additional 3 weeks, after which lungs were collected and analyzed for the occurrence of



metastatic lesions. Macrometastasis in the lungs could readily be visualized by eye at the day of harvesting (Fig. 7c, arrowheads). Prior to collection, one EV-control mouse was found dead in the cage without obvious antemortem signs of illness, therefore its lungs were excluded for analysis. Comparison of the number of macrometastasis between the two groups revealed that mice injected with C/EBP $\alpha$ -overexpressing organoids (C/EBP $\alpha$ -mice) displayed significantly reduced numbers of metastatic lesions, suggesting that C/EBP $\alpha$  is sufficient to prevent tumor metastasis in vivo (Fig. 7c). Immunostaining confirmed that primary tumors of C/EBP $\alpha$ -mice expressed higher levels of C/EBP $\alpha$

compared with control mice (Supplementary Fig. 7b). C/EBP $\alpha$ -positive cells could be detected in the surrounding tissue of the primary tumor of mice injected with control organoids, however no overlap with YFP-tumor tissue was observed, suggesting that some stromal cells or infiltrating cells endogenously express C/EBP $\alpha$  (Supplementary Fig. 7b, arrowheads). Immunohistochemistry analysis of various depths of lung tissue confirmed that mice injected with C/EBP $\alpha$ -overexpressing organoids display reduced number of metastatic lesions compared with control mice (Fig. 7d). Since there were still some metastases present in the C/EBP $\alpha$ -mice, it is relevant to evaluate whether

**Fig. 6 Restoration of C/EBP $\alpha$  rescues epithelial phenotype.** **a** Schematic representation of the experimental set up. HMLE cells were treated with 2.5 ng/ml of TGF- $\beta$  for 15 days, and posteriorly C/EBP $\alpha$  expression was induced upon 200 ng/ml of doxycycline for 64 h. **b** The results from qRT-PCR showing the expression levels of known EMT markers in EV or CEBPA-inducible HMLE cells treated with TGF- $\beta$ , doxycycline or both. Data represented as mean  $\pm$  SD of four independent experiments. *p*-values were calculated using two-way ANOVA with Tukey's multiple comparisons test. \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001. ns not significant. **c** Immunoblotting analysis of the expression levels of well-established EMT markers in TGF- $\beta$ -treated HMLE cells in the absence or presence of C/EBP $\alpha$  overexpression. Data are representative of three independent experiments. **d** Confocal microscopy visualization of MCF10A-epithelial spheroids treated with 5 ng/ml of TGF- $\beta$  for 4 days. Data are representative of three independent experiments. Scale bar: 45  $\mu$ m. **e** Schematic representation of the 3D spheroid-formation assay. MCF10A cells expressing empty vector (EV) or doxycycline-inducible C/EBP $\alpha$  were stimulated with TGF- $\beta$  (5 ng/ml), doxycycline (200 ng/ml), or both as indicated. **f** Confocal microscopy visualization of epithelial spheroids using MCF10A cells expressing empty vector or doxycycline-inducible C/EBP $\alpha$  in the presence of TGF- $\beta$ , doxycycline (DOX), or both. Cells were stained for C/EBP $\alpha$  (green) and phalloidin (white). DAPI was used to visualize the nucleus (blue). Data are representative of three independent experiments. Scale bar: 45  $\mu$ m. **g** Quantification of the number of spheres formed in MCF10A cells expressing empty vector or doxycycline-induced C/EBP $\alpha$  upon the treatment of TGF- $\beta$ , doxycycline, or both. The number of spheres are relative to empty vector untreated. Data are represented as mean  $\pm$  SD of three independent experiments. *p*-values were calculated using two-way ANOVA with Tukey's multiple comparisons test. \*\**p* < 0.01, \*\*\*\**p* < 0.0001.  $\beta$ 2m beta-2-microglobulin, EV empty vector, UNT untreated.

these lesions were positive for C/EBP $\alpha$  expression. Nuclear C/EBP $\alpha$  staining was absent in the YFP-positive metastatic lesions of EV-mice and greatly reduced in the metastasis of mice injected with C/EBP $\alpha$ -overexpressing organoids compared with its primary tumor tissue (Fig. 7e). This suggests that cells with the ability to metastasize are likely C/EBP $\alpha$ -negative/low. In order to confirm that reduction of metastatic potential observed in C/EBP $\alpha$ -expressing tumors was not related to the variability in tumor size, primary tumor of both groups was measured on the day of harvesting, and number of metastatic lesions were assessed (Supplementary Fig. 7c). Similarly, mice injected with C/EBP $\alpha$ -expressing organoids displayed a reduction in the number of metastatic lesions compared with control mice, supporting that C/EBP $\alpha$  expression impairs metastasis (Supplementary Fig. 7d). Taken together, these data suggest that in vivo, C/EBP $\alpha$  can function as a tumor suppressor in breast cancer, impairing metastasis, and thereby breast cancer progression.

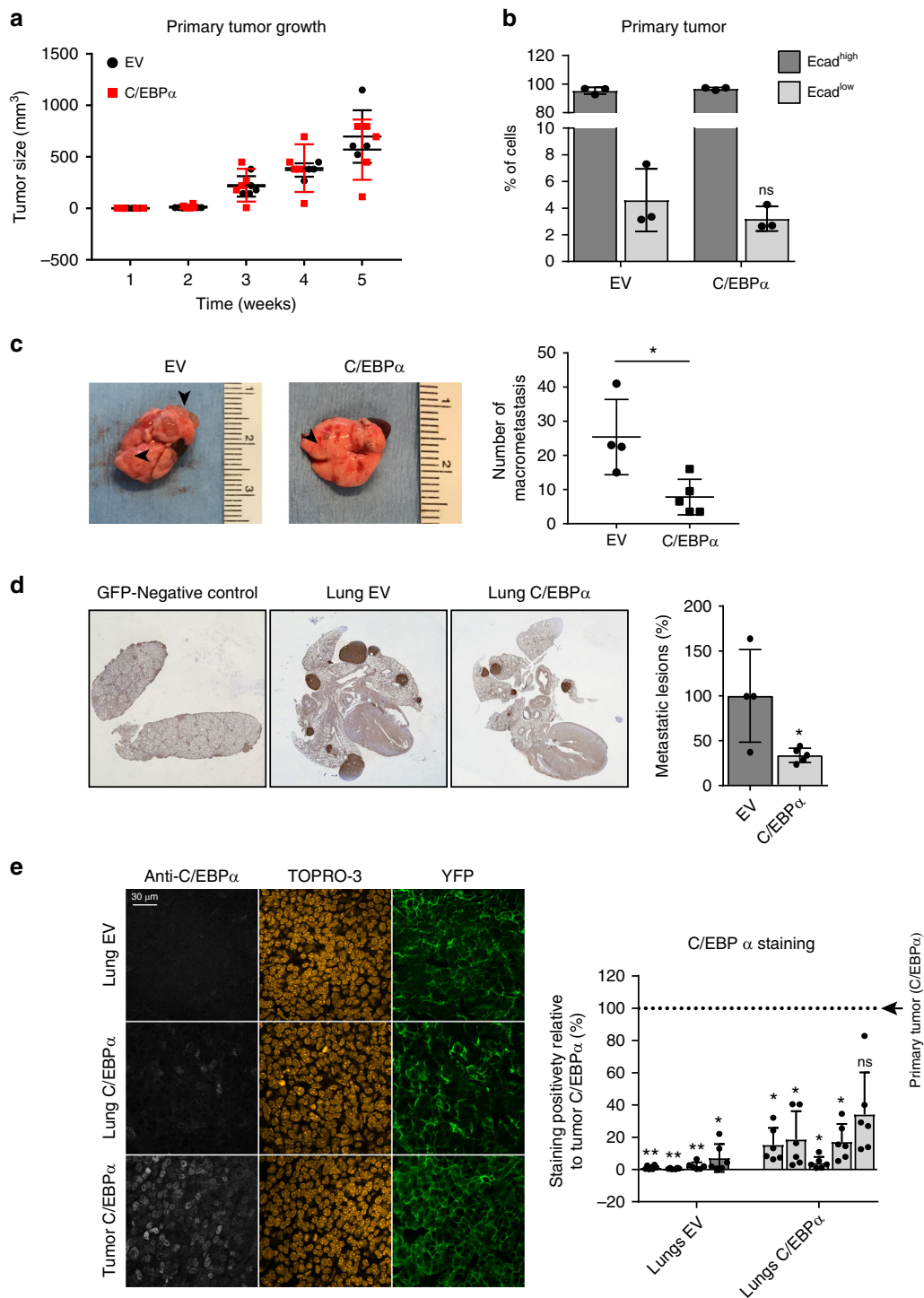
## Discussion

EMT-inducing signals, such as the TGF- $\beta$  pathway, contribute to the loss of epithelial features and acquisition of mesenchymal traits, and although it has been a subject of debate, several studies have reported that such cellular reprogramming impacts cell motility, extracellular matrix remodeling, cell extravasation, and chemotherapy resistance<sup>4,5,29,30</sup>. Whereas several studies have focused on understanding the role of the transcription factors induced during TGF- $\beta$ -mediated EMT, knowledge regarding the regulation of epithelial transcription factors during this process is limited. Here, we demonstrate that repression of the transcription factor C/EBP $\alpha$  is essential for TGF- $\beta$ -mediated EMT in human mammary epithelial cells. C/EBP transcription factors play a pivotal role during terminal differentiation of a variety of cells types, including mammary epithelial cells<sup>12</sup>. Decreased levels of C/EBP $\alpha$  have been found in several types of solid tumors, including breast cancer, however the mechanisms underlying this downregulation remain in general unknown<sup>17</sup>. By performing unbiased global RNA-sequencing analysis on TGF- $\beta$ -treated HMLE cells, we identified C/EBP $\alpha$  expression as being one of the most repressed transcription factors, whereas no alterations in *CEBPB* expression were observed. These results contrast with previous observations that showed reduced *CEBPB* expression upon TGF- $\beta$ -mediated EMT in murine NMuMG cells<sup>31</sup>. These differences can be potentially explained due to species differences and/or context-dependent upstream signaling regulating C/EBPs expression<sup>31</sup>.

TGF- $\beta$ -mediated repression of C/EBP $\alpha$  expression was maintained during EMT. Since *CEBPA* mRNA levels were rapidly inhibited upon TGF- $\beta$  treatment, we hypothesized that SMAD3 may be responsible for repression of *CEBPA* transcription.

Although SMADs are generally transcriptional activators, there are examples of genes that are transcriptionally repressed by this transcription factor family<sup>32–34</sup>. Similarly, we observed that occupancy of SMAD3 on *CEBPA* locus is enriched upon TGF- $\beta$  treatment, and depletion of SMAD3 impaired TGF- $\beta$ -mediated repression of *CEBPA*, supporting a role for SMAD3 as a transcriptional repressor of C/EBP $\alpha$  expression during TGF- $\beta$ -induced EMT. Differences in the kinetics of C/EBP $\alpha$  mRNA and protein expression upon TGF- $\beta$  treatment suggest that C/EBP $\alpha$  activity might also be regulated independently of protein degradation. We observed that short exposure to TGF- $\beta$  induces cytoplasmic shuttling of C/EBP $\alpha$  protein in HMLE cells, suggesting that C/EBP $\alpha$  is posttranscriptionally regulated during TGF- $\beta$ -mediated EMT, and together with its transcriptional repression enables a rapid induction of the EMT program.

Rapid reduction of C/EBP $\alpha$  expression during EMT appears to be critical to allow the progression of this program. Constitutive expression of C/EBP $\alpha$  during the initial 24 h of TGF- $\beta$  stimulation was sufficient to repress genes important for the mesenchymal transition, including *CDH2*, *MMP2*, *CDH11*, and *FBLN5*. Proteins encoded by these genes regulate cell motility, extracellular matrix degradation, and transendothelial migration<sup>35–38</sup>. Indeed, C/EBP $\alpha$  overexpression was sufficient to impair TGF- $\beta$ -induced cell migration. In addition, constitutive expression of C/EBP $\alpha$  during TGF- $\beta$ -driven EMT impaired this program, as showed by the stable inhibition of crucial mesenchymal markers, including N-cadherin and fibronectin, and maintenance of high levels of epithelial marker E-cadherin. Induction of N-cadherin expression plays a pivotal role in EMT, and has been shown to be sufficient to promote cell migration and invasion in several cancer cell lines, regardless of E-cadherin expression<sup>39,40</sup>. In addition, mice expressing mutant N-cadherin die on day 10 after gestation due to impairment in heart formation, in which EMT is a critical process<sup>41</sup>. Therefore, repression of N-cadherin is essential to impair EMT and cancer progression. Furthermore, higher levels of E-cadherin were observed in TGF- $\beta$ -treated C/EBP $\alpha$ -HMLE cells compared with the control cells. In addition to its role in preserving cell–cell/cell–basal membrane adhesion and cell polarity<sup>41</sup>, studies showed that transmembrane E-cadherin-binding  $\beta$ -catenin is important to prevent nuclear translocation of  $\beta$ -catenin, thus preventing regulation of genes involved in promoting cell survival, proliferation, migration, and angiogenesis<sup>41–44</sup>. Taken together, by contributing to the repression of N-cadherin and maintenance of E-cadherin expression, C/EBP $\alpha$  is a critical negative regulator of TGF- $\beta$ -induced EMT. Overexpression of C/EBP $\alpha$  in lung adenocarcinoma cells or enhanced C/EBP $\alpha$  expression by short-activating RNAs (saRNA) in hepatocellular carcinoma was observed to induce similar phenotypes<sup>45,46</sup>. This suggests that C/EBP $\alpha$  plays an



important and conserved role in preventing unwarant EMT in multiple tissues.

Since *C/EBPα* depletion evoked features reminiscent of spontaneous EMT, we hypothesized that *C/EBPα* may act as a transcriptional repressor of a master regulator of EMT. The expression of EMT-inducers, including Snail, Twist, or Zeb1 transcription factors, can differentially contribute to the EMT program depending on the cell, tissue type and the signaling pathway triggering EMT<sup>29</sup>. Depletion of *CEBPA* in both HMLE and MCF10A cells resulted in strong upregulation of *ZEB1*, whereas no significant changes in other EMT factors were observed. The Zeb1 transcription factor acts as both transcriptional activator and

repressor by binding regulatory sequences at E-boxes. In several cellular contexts, TGF- $\beta$  has been shown to activate Zeb1 resulting in the repression of E-cadherin and induction of N-cadherin and MMPs expression, whereas no major alterations in vimentin expression was observed<sup>9</sup>. In accordance with these observations, constitutive expression of *C/EBPα* during TGF- $\beta$ -mediated EMT contributed to repression of ZEB1 (Supplementary Fig. 3b), N-cadherin and MMP2 as well as maintenance of E-cadherin levels, however no significant repression of vimentin was detected (Fig. 4b). These data, together with analysis of *ZEB1* regulatory region showing highly conserved *C/EBPα*-binding motifs, strongly suggests that *C/EBPα* contributes to epithelial homeostasis and

**Fig. 7 Ectopic C/EBP $\alpha$  expression in breast tumor organoids inhibits lung metastasis.** Tumor organoids derived from *MMTV-PyMT; MMTV-Cre; R26R-YFP; E-cad-mCFP* mice were transduced with EV control or C/EBP $\alpha$ -overexpressing construct and injected in NGS mice. **a** Tumor volume of primary mammary tumor.  $n = 5$  per experimental group.  $p$ -values were calculated using two-way ANOVA with Tukey's multiple comparisons test (not significant). **b** Primary tumors of mice injected with control or C/EBP $\alpha$ -overexpressing tumor organoids ( $n = 3$  per group) were dissociated to single cells, stained for extracellular E-cadherin, and analyzed for YFP, CFP, and E-cadherin expression by flow cytometry analyses. YFP-positive and DAPI-negative cells were further analyzed for low and high E-cadherin expression using both CFP and extracellular E-cadherin expression. Statistical significance was calculated using unpaired two-tailed Student's  $t$  test (ns not significant). **c** Representative lungs obtained from mice injected with EV or C/EBP $\alpha$ -overexpressing organoids ( $n = 4$ –5 per group). Arrowheads indicate an example of macrometastasis observed in the lungs. Quantification of the number of macrometastasis visible on the surface of the lungs from control and C/EBP $\alpha$ -mice is displayed. Data represented as mean  $\pm$  SD.  $p$ -values were calculated using unpaired two-tailed Student's  $t$  test.  $*p < 0.05$ . **d** GFP staining of the metastatic lungs and quantification of the number of metastatic lesions present in various depths of the lung tissue. Data represented as mean  $\pm$  SD of 4–5 mice per group.  $p$ -values were calculated using unpaired two-tailed Student's  $t$  test.  $*p < 0.05$ . **e** Immunofluorescence staining and quantification of C/EBP $\alpha$  expression in sections of the metastatic lungs of both experimental groups compared with primary tumor tissue derived from mice injected with C/EBP $\alpha$ -overexpressing organoids (4–5 mice per group). Scale bar: 200  $\mu$ m. Each column represents the average value for each individual mouse relative to the positive control (primary tumor tissue developed from injected C/EBP $\alpha$ -overexpressing organoids). Six different sections were quantified per mouse. Data represented as mean  $\pm$  SD.  $p$ -values were calculated using unpaired two-tailed Student's  $t$  test.  $*p < 0.05$ ,  $**p < 0.01$ . ns not significant.

impairment of TGF- $\beta$ -induced EMT through repression of *ZEB1*. Future studies will reveal whether C/EBP $\alpha$  is a direct repressor of *ZEB1*, or if such modulation requires intermediate steps of regulation.

The occurrence of metastasis is unavoidable for a great majority of the cancer patients, and despite the significant advances in treatment, metastases are still the major cause of cancer-related deaths. Eradication of metastasis often fails due to chemotherapy resistance. Recent studies have shown that cells that have undergone EMT and display mesenchymal traits are more insensitive to chemotherapy treatment<sup>3,47</sup>. Therefore, identification of novel molecular mechanisms underlying mesenchymal-to-epithelial switch (also known as MET) may contribute to the development of future cancer therapeutics and increase success rate of cancer treatment. We demonstrated that conditional activation of C/EBP $\alpha$  in TGF- $\beta$ -induced mesenchymal cells is sufficient to strongly reduce expression of mesenchymal markers and partially restored expression levels of E-cadherin. Likewise, in a more complex biological setting, conditional activation of C/EBP $\alpha$  was sufficient to induce reorganization of epithelial spheroid architecture. These data suggest that restoration of C/EBP $\alpha$  expression in mesenchymal cells is sufficient to promote MET, thereby potentially restoring sensitivity toward chemotherapy. Generation of small molecules able to reach metastatic regions and activate *CEBPA* expression would therefore be strong therapeutic candidates.

We have demonstrated that C/EBP $\alpha$  is a direct transcriptional target of SMAD3, and its repression by TGF- $\beta$  signaling is pivotal for the progression of TGF- $\beta$ -induced EMT. Therefore, C/EBP $\alpha$  expression levels represent an important determinant of breast cancer progression. Moreover, by using an established mice model of EMT-induced breast cancer, we validated the anti-metastatic role of C/EBP $\alpha$  in vivo. Although the vast majority of GFP-positive cells colonizing the lungs of C/EBP $\alpha$ -mice were negative for C/EBP $\alpha$  expression, some cells showed positive nuclear staining at metastatic regions. Nonetheless, we cannot exclude that C/EBP $\alpha$  is transcriptionally inactive in these cells. As discussed previously, regulation of C/EBP $\alpha$  activity by post-transcriptional modifications has been shown to impact C/EBP $\alpha$  transcriptional output in liver cancer<sup>16,48</sup>. Importantly, our observations are supported by clinical data of breast cancer patients where *CEBPA* expression is found higher in primary normal tissue compared with primary tumor and markedly lower in primary tumors with metastasis compared with primary tumor without metastasis (Supplementary Fig. 8a, b). These data support the notion that C/EBP $\alpha$  is an inhibitory factor for tumor metastasis formation in a clinical setting. Moreover, we have also evaluated C/EBP $\alpha$  expression levels in various breast cancer cell

lines and this also showed lower C/EBP $\alpha$  levels in metastatic breast tumor cell lines, such as MDA-MB-231 and MDA-MB-436, in comparison with untransformed HMLE cells (Supplementary Fig. 8c, d).

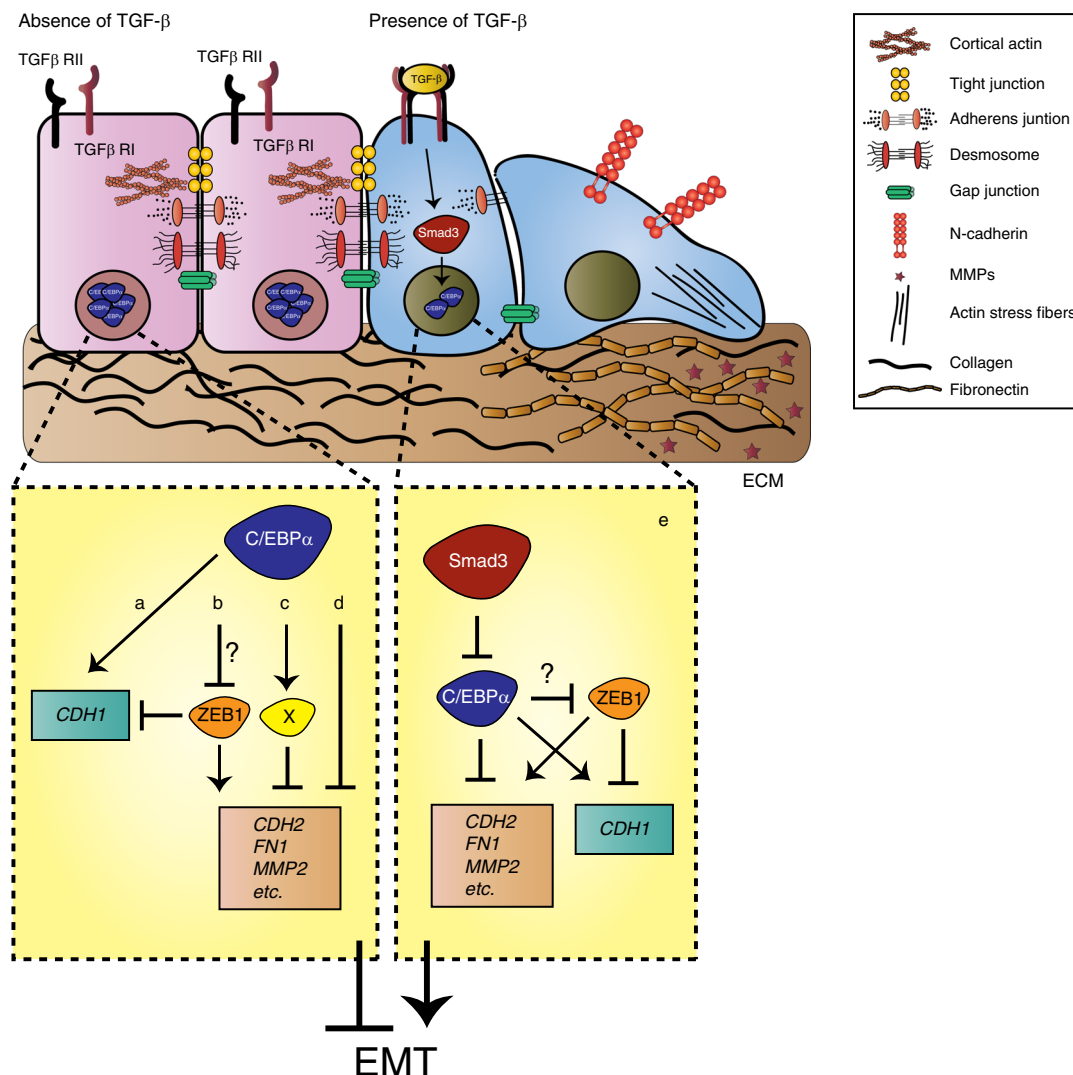
Based on our data, we propose a model that TGF- $\beta$ -activated SMAD3 in epithelial cells promotes loss of C/EBP $\alpha$  expression resulting in the transdifferentiation into mesenchymal cells (Fig. 8). Maintenance of epithelial traits by C/EBP $\alpha$  can be envisioned by its direct involvement in the transcriptional activation of *CDH1* (E-cadherin) and repression of *ZEB1* (Supplementary Fig. 8e; Fig. 8a, b) and/or by the activation of an intermediate player responsible for transcriptionally repressing pro-oncogenic genes, such as *CDH2* and *MMP2* (Fig. 8c). In addition, we cannot exclude that C/EBP $\alpha$  itself cannot directly bind to the regulatory regions of these genes promoting its repression, as C/EBP transcription factors are both activators and repressors<sup>49</sup> (Fig. 8d). Upon TGF- $\beta$  pathway activation, nuclear SMAD3 can bind and repress *CEBPA*. Loss of C/EBP $\alpha$  expression permits the activation of EMT program (Fig. 8e).

Taken together, our findings demonstrate that C/EBP $\alpha$  plays a pivotal role in the maintenance of epithelial homeostasis of human mammary cells and restoration of C/EBP $\alpha$  expression impairs EMT and the development of metastasis, supporting the view that cells require plasticity between epithelial and mesenchymal state in order to metastasize<sup>44</sup>.

## Methods

**Cell culture.** Non-transformed immortalized human mammary epithelial cells expressing hTERT and SV40 large T and small t antigens (classified as HMLE cells and kindly provided by Prof. Dr. Robert Weinberg) were cultured in the MEGM medium (cc-3150, Lonza): DMEM/F12 media (Invitrogen) (1:1) supplemented with insulin (10  $\mu$ g/ml, Sigma), EGF (20 ng/ml, Peprotech), hydrocortisone (0.5 mg/ml, Sigma), and 1% penicillin-streptomycin (Invitrogen). Mesenchymal-like phenotype cell cultures were obtained by supplementing the normal culture medium with 2.5 ng/ml of TGF- $\beta$ 1 (R&D Systems, 240-B-010). MCF10A cells (kindly provided by Dr. Patrick Derksen) were cultured in the DMEM/F12 media (Invitrogen) supplemented with horse serum (5% final, Invitrogen) insulin (10  $\mu$ g/ml, Sigma), EGF (20 ng/ml, Peprotech), hydrocortisone (0.5  $\mu$ g/ml, Sigma), 1% penicillin-streptomycin (Invitrogen), and cholera toxin (100 ng/ml, Sigma).

**Generation of C/EBP $\alpha$  cell lines.** To generate HMLE cells expressing constitutive C/EBP $\alpha$  and respective control, HMLE cells were transduced with pLZRS eGFP retroviral construct containing the coding sequence of human *CEBPA* or empty vector, respectively. Retrovirus particles were produced by transfection of the retroviral packing cell line Phoenix-ampho with 10  $\mu$ g of DNA using poly-ethylenimine (PEI; Polysciences). Twenty-four hours after transfection, medium was replaced by MEGM:DMEM/F12 media. Viral supernatants were collected 24 h after, and filtered through a 0.22- $\mu$ m filter and added to the HMLE cells. Cells were expanded and sorted according to their GFP positivity using flow cytometer (BD Biosciences). Cells were expanded for 2 weeks after sorting before experiments. The expression of C/EBP $\alpha$  p42 isoform was confirmed by immunoblotting analysis.



**Fig. 8 Schematic model for C/EBP $\alpha$  function during epithelial homeostasis and TGF- $\beta$ -mediated EMT.** In non-transformed mammary epithelial cells, epithelial homeostasis is ensured by high levels of C/EBP $\alpha$ , which negatively regulates the expression of pro-mesenchymal proteins. C/EBP $\alpha$ -mediated repression of unwarranted EMT might be due to (a) direct transcription of epithelial proteins such as E-cadherin and (b) suppression of the EMT-inducer ZEB1. ZEB1 itself is able to directly repress CDH1 (E-cadherin) expression and induce the transcription of several mesenchymal genes, including CDH2 (N-cadherin) and MMP2. c Also, C/EBP $\alpha$ -mediated activation of an intermediate transcriptional repressor may be responsible for impairing EMT and/or (d) direct transcriptional repression by C/EBP $\alpha$ . Upon deregulation of TGF- $\beta$  signaling, induction of EMT is induced by SMAD3-mediated inhibition of C/EBP $\alpha$  and (e) repression of EMT is alleviated.

To generate CEBPA-depleted HMLE cells and respective control, a pLKO.1-puro lentiviral construct expressing shRNA targeting *CEBPA* (shCEBPA\_1: Sigma-Aldrich, TRCN0000007304 and shCEBPA\_2: TRCN0000007306) or expressing shRNA control (Sigma-Aldrich, SHC002) were used, respectively. Lentivirus particles were produced by transfection of HEK293T cells with 3.25  $\mu$ g of Pax2, 1.8  $\mu$ g of pMD2.G, and 5  $\mu$ g of shRNA constructs. Twenty-four hours after transfection, medium was replaced by MEGM:DMEM/F12 media. Viral supernatants were collected 24 h after, and filtered through a 0.22- $\mu$ m filter and added to the HMLE cells. Cells were selected using 1  $\mu$ g/ml of puromycin (invivoGen). Cells were maintained in 1  $\mu$ g/ml of puromycin through the cell culture, except during experimental conditions where 0.3  $\mu$ g/ml of puromycin was used.

To generate conditionally regulated C/EBP $\alpha$  cells, HMLE and MCF10A cells were transduced with a pINDUCER21 eGFP lentiviral construct containing the coding sequence of human *CEBPA* or empty vector as a control. Lentivirus particles were produced by transfection of HEK293T cells with 3.25  $\mu$ g of Pax2, 1.8  $\mu$ g of pMD2.G, and 5  $\mu$ g of shRNA constructs. Twenty-four hours after transfection, medium was replaced by MEGM:DMEM/F12 media or DMEM/F12. Viral supernatants were collected 24 h after, and filtered through a 0.22- $\mu$ m filter and added to the HMLE cells or MCF10A cells. Cells were expanded and sorted according to their GFP positivity using flow cytometer (BD biosciences).

Knockdown of *CEBPA* in HMLE cells using siRNAs was performed using 25 nM human *CEBPA* siRNA (Thermo Scientific, ON-TARGET plus SMARTpool) with Lipofectamine RNAiMAX reagents (Invitrogen), 72 h before harvesting samples.

All experiments in this paper have been carried out with pooled multiclonal populations.

**RNA sequencing.** HMLE cells were treated with TGF- $\beta$  for 24 h after which RNA was isolated using the RNeasy Isolation kit (Qiagen) and according to the manufacturer's protocol. Purified RNA was subsequently repurified using mRNA-ONLY Eukaryotic mRNA Isolation Kit (Epicentre (Illumina, Inc.), Madison, WI, USA). Next, sequencing libraries were constructed using SOLiD Total RNA-Seq Kit (Applied Biosystems Life Technologies) according to the standard protocol recommendations for low input and sequenced on SOLiD Wildfire sequencer (Applied Biosystems Life Technologies) in a multiplexed way to produce 50-bp-long reads. Sequencing reads were mapped against the reference genome (hg19 assembly) using the BWA package. Only uniquely placed reads were used for further analysis. Cisgenome v2.0 software package was used to calculate reads per 1000 base pairs of transcript per million reads sequenced (RPKM) values for all RefSeq annotated genes. RPKMs were quantile normalized throughout all samples and presented as  $\log_2(\text{RPKM})$  after adding small number to RPKM (0.1) to avoid

**Table 1** List of primers used for qRT-PCR.

Target	Sequence	Target	Sequence
hFN1 Fw	TGGCACCCACGCTCAGATACA	hB2m Fw	ATGAGTATGCCTGCCGTGTGA
hFN1 Rv	CTCGCCAGGCAGGTTGACGG	hB2m Rv	GGCATCTTCAAACCTCCATG
hVIM Fw	ACCAACGACAAAGCCCGCGT	hCDH1 Fw	CACCACGTACAAGGGTCAGGTGC
hVIM Rv	CAGAGACGCATTGTCAACATCCTGT	hCDH1 Rv	CAGCCTCCACGCTGGGGTAT
hTJP1 Fw	TGCGCTTACCACACTGTGATCCT	hSOX4 Fw	GTCCGCGCCTTGTACAGCGA
hTJP1 Rv	CCGACCATGGTTCAGGGGCA	hSOX4 Rv	GGCCTCGAGCTGGGAATCGC
hFBLN5 Fw	AATAAAACACCCGCGAGCCC	hSNAI1 Fw	TCCGGACCCACACTGGCGAGAA
hFBLN5 Rv	CACTGTGCCTGTGCATTCCC	hSNAI1 Rv	CCTGAGCAGCCGACTCTTGGT
hCDH11 Fw	TCAAGGGCCCCAGAAATCAC	hSNAI2 Fw	CTGGGCGCCCTGAAGATGCAT
hCDH11 Rv	TTGAGCTCATCACGTCAGGG	hSNAI2 Rv	GGCTTCTCCCCGTGTGAGTTCTA
hCDH2 Fw	AGTACCCTGGTCAAACCAATCGA	hTWIST1 Fw	CGCTCGCCGCTCGAGAGATG
hCDH2 Rv	TGCAGTTGACTGAGCGGGTG	hTWIST1 Rv	TCGTGTTGCTCAGGCTGTCTG
hMMP2 Fw	GGAGGCGCTAATGGCCC	hZEB1 Fw	AGCGCTTCTCACACTCTGGGTCTT
hMMP2 Rv	GGTATTGCACTGCCAATCTTTGT	hZEB1 Rv	TGGCACCCACGTGCTCATTCCG
hCEBPB Fw	CGACGAGTACAAGATCCGGC	hTWIST2 Fw	CGCCAGGGGTGTCCGTC
hCEBPB Rv	TGCTTGAACAAGTTCCCGCAG	hTWIST2 Rv	CGGGTCTTCTGTCCGATGTC
hCEBPA 1 Fw	CGCGCGCGCGGACTTT	hEPCAM Fw	CCATGTGCTGGTGTGTGAA
hCEBPA 1 Rv	GGTAGCCGCGCGCCGCGCA	hEPCAM Rv	TGTGTTTTAGTTCAATGAT
hCEBPA 2 Fw	TTGTATCTGGCCTCTGTGCC	hOCLN Fw	CCGAGTTTCAGGTGAATTGG
hCEBPA 2 Rv	GCCGACGGAGAGTCTCATT	hOCLN Rv	GGAGTGTAGTGTGGTGT

log<sub>2</sub> transformation of zero values. A fold change of value  $\geq 1.5$  ( $\log_2FC \geq 0.58$ ) and  $\leq 0.72$  ( $\log_2FC \leq -0.48$ ) were used as thresholds for differential expression analysis of the RNA-Seq data.

**Quantification of RNA expression (qRT-PCR).** mRNA was extracted from cells using the RNeasy Isolation Kit (Qiagen). According to the manufacturer's protocol for single-stranded cDNA synthesis, 500 ng of the total RNA was reverse transcribed using iScript cDNA synthesis kit (BIO-Rad). cDNA samples were amplified using SYBR green supermix (BIO-Rad), in a white 96-multiwell plate by Light-Cycler 96 system instrument (Roche) according to the manufacturer's protocol. To quantify the data, the comparative Ct method was used. Relative quantity was defined as  $2^{-\Delta\Delta Ct}$  and  $\beta$ -2-microglobulin was used as a reference gene. qRT-PCR primers are available below (Table 1):

**Immunoblotting analysis.** Cells were washed with PBS and lysed in Laemmli buffer (0.12 mol/L Tris-HCL (pH 6.8), 4% SDS, and 20% glycerol). Protein concentration was determined using Lowry protein assay. Equal amounts of sample (30–40  $\mu$ g) were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and electrophoretically transferred to the polyvinylidene difluoride membrane (Milipore). The membranes were blocked with 5% milk protein in TBST (0.3% Tween, 10 mM Tris pH 8.0, and 150 mM NaCl in H<sub>2</sub>O) and probed with antibodies overnight at 4°. Immunocomplexes were detected using Odyssey imaging system (LI-COR Biosciences) or ECL (GE Healthcare), which were consequently exposed to Kodak XB films (Rochester). The following primary antibodies were used: anti-C/EBP $\alpha$  (Santa Cruz Biotechnology, sc-61), anti-N-cadherin (BD Biosciences, 610921), anti-E-cadherin (BD Biosciences, 610182), anti-Fibronectin (BD Biosciences, 610077), anti- $\alpha$ -tubulin (Sigma, T90026).

**Wound-healing migration assays.** Twenty-four-well plates were coated with collagen I, collagen IV, or fibronectin (all 50  $\mu$ g/ml) for 2 h at 37 °C. In total, 200,000 control or C/EBP $\alpha$ -overexpressing HMLE cells were plated per well on the coated plates. When a confluent monolayer was achieved, migration was assessed by making a single straight scratch in the monolayer using a 20–200  $\mu$ l pipette tip at the end of the day. Wells were washed twice with PBS to remove loose cells, and 0.5 ml of medium containing TGF- $\beta$  (2.5 ng/ml) was added, followed by a 20 h incubation at 37 °C, 5% CO<sub>2</sub>, and ambient oxygen level (20%). Images were recorded at  $t = 0$  and  $t = 20$  h at similar locations, after which the scratch surface was determined using ImageJ software to determine wound closure.

**Transwell migration assays.** Transwell assays were performed according to the manufacturer's protocol (CLS3464-12EA, Corning). Briefly, 30,000 scrambled control or CEBPA shRNA MCF10A cells were plated in the inner compartment of the transwell plate in 300  $\mu$ l 1% FBS MCF10A medium, with 700  $\mu$ l 10% FBS MCF10A medium in the outer compartment. The plate was subsequently incubated for 24 h at 37 °C and 5% CO<sub>2</sub>. Subsequently, media was removed and cells were washed, fixed, and stained with DAPI to quantify the number of migrated cells using ImageJ software.

**Confocal microscopy.** Cells were cultured on microscope glasses (Sigma-Aldrich). Coverslips were washed twice with PBS and fixed using PBS containing 4% paraformaldehyde (Merck) for 20 min at room temperature. Cells were washed twice with PBS and permeabilized with PBS containing 0.25% Triton (Sigma) for 5 min at room temperature. Posteriorly, cells were pre-incubated with PBS containing 2% of bovine serum albumin (Sigma) for 60 min at room temperature. After the blocking step, cells were incubated overnight at 4° with anti-C/EBP $\alpha$  (Abcam, ab128482), anti-E-cadherin (Abcam, ab1416), and anti-Fibronectin (BD Biosciences, 610077). Cells were washed twice in PBS, and respective secondary antibodies (BD Biosciences) were incubated for 60 min, at room temperature. Lastly, cells were washed twice in PBS and coverslips were mounted in Prolong Gold anti-fade reagent with DAPI (Invitrogen). Confocal images were acquired using a Zeiss LSM 700 fluorescence microscope (Zeiss).

**Chromatin immunoprecipitation—sequencing.** Chromatin immunoprecipitation was performed in HMLE cells using 5 of 15-cm dishes per condition. HMLE cells were either treated with TGF- $\beta$  (2.5 ng/ml; R&D Systems, 240-B-010) for 16 h or left untreated. Subsequently, double cross-linking was performed using Di(N-succinimidyl) glutarate (DSG) for 45 min followed by a 30 min incubation with formaldehyde. The reaction was quenched using incubation with 0.1 M glycine for 5 min, after which cells were washed in PBS and nuclear extracts were generated. Sonication was subsequently performed for 8 min at maximum output (ultrasonicator, Covaris), after which immunoprecipitation was performed with 1  $\mu$ g of the rabbit anti-SMAD3 (ab28379) antibodies coupled to protein A/G sepharose beads (Santa Cruz Biotechnology). Sequencing libraries were generated using the TruSeq LT kit (Illumina), and sequencing was performed on NextSeq platform (Illumina). Sequencing reads were mapped to the reference genome assembly (hg19) using Bowtie2, and peak calling was performed using MACS2. Data analysis was performed using the HOMER software package.

**3D matrigel assay.** Eight-well glass bottom chamber (Ibidi) was coated with 40  $\mu$ l of cultrex matrigel (R&D Systems) for 1 h at 37°. MCF10A cells were plated (5000 cells per chamber) in MCF10A assay medium (DMEM/F12 media (Invitrogen) supplemented with horse serum (2% final, Invitrogen), insulin (10  $\mu$ g/ml, Sigma), hydrocortisone (0.5  $\mu$ g/ml, Sigma), cholera toxin (100 ng/ml, Sigma), and 1% penicillin–streptomycin (Invitrogen)<sup>26</sup>. At specific experimental points, 5 ng/ml of TGF- $\beta$  was added. For immunofluorescence staining, medium was gently removed and cells were fixed using PBS containing 4% paraformaldehyde (Merck) for 15 min at room temperature. Cells were washed twice with PBS and permeabilized with PBS containing 0.20% Triton (Sigma), 1% dimethyl sulfoxide (DMSO; Santa Cruz Biotechnology), and 2% bovine serum albumin (Sigma) for 60 min at room temperature. After the permeabilization/blocking step, cells were incubated overnight at 4° with anti-C/EBP $\alpha$  (Cell Signaling Technology, 2295). On the following day, cells were left for 30 min at room temperature and thereafter washed twice in PBS. Respective secondary antibody (BD Biosciences) and phalloidin (Cell Signaling Technology, 8940) were incubated for 2 h at room temperature. Lastly, cells were washed twice in PBS, and 15  $\mu$ l of Prolong Gold anti-fade reagent with DAPI (Invitrogen) was added. Confocal images were acquired using a Zeiss LSM 700 fluorescence microscope (Zeiss).

**Cell growth analysis.** Equal numbers of control or C/EBP $\alpha$ -overexpressing HMLE cells were seeded on a tissue-culture-treated six-well plate. After 4 days, cell dissociation using trypsin was performed, and the number of cells in both experimental groups was analyzed using trypan blue stain.

**Luciferase assays.** For the luciferase assays, HEK293T cells at 50% confluency were transfected in 24-well plates with 0.1  $\mu$ g of CDH1 promoter luciferase reporter (pGL3 CDH1) or empty-vector control (pGL3), with 0.1  $\mu$ g of pcDNA3 empty vector or pcDNA3- HA-CEBPA and 0.02  $\mu$ g pRL-TK Renilla (Promega) as a transfection control. The cells were lysed in 50  $\mu$ l passive lysis buffer 3 days post transfection. The soluble fraction was subsequently assayed for luciferase activity with a Dual-Luciferase Reporter Assay System (Promega, USA).

**PyMT tumor organoids.** PyMT tumor organoids were generated from MMTV-PyMT;MMTV-Cre;R26R-YFP;E-Cad-mCFP mice on a FVB genetic background. Tumors were harvested and enzymatically digested using trypsin (from bovine pancreas, Sigma) and collagenase A (Roche). The digested tumors were spun down in several steps until only the cell fragments of 200 to 1000 cells were left. These organoids were embedded in BME (reduced growth factor basement membrane extract type 2, PathClear)<sup>4,50</sup>. Organoids were maintained in medium consisting of DMEM/F12 Glutamax supplemented with HEPES (1 M Gibco), 1% penicillin-streptomycin, FGF (Life technologies) and B27 (50 $\times$  Gibco).

Lentiviral particles from pLEX307-C/EBP $\alpha$  overexpressing construct and pLEX307-empty-vector construct were produced using standard protocols and as detailed above. Viral supernatants were concentrated using a centrifugal filter unit (Milipore, UFC901024) for 60 min at 4 $^{\circ}$ . The organoids were digested using Trypsin (Invitrogen Technologies), and mechanical disruption was used to generate small fragments of ~30–50 cells. The concentrated virus was diluted in DMEM/F12 Glutamax (Gibco) supplemented with HEPES (1 M Gibco), penicillin-streptomycin, 8  $\mu$ g/ml polybrene (Sigma), and 10  $\mu$ M Y27632 (Sigma). The organoids and virus-medium were spun in a 48-well plate at 600 rcf for 60 min at 32 $^{\circ}$  and afterwards placed in a 37 $^{\circ}$ C, 5% CO $_2$  incubator for 3 h. Next, the organoids were washed and plated. The organoids were selected using puromycin (100 ng/ml) Life Technologies).

**Mice.** Non-obese diabetic SCID IL-2 receptor gamma chain knockout (NSG) mice (own colony) were housed under IVC conditions. Mice received food and water ad libitum. All experiments were carried out in accordance with the guidelines of the Animal Welfare Committee of the Royal Netherlands Academy of Arts and Sciences, The Netherlands.

**Organoid transplantation and mastectomy in mice.** C/EBP $\alpha$ -overexpressing organoids or empty-vector control organoids were harvested and made into smaller pieces using mechanical disruption. The organoids were diluted in PBS. NSG mice were sedated using isoflurane inhalation anesthesia (1.5% to 2% isoflurane/O $_2$  mixture), and the organoids were injected beneath the 4th nipple. If the mouse developed a tumor of 1500 mm $^3$ , a mastectomy was performed while sedated using isoflurane inhalation anesthesia (1.5% to 2% isoflurane/O $_2$  mixture). All the mice were killed 5–8 weeks after injection. Metastatic lesions in the lungs were counted by two independent researchers.

**Mouse tumor and tissue processing for histology.** Part of tissues were fixed in periodate-lysine-paraformaldehyde (PLP) buffer (2.5 ml 4% PFA + 0.0212 g NaIO $_4$  + 3.75 ml L-Lysine + 3.75 ml P-buffer (pH 7.4)) O/N at 4 $^{\circ}$ C. The following day, the fixed tumors and tissues were washed twice with P-buffer and placed for at least 6 h in 30% sucrose at 4 $^{\circ}$ C. The tumors and tissues were then embedded in tissue freezing medium (Leica Microsystems) and stored at -80 $^{\circ}$ C before cryosectioning. Other part of the tissues were fixed in 4% paraformaldehyde, dehydrated, and embedded in paraffin.

**GFP staining of paraffin sections.** Paraffin sections of 4  $\mu$ m were firstly immersed in xylene for 10 min at room temperature and subsequently immersed in different percentages of ethanol (100%, 95%, 70%, and 50%) for ~2 min each step. Tissue was rinsed in deionized water and boiled in citrate buffer (10 mM C $_6$ H $_5$ Na $_3$ O $_7$ ·2H $_2$ O pH 6.0) for 20 min. Slides were put aside to cool down and thereafter washed in PBS. After that, slides were blocked in PO-block buffer (95 ml PBS with 5 ml 30% H $_2$ O $_2$  (37%)) for 15 min at room temperature. Washing step in PBS was performed, and second block in PBS containing 2.5% bovine serum albumin (Sigma) for 30 min at room temperature was followed. Subsequently, tissue was incubated with anti-GFP antibody (Abcam, ab6673) in PBS containing 0.5% bovine serum albumin (Sigma) for 3 days at 4 $^{\circ}$ . After primary antibody incubation, tissues were washed in PBS and incubated with rabbit anti-goat antibody (BIO-RAD, STAR194) for 1 h at room temperature. After that, samples were washed in PBS and incubated with HRP-labeled anti-rabbit (ImmunoLogic, DPVB110HRP) for 30 min at room temperature. Tissues were washed in PBS and incubated with 3,3' diaminobenzidine (DAB) containing 0.1% of H $_2$ O $_2$  (from a 30% stock) for 10 min at room temperature. Washing steps in PBS were followed, and tissues were incubated with hematoxyline counterstaining for 5 min at room

temperature. Lastly, samples were immersed in different percentages of ethanol (50%, 70%, 95%, and 100%) for ~2 min each step, followed by an incubation with xylene for 10 min at room temperature and subsequently mounted with pertex. Tissues were analyzed with an inverted bright-field microscope at 10 and 40 magnifications to morphologically identify metastasis. Metastatic lesions in the lungs were counted by two independent researchers.

**Immunofluorescence staining of frozen tissue sections.** Frozen tissues sections of 16  $\mu$ m were rehydrated in 0.1 M Tris pH 7.4 for 10 min at room temperature and subsequently permeabilized in 0.1 M Tris pH 7.4 containing 0.25% Triton (Sigma) for 5 min at room temperature. Posteriorly, blocking was performed in 0.1 M Tris pH 7.4 containing 2% bovine serum albumin for 45 min at room temperature. After blocking step, tissues were incubated overnight at 4 $^{\circ}$  with anti-C/EBP $\alpha$  (Cell Signaling Technology, 2295). In the following day, samples were washed twice in 0.1 M Tris pH 7.4, and incubated with donkey anti-rabbit secondary antibody Alexa Fluor 555 (Thermo Scientific, A31572) and TOPRO-3 for 2 h at room temperature. Lastly, tissues were washed twice in 0.1 M Tris pH 7.4 and mounted using vectashield anti-fade mounting medium (Vector labs). Confocal images were acquired using a Zeiss LSM 700 fluorescence microscope (Zeiss) at 63 magnification.

**Flow cytometry on mouse material.** Tumors were collected and minced manually on ice using sterile scalpels. The tumor mass was digested in PBS supplemented with 25  $\mu$ g/ml DNase I (Roche) and 5 Wünsch units TH Liberase /ml (Roche) at 37 $^{\circ}$ C for 35 min, followed by mashing through a 70- $\mu$ m filter (BD Falcon) while adding DMEM/F12 + GlutaMAX (GIBCO, Invitrogen Life Technologies) supplemented with 5% (v/v) fetal bovine serum (Sigma), 100  $\mu$ g/ml streptomycin, and 100 U/ml penicillin (Invitrogen Life Technologies), 5 ng/ml insulin (I0516-5ML Sigma, St. Louis, MO, USA), 5 ng/ml EGF (Invitrogen), and 25  $\mu$ g/ml DNase. After spin down (4 min at 500 RCF at RT), the pellet was resuspended in 6 ml 5 mM EDTA/PBS, after which a Ficoll gradient (Histopaque-1077, Sigma) was used to select for live cells (30 min at 400 RCF at RT, break 1). Cells were washed once in 5 mM EDTA/PBS and centrifuged (4 min at 500 RCF at RT) before proceeding with antibody labeling. The cells were blocked in 80% FACS buffer (5 mM EDTA in PBS supplemented with 5% fetal calf serum)/20% serum mix (50/50 normal goat serum (monx10961, Monosan) and Fc $\gamma$ II/III receptor blocking serum 2.4G2 (kind gift from Kiki Tesselar, UMCU, The Netherlands)) for 10 min on ice before labeling with E-cad-eFluor660 (DECMA-1, eBioscience). The cells were analyzed on FACS Jazz (BD biosciences) using the following strategy: a broad FSC SSC gate was followed by a gate excluding doublets, after which immune cells and megakaryocytes were excluded in a dump channel. YFP-positive and DAPI-negative tumor cells were gated using E-cad-mCFP and E-cadherin 647 lasers.

**Statistical analysis.** Data represented as mean  $\pm$  SD of at least three independent experiments. Differences were analyzed by unpaired two-tailed *t* test between two groups and by two-way ANOVA for differences between more than two groups. Exceptionally, qPCR results obtained on Fig. 2b were analyzed using unpaired one-tailed *t* test. ns indicates nonsignificant. \**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001; \*\*\*\**p* < 0.0001.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Data presented on Fig. 2a and S1c were generated by analyzing the data available under the accession numbers GSE104761 and GSE24202, respectively. The data sets generated during this study (RNA-seq) were deposited under the accession number GSE143612. Uncropped western blot images can be found in the Supplementary Information section (Supplementary Figs. 9–15).

Received: 29 July 2017; Accepted: 20 January 2020;  
Published online: 07 February 2020

## References

- Lim, J. & Thiery, J. P. Epithelial-mesenchymal transitions: insights from development. *Development* **139**, 3471–3486 (2012).
- Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* **119**, 1420–1428 (2009).
- Fischer, K. R. et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **527**, 472–476 (2015).
- Beerling, E. et al. Plasticity between epithelial and mesenchymal states unlinks EMT from metastasis-enhancing stem cell capacity. *Cell Rep.* **14**, 2281–2288 (2016).



5. Krebs, A. M. et al. The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. *Nat. Cell Biol.* **19**, 518–529 (2017).
6. Du, B. & Shim, J. Targeting epithelial–mesenchymal transition (EMT) to overcome drug resistance in cancer. *Molecules* **21**, 965 (2016).
7. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196 (2014).
8. Cantelli, G., Crosas-Molist, E., Georgouli, M. & Sanz-Moreno, V. TGF $\beta$ -induced transcription in cancer. *Semin. Cancer Biol.* <https://doi.org/10.1016/j.semcancer.2016.08.009> (2016).
9. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196 (2014).
10. Park, J. & Schwarzbauer, J. E. Mammary epithelial cell interactions with fibronectin stimulate epithelial–mesenchymal transition. *Oncogene* **33**, 1649–1657 (2014).
11. Miller, M., Shuman, J. D., Sebastian, T., Dauter, Z. & Johnson, P. F. Structural basis for DNA recognition by the basic region leucine zipper transcription factor CCAAT/enhancer-binding protein alpha. *J. Biol. Chem.* **278**, 15178–15184 (2003).
12. Ramji, D. P. & Foka, P. CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem. J.* **365**, 561–575 (2002).
13. Leroy, H. et al. CEBPA point mutations in hematological malignancies. *Leukemia* **19**, 329–334 (2005).
14. Nerlov, C. C/EBPalpha mutations in acute myeloid leukaemias. *Nat. Rev. Cancer* **4**, 394–400 (2004).
15. Vogelstein, B. et al. Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
16. Lourenço, A. R. & Coffey, P. J. A tumor suppressor role for C/EBP $\alpha$  in solid tumors: more than fat and blood. *Oncogene*. <https://doi.org/10.1038/nc.2017.151> (2017).
17. Gery, S. et al. Down-regulation and growth inhibitory role of C/EBP $\alpha$  in breast cancer. *Clin. Cancer Res.* **11**, 3184–3190 (2005).
18. Vervoort, S., Lourenço, A. R., van Bostel, R. & Coffey, P. J. SOX4 mediates TGF $\beta$ -induced expression of mesenchymal markers during mammary cell epithelial to mesenchymal transition. *PLoS ONE* **8**, e53238 (2013).
19. Mani, S. A. et al. The epithelial–mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715 (2008).
20. Martis, P. C. et al. C/EBP $\alpha$  is required for lung maturation at birth. *Development* **133**, 1155–1164 (2006).
21. Taube, J. H. et al. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl Acad. Sci. USA* **107**, 15449–15454 (2010).
22. Zhang, J. et al. TGF $\beta$ -induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci. Signal.* **7**, ra91 (2014).
23. Maeda, M., Johnson, K. R. & Wheelock, M. J. Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition. *J. Cell Sci.* **118**, 873–887 (2005).
24. Morikawa, M., Koinuma, D., Miyazono, K. & Heldin, C.-H. Genome-wide mechanisms of Smad binding. *Oncogene*. <https://doi.org/10.1038/nc.2012.191> (2012).
25. Aiello, N. M. & Kang, Y. Context-dependent EMT programs in cancer metastasis. *J. Exp. Med.* **216**, 1016–1026 (2019).
26. Debnath, J., Muthuswamy, S. K. & Brugge, J. S. Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* **30**, 256–268 (2003).
27. Naber, H. P. H., Wiercinska, E., ten Dijke, P. & van Laar, T. Spheroid assay to measure TGF $\beta$ -induced invasion. *J. Vis. Exp.* <https://doi.org/10.3791/3337> (2011).
28. Arising, K. R. & Fischer. Upholding a role for EMT in breast cancer metastasis. *Nat. Publ. Gr.* **547**, 472–476 (2017).
29. Xu, J., Lamouille, S. & Derynck, R. TGF $\beta$ -induced epithelial to mesenchymal transition. *Cell Res.* **19**, 156–172 (2009).
30. Moustakas, A. & Heldin, C.-H. Mechanisms of TGF $\beta$ -induced epithelial–mesenchymal transition. *J. Clin. Med.* **5**, 63 (2016).
31. Johansson, J. et al. MiR-155-mediated loss of C/EBP $\beta$  shifts the TGF $\beta$  response from growth inhibition to epithelial–mesenchymal transition, invasion and metastasis in breast cancer. *Oncogene* **32**, 5614–5624 (2013).
32. Massague, J. Smad transcription factors. *Genes Dev.* **19**, 2783–2810 (2005).
33. Fang, W. Bin et al. TGF $\beta$  negatively regulates CXCL1 chemokine expression in mammary fibroblasts through enhancement of Smad2/3 and suppression of HGF/c-Met signaling mechanisms. *PLoS ONE* **10**, e0135063 (2015).
34. Marchildon, F., St-Louis, C., Akter, R., Roodman, V. & Wiper-Bergeron, N. L. Transcription factor Smad3 is required for the inhibition of adipogenesis by retinoic acid. *J. Biol. Chem.* **285**, 13274–13284 (2010).
35. Wheelock, M. J., Shintani, Y., Maeda, M., Fukumoto, Y. & Johnson, K. R. Cadherin switching. *J. Cell Sci.* **121**, 727–735 (2008).
36. Pishvaian, M. J. et al. Cadherin-11 is expressed in invasive breast cancer cell lines. *Cancer Res.* **59**, 947–952 (1999).
37. Radisky, E. S. & Radisky, D. C. Matrix metalloproteinase-induced epithelial–mesenchymal transition in breast cancer. *J. Mammary Gland Biol. Neoplasia* **15**, 201–212 (2010).
38. Lee, Y.-H., Albig, A. R., Regner, M., Schiemann, B. J. & Schiemann, W. P. Fibulin-5 initiates epithelial–mesenchymal transition (EMT) and enhances EMT induced by TGF-beta in mammary epithelial cells via a MMP-dependent mechanism. *Carcinogenesis* **29**, 2243–2251 (2008).
39. Nieman, M. T., Prudoff, R. S., Johnson, K. R. & Wheelock, M. J. N-cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression. *J. Cell Biol.* **147**, 631–644 (1999).
40. Hazan, R. B., Phillips, G. R., Qiao, R. F., Norton, L. & Aaronson, S. A. Exogenous expression of N-cadherin in breast cancer cells induces cell migration, invasion, and metastasis. *J. Cell Biol.* **148**, 779–790 (2000).
41. Gheldof, A. & Bex, G. Cadherins and epithelial-to-mesenchymal transition. *Prog. Mol. Biol. Transl. Sci.* **116**, 317–336 (2013).
42. Mascuchań, T. N. H., Shawber, C. J., Funahashi, Y., Li, C.-M. & Kitajewski, J. Wnt/ $\beta$ -catenin signaling induces proliferation, survival and interleukin-8 in human endothelial cells. *Angiogenesis* **8**, 43–51 (2005).
43. Chen, S. et al. Wnt-1 signaling inhibits apoptosis by activating beta-catenin/T cell factor-mediated transcription. *J. Cell Biol.* **152**, 87–96 (2001).
44. Schmalhofer, O., Brabletz, S. & Brabletz, T. E-cadherin, beta-catenin, and ZEB1 in malignant progression of cancer. *Cancer Metastasis Rev.* **28**, 151–166 (2009).
45. Lu, J. et al. C/EBP $\alpha$  suppresses lung adenocarcinoma cell invasion and migration by inhibiting  $\beta$ -catenin. *Cell. Physiol. Biochem.* **42**, 1779–1788 (2017).
46. Huan, H. et al. C/EBP $\alpha$  short-activating RNA suppresses metastasis of hepatocellular carcinoma through inhibiting EGFR/ $\beta$ -catenin signaling mediated EMT. *PLoS ONE* **11**, e0153117 (2016).
47. Zheng, X. et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* **527**, 525–530 (2015).
48. Datta, J. et al. Metallothionein expression is suppressed in primary human hepatocellular carcinomas and is mediated through inactivation of CCAAT/enhancer binding protein alpha by phosphatidylinositol 3-kinase signaling cascade. *Cancer Res.* **67**, 2736–2746 (2007).
49. Lektrom-Himes, J. & Xanthopoulos, K. G. Biological role of the CCAAT/enhancer-binding protein family of transcription factors. *J. Biol. Chem.* **273**, 28545–28548 (1998).
50. Nguyen-Ngoc, K.-V. et al. ECM microenvironment regulates collective migration and local dissemination in normal and malignant mammary epithelium. *Proc. Natl Acad. Sci. USA* **109**, E2595–E2604 (2012).

## Acknowledgements

We would like to thank Prof. Robert A. Weinberg and Dr. Patrick Derksen for providing us the HMLE and MCF10A cells, respectively. A.R.L. was supported by a FCT (Fundação para a Ciência e a Tecnologia) fellowship. A.R.L., S.J.V., M.G.R. and C.L.F. were supported by a grant from the Dutch Cancer Society (KWF).

## Author contributions

P.J.C. supervised the study. A.R.L., M.G.R., D.S., J.v.R. and P.J.C. designed experiments. A.R.L., C.L.F., C.E.P., A.S.M. and D.S. performed experiments. S.J.V. performed NGS analysis. A.R.L., M.G.R. and P.J.C. interpreted experiments. A.R.L. and P.J.C. wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-14556-x>.

**Correspondence** and requests for materials should be addressed to P.J.C.

**Peer review information** *Nature Communications* thanks Erik Thompson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020