Cancer Cell
Article

Sox4 Is a Master Regulator of Epithelial-Mesenchymal Transition by Controlling Ezh2 Expression and Epigenetic Reprogramming

Neha Tiwari,[1,4] Vijay K. Tiwari,[2,5] Lorenz Waldmeier,[1] Piotr J. Balwierz,[3] Phil Arnold,[3] Mikhail Pachkov,[3] Nathalie Meyer-Schaller,[1] Dirk Schübeler,[2] Erik van Nimwegen,[3] and Gerhard Christofori[1,*]
[1]Department of Biomedicine, University of Basel, 4058 Basel, Switzerland
[2]Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland
[3]Biozentrum, University of Basel, and Swiss Institute of Bioinformatics, 4056 Basel, Switzerland
[4]Present address: Institute of Physiological Chemistry, University Medical Center, Johannes Gutenberg University Mainz, 55131 Mainz, Germany
[5]Present address: Institute for Molecular Biology gGmbH (IMB), 55128 Mainz, Germany
*Correspondence: gerhard.christofori@unibas.ch
http://dx.doi.org/10.1016/j.ccr.2013.04.020

SUMMARY

Gene expression profiling has uncovered the transcription factor Sox4 with upregulated activity during TGF-β-induced epithelial-mesenchymal transition (EMT) in normal and cancerous breast epithelial cells. Sox4 is indispensable for EMT and cell survival in vitro and for primary tumor growth and metastasis in vivo. Among several EMT-relevant genes, Sox4 directly regulates the expression of *Ezh2*, encoding the Polycomb group histone methyltransferase that trimethylates histone 3 lysine 27 (H3K27me3) for gene repression. Ablation of Ezh2 expression prevents EMT, whereas forced expression of Ezh2 restores EMT in Sox4-deficient cells. Ezh2-mediated H3K27me3 marks associate with key EMT genes, representing an epigenetic EMT signature that predicts patient survival. Our results identify Sox4 as a master regulator of EMT by governing the expression of the epigenetic modifier Ezh2.

INTRODUCTION

Epithelial-mesenchymal transition (EMT) is a cellular mechanism known to constitute the core of normal embryonic development (Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and Weinberg, 2009; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006). Similar, yet pathophysiological transitions occur during the progression of epithelial tumors, endowing cancer cells with increased motility and invasiveness to seed metastasis and, sometimes, after metastatic dissemination to redifferentiate into epithelial structures by mesenchymal-epithelial transition (Kang and Massagué, 2004; Thiery and Morgan, 2004). Multiple oncogenic events and signaling pathways, mediated, for example, by transforming growth factor β (TGF-β), hepatocyte growth factor (HGF), Wnt and Notch signaling, or oncogenic Src or Ras activation, are implicated in the induction of EMT (Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and Weinberg, 2009; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006).

Sox4 is a member of the Sox (SRY-related HMG-box) family of transcription factors with a critical role in embryonic development and in cell-fate determination during organogenesis of the heart (Restivo et al., 2006; Schilham et al., 1996), pancreas (Lioubinski et al., 2003; Wilson et al., 2005), and brain (Cheung et al., 2000; Hong and Saint-Jeannet, 2005), and in B and T lymphocyte differentiation (Cheung et al., 2000; Hong and Saint-Jeannet, 2005; Lioubinski et al., 2003; Schilham et al., 1996, 1997; van de Wetering et al., 1993; Wilson et al., 2005). Sox4 gene expression is upregulated in many cancer types, and increased Sox4 activity contributes to cellular transformation (Liu et al., 2006; Shin

Significance

An epithelial-mesenchymal transition (EMT) is a key process during organismal development and the progression of epithelial tumors to metastatic cancers. Here, we identify the transcription factor Sox4 as a master regulator of EMT. Sox4 appears to act highly upstream in the epistatic hierarchy of EMT regulation. It controls a number of EMT-relevant genes in addition to *Ezh2*. Conversely, Ezh2 is able to functionally replace Sox4 during EMT and regulates the expression of a number of EMT-associated genes. *Ezh2* thus represents one critical Sox4 target gene during EMT. The results exemplify an important interplay between transcriptional and epigenetic control during EMT and suggest that the inhibition of Ezh2 could be an attractive avenue for the therapeutic intervention of malignant tumor progression.

et al., 2004), cell survival (Aaboe et al., 2006; Ahn et al., 2002; Liu et al., 2006; Pramoonjago et al., 2006), and metastasis (Liao et al., 2008; Tavazoie et al., 2008). For example, restoration of miR-335 expression suppresses lung and bone metastasis in human cancer cells by interfering with the expression of Sox4 (Tavazoie et al., 2008). Interestingly, Sox4 also directly modulates key cellular regulators, including the genes encoding for epidermal growth factor receptor (EGFR), tenascin-C, heat shock protein 70 (Hsp70), frizzled homolog 5 (Fzd5), Delta-like 1 (Dll1), and Patched-1 (Ptch1), the transcriptional regulators Mll, Foxa1, Znf281, and Nkx3-1, and components of the RNAi machinery, including Dicer, Argonaute 1, and RNA helicase A (Scharer et al., 2009). Finally, Sox4 regulates Wnt signaling by directly binding to β-catenin and Tcf family members (Sinner et al., 2007). Most recently, Sox4 has been reported to induce EMT and to cooperate with oncogenic Ras in breast cancer progression (Zhang et al., 2012); however, its direct transcriptional target genes during EMT have remained elusive. Here, we demonstrate a central role of Sox4 in EMT as well as in primary tumor growth and metastasis by directly regulating the expression of a number of genes with critical functions in EMT, including *Ezh2*.

The histone methyltransferase Ezh2 (Enhancer of Zeste homolog 2) is a component of the Polycomb (PcG) repressive complex 2 (PRC2), which epigenetically regulates genes involved in cell fate decisions. Ezh2 specifically trimethylates nucleosomal histone H3 at lysine 27 (H3K27me3), an epigenetic modification associated with gene silencing (Sparmann and van Lohuizen, 2006). A conditional knockout of the *Ezh2* gene in basal keratinocytes leads a precocious acquisition of an epidermal barrier function in the embryo (Ezhkova et al., 2009), while a conditional knockout of *Ezh2* in B-lymphocytes results in improper IgH rearrangements (Su et al., 2003), suggesting a role for Ezh2 methyltransferase in cell differentiation and maturation. Ezh2 is found to be highly expressed in a variety of cancer types where the genomic loss of miR-101 leads to increased expression of Ezh2 and concomitant deregulation of epigenetic pathways, altogether resulting in cancer progression (Varambally et al., 2008). Furthermore, Ezh2 can induce EMT and increase the metastatic potential of prostate cancer cells by downregulation of DAB2IP, a tumor-suppressive Ras GTPase-activating protein (RasGAP) (Chen et al., 2005; Min et al., 2010). Finally, low Ezh2 expression levels correlate with metastasis-free survival in breast cancer (Cao et al., 2008; Kleer et al., 2003). However, whether and how Ezh2-mediated epigenetic mechanisms contribute to the transcriptional reprogramming that accompanies EMT is still poorly understood.

RESULTS

Sox4 Is Required for EMT

To identify critical genes underlying early, intermediate, and late stages of EMT, we induced EMT in the untransformed normal murine mammary gland (NMuMG) cell line by treatment with TGF-β for 0, 1, 4, 7, 10, and 20 days (data not shown). During this time course, the cells underwent progressive EMT and acquired a complete mesenchymal morphology (Lehembre et al., 2008). Motif activity response analysis (MARA) (Suzuki et al., 2009) of gene expression data derived from the EMT time course predicted several transcription factor binding motifs to be important regulators of the EMT expression dynamics, including a motif bound by Sox transcription factors (Figure S1A available online). Gene expression profiling and quantitative RT-PCR experiments revealed that Sox4 and Sox9 were the only Sox family members significantly upregulated in their expression during TGF-β-induced EMT in NMuMG cells (Figures S1B–S1D). Sox4 expression was also robustly induced during EMT in a number of other cellular EMT systems, including TGF-β-induced EMT in Py2T murine breast cancer cells derived from a tumor of MMTV-PyMT transgenic mice (Waldmeier et al., 2012), in MCF10A human breast epithelial cells, in EpRas murine mammary epithelial cells (Figures S1D–S1G), and in human mammary epithelial cells (Kloeker et al., 2004). Sox4 expression was also increased in MCF7 human breast cancer cells that exert full EMT upon shRNA-mediated knockdown of E-cadherin expression (Lehembre et al., 2008; Figure S1H).

To directly assess the role of Sox4 and Sox9 in EMT, NMuMG cells were transfected with a pool of two different siRNAs against Sox4 (siSox4) and Sox9 (siSox9), as well as siSox4 and siSox9 together, to transiently ablate the expression of Sox4 and Sox9 in the absence and presence of TGF-β, resulting in an efficient repression of Sox4 and Sox9 expression (Figures 1A; Figures S1I and S1J). As reported recently (Zhang et al., 2012), Sox4-ablated NMuMG cells were not able to undergo EMT and largely retained their epithelial morphology during TGF-β treatment, whereas cells transfected with control-siRNA (siControl) changed to a mesenchymal, fibroblastoid phenotype (Figure 1B). In contrast, the single ablation of Sox9 did not have any effect on TGF-β-induced EMT, while the concomitant ablation of Sox4 and Sox9 mimicked the single ablation of Sox4 (Figures S1K and S1L). Immunofluorescence staining revealed that Sox4 depletion in the absence of TGF-β did not substantially affect the epithelial morphology of NMuMG cells (Figure 1C). In the presence of TGF-β, Sox4-depleted cells maintained the epithelial markers E-cadherin and ZO-1 at the cell membrane, and the expression of the mesenchymal marker N-cadherin was decreased compared to that observed for TGF-β-treated siControl cells (Figure 1C). Moreover, TGF-β-mediated remodeling of the cytoskeleton from cortical actin to stress fibers, as determined by phalloidin staining, and focal adhesion formation, as determined by paxillin staining, were significantly reduced in siSox4 cells (Figure 1C; Figure S1M). The failure to lose epithelial and gain mesenchymal marker expression in TGF-β-treated, Sox4-depleted cells was further confirmed by immunoblotting analysis (Figure 1D). Notably, Sox4 depletion in NMuMG cells treated previously with TGF-β for 15 days caused these mesenchymal cells to revert to an epithelial phenotype (Figure 1E). Similar to NMuMG cells, shRNA-mediated stable knockdown of Sox4 in Py2T cells markedly delayed TGF-β-induced EMT (Figures S1N–S1P). Together, these results demonstrate that Sox4, but not Sox9, is required for the induction and maintenance of TGF-β-induced EMT in murine mammary epithelial cells.

Sox4 Is Required and Sufficient for Cell Survival and Cell Migration

Sox4 has been implicated in cell survival in a variety of cancers (Hur et al., 2010; Pramoonjago et al., 2006; Shen et al., 2010). Indeed, Sox4-depleted NMuMG and Py2T cells also showed a

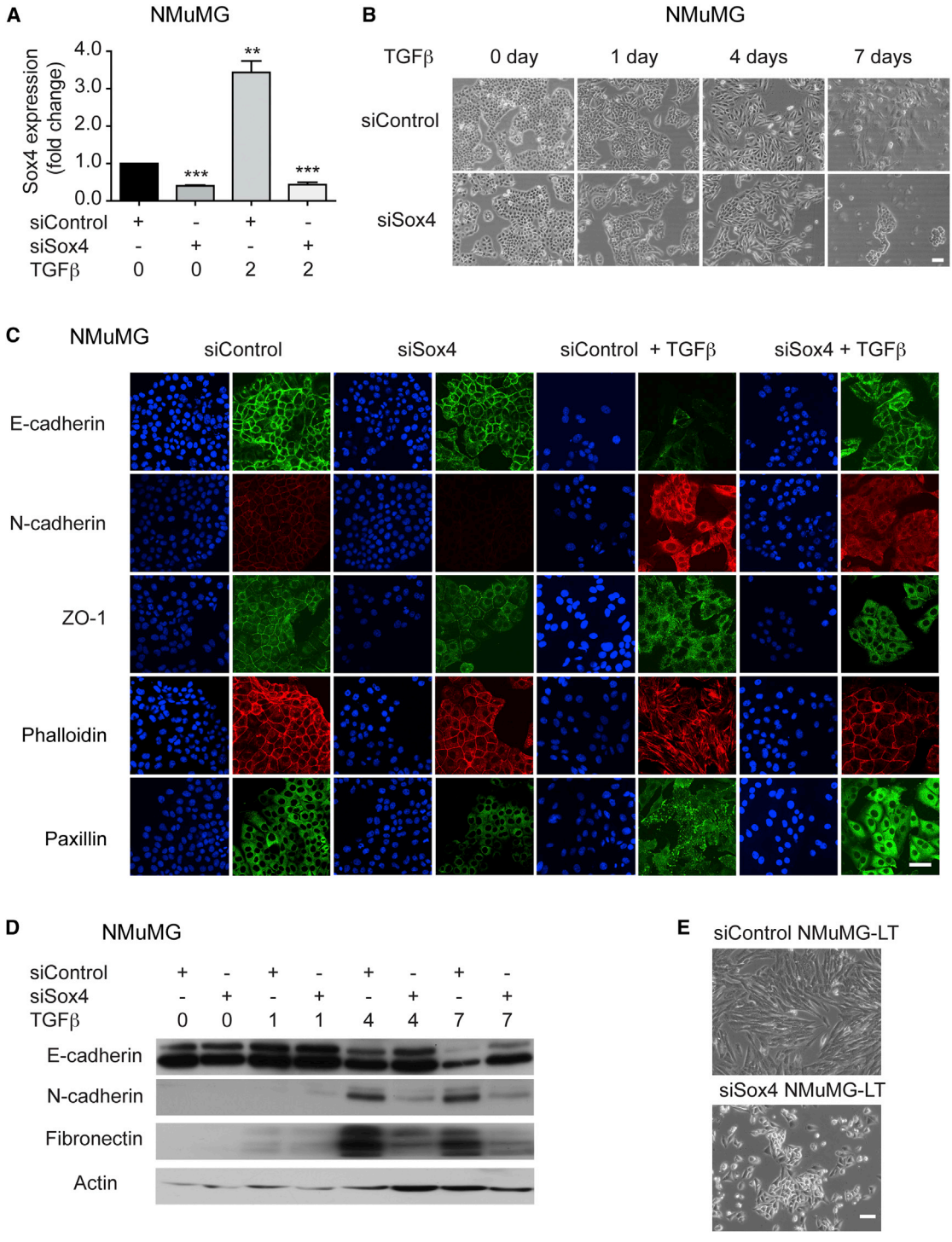

Figure 1. Sox4 Is Required for EMT

(A) Sox4 mRNA levels were determined by quantitative RT-PCR in NMuMG cells transfected with either control siRNA (siControl) or with siRNA against Sox4 (siSox4) in the absence or presence of TGF-β for 2 days. Fold changes are relative to those of cells transfected with control siRNA in the absence of TGF-β.

(B) NMuMG cells transfected with either siControl or siSox4 were treated with TGF-β for the number of days indicated, and their morphology was evaluated by phase-contrast microscopy. Scale bar, 100 μm.

(C) Immunoflorescence microscopy analysis of changes in the localization and expression levels of marker proteins during EMT. NMuMG cells transfected with either siControl or siSox4 were left untreated or treated with TGF-β for 2 days and were stained with antibodies against the epithelial markers E-cadherin and ZO-1, against the mesenchymal marker N-cadherin, against paxillin to detect focal adhesion plaques, and with phalloidin to visualize the actin cytoskeleton. Scale bar, 50 μm.

(legend continued on next page)

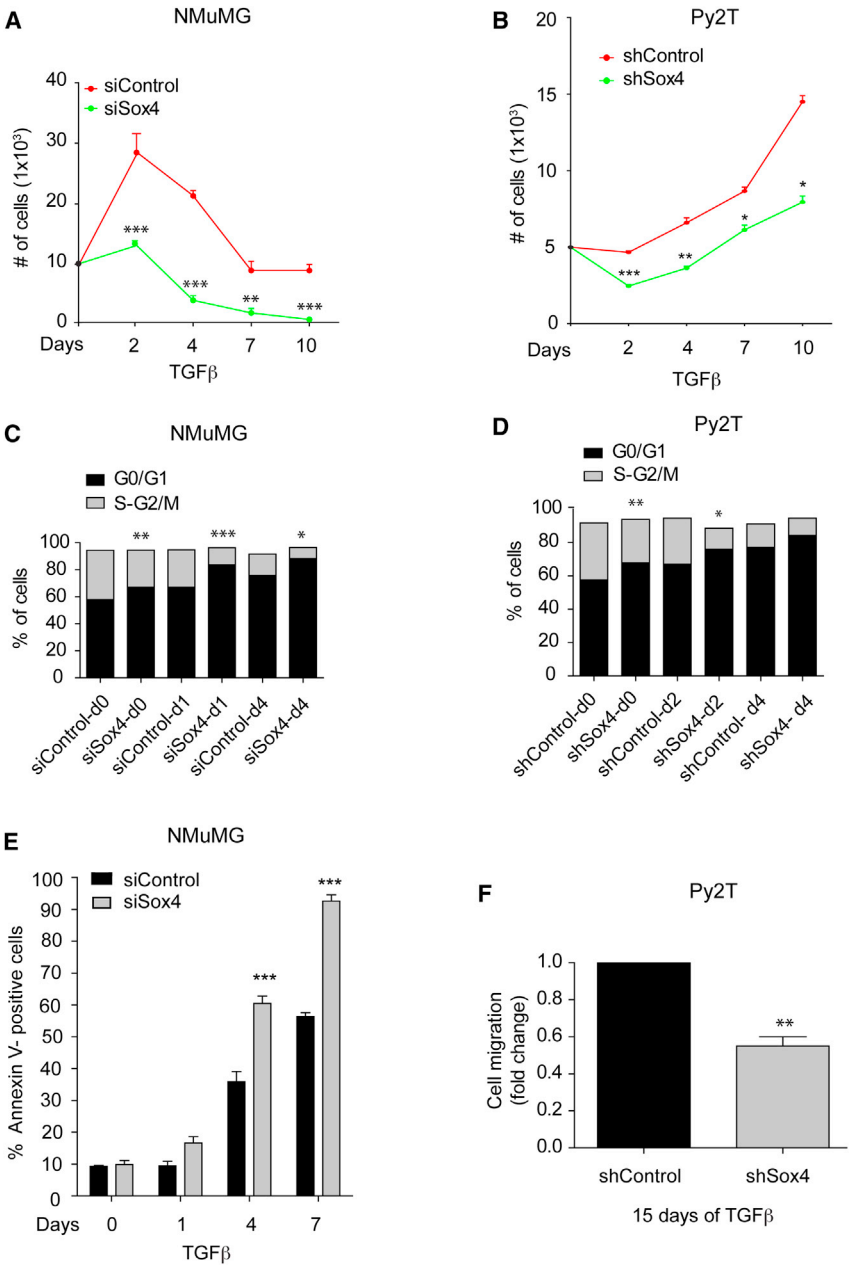

Figure 2. Sox4 Is Required for Cell Survival and Cell Migration during TGF-β-Induced EMT

(A and B) siRNA or shRNA-mediated ablation of Sox4 expression (siSox4 and shSox4, respectively) during TGF-β treatment of NMuMG (A) and Py2T (B) cells results in a significant decrease in cell numbers compared to transfection with control siRNA or shRNA (siControl and shControl, respectively).

(C and D) NMuMG cells (C) or Py2T cells (D) transfected with either siControl and shControl or siSox4 and shSox4 were treated with TGF-β for the days indicated (d0 to d4). Cells were stained with propidium iodide (PI), and the percentages of cells in G0/G1 and S-G2/M phases of the cell cycle were determined by flow cytometry.

(E) NMuMG cells transfected with either siControl or siSox4 were treated with TGF-β for the days indicated, and the rates of apoptosis were determined by Annexin-V staining and flow cytometry.

(F) shSox4-expressing and shControl-expressing Py2T murine breast cancer cells were treated for 15 days with TGF-β, and cell migration was determined in a modified Boyden chamber migration assay 20 hr after seeding and using 20% FBS as a chemoattractant.

Statistical values were calculated using an unpaired, two-tailed t test. *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. Error bars indicate the mean ± SD.

significant reduction in cell growth in comparison to control cells and displayed a discernable sensitivity toward TGF-β-mediated growth inhibition (Figures 2A and 2B). Propidium iodide staining followed by flow-cytometry-based cell cycle analysis revealed a G0/G1 cell cycle arrest after Sox4 depletion (Figures 2C and 2D). Annexin-V staining demonstrated a marked increase in the levels of apoptosis in siSox4-transfected NMuMG cells (Figure 2E), but not in Py2T cells (data not shown). These results indicate that Sox4 is required for cell survival and proliferation during TGF-β-induced EMT and that the prosurvival function of Sox4 only becomes critical in nontransformed cells. Cell migration is a characteristic feature of EMT and metastasis (Brabletz et al., 2005; Christofori, 2006; Grünert et al., 2003; Huber et al., 2005; Thiery and Sleeman, 2006). Transwell migration assays with the apoptosis-resistant Py2T breast cancer cells revealed that shSox4 Py2T cells migrated significantly less than shControl cells after long-term treatment with TGF-β (Figure 2F).

We next asked whether Sox4 was sufficient to induce EMT. Transient expression of HA-tagged Sox4 in NMuMG cells at the levels observed during TGF-β-induced EMT increased filopodia formation and cell scattering (Figure 3A) accompanied by the gain of the mesenchymal markers N-cadherin and fibronectin (Figures 3B and 3C). Although its mRNA and protein levels were not apparently affected, E-cadherin was displaced at the cell junctions of HA-Sox4-expressing cells (Figure 3D). Stable expression of

(D) Expression of E-cadherin, N-cadherin, and fibronectin was determined by immunoblotting during TGF-β-induced EMT in NMuMG cells transfected with either siSox4 or siControl. Immunoblotting for actin was used as a loading control.

(E) Morphology of long-term TGF-β-treated NMuMG (NMuMG-LT) cells transfected with either siControl or siSox4, as evaluated by phase-contrast microscopy. Scale bar, 100 μm.

Statistical values were calculated using a paired, two-tailed t test. **$p \leq 0.01$; ***$p \leq 0.001$. Error bars indicate the mean ± SD.

See also Figure S1.

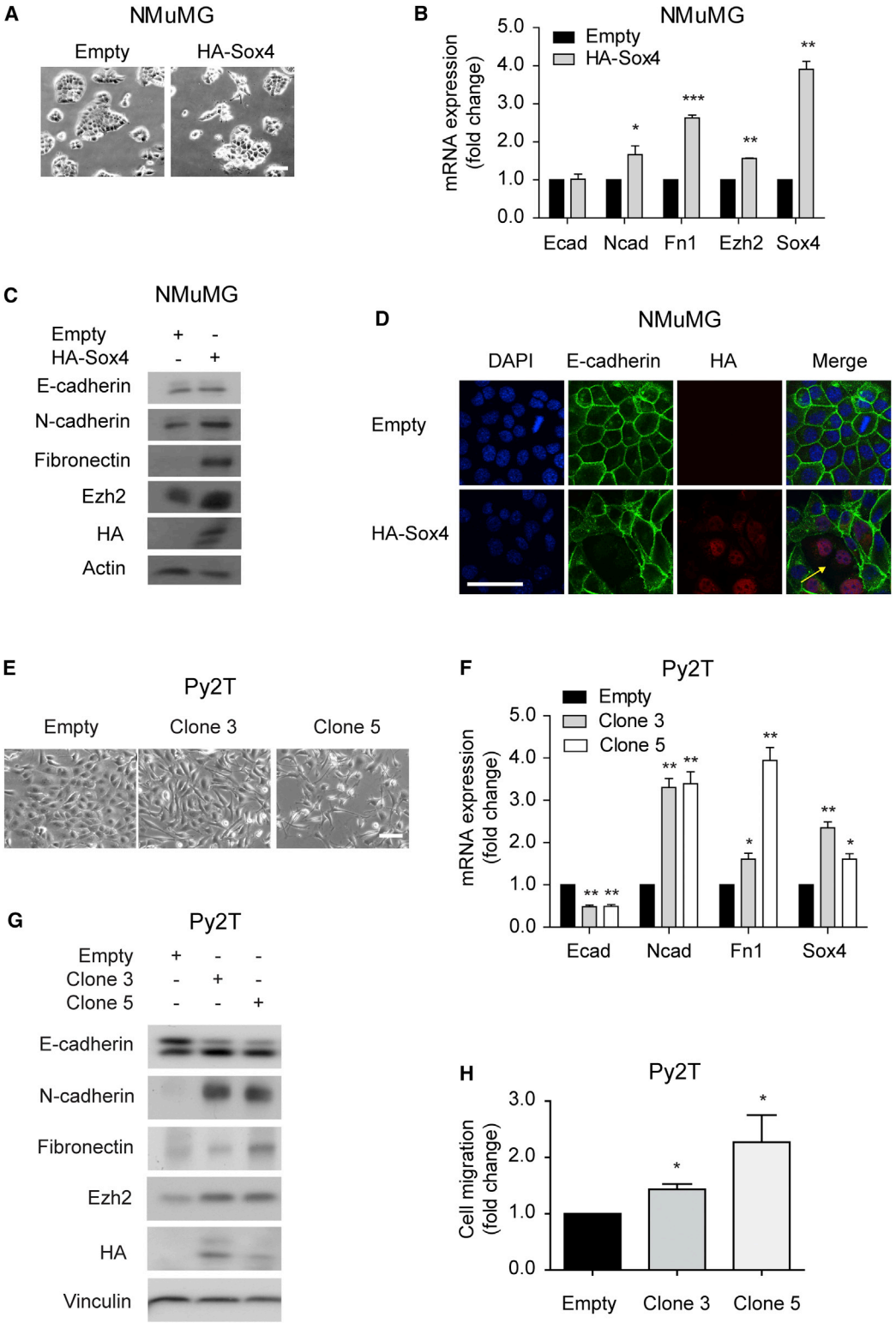

Figure 3. Ectopic Expression of Sox4 Promotes EMT and Cell Migration

(A) NMuMG cells were transiently transfected with pcDNA3-HA and pcDNA3-HA-Sox4 constructs, and cellular morphology was evaluated. Note the formation of filopodia and the cell dispersal upon Sox4 expression (HA-Sox4) as compared to empty vector control (Empty). Scale bar, 100 μm.

(B) The expression of E-cadherin (Ecad), N-cadherin (Ncad), fibronectin (Fn1), Ezh2, and Sox4 was determined by quantitative RT-PCR in NMuMG cells transiently transfected with HA-Sox4 or with empty vector (Empty) in the absence of TGF-β. Fold changes are relative to those in cells transfected with empty vector.

(legend continued on next page)

HA-tagged Sox4 in Py2T cells also induced mesenchymal cell morphology in the absence of TGF-β, with a moderate loss of E-cadherin expression but a marked gain in N-cadherin and fibronectin expression (Figures 3E–3G). Furthermore, the forced expression of Sox4 promoted Py2T cell migration (Figure 3H). These results suggest that Sox4 is not only required but also sufficient for the induction of an EMT.

Sox4 Is Required for Tumorigenesis and Metastatic Spread

The critical role of Sox4 in EMT raised the question of whether Sox4 contributes to malignant tumor progression and metastasis. First, computational analysis revealed a significant correlation between high Sox4 expression and poor metastasis-free survival in the "Schmidt" database of 200 early-stage, lymph-node-negative breast cancer patients (Schmidt et al., 2008) (Figure 4A). The analysis of databases assembled of late-stage breast cancer patient data did not reveal a significant prognostic value for Sox4 expression (data not shown), suggesting that Sox4 may play a critical role in the early stages of the malignant progression of breast cancer.

Next, we investigated the functional contribution of Sox4 to primary tumor growth and metastasis formation in vivo. Py2T cells stably expressing an shRNA against Sox4 (shSox4) and control-transfected cells (shControl), both expressing the firefly luciferase gene, were orthotopically implanted into mammary fat pads of nude mice, and tumor growth and metastasis in lymph nodes, lungs, and livers were quantified. Quantitative RT-PCR (Figure S2A) and immunostaining of histological tumor sections (Figure 4B) documented an efficient depletion of Sox4 expression in the cultured cell line before implantation and in primary tumors, respectively. Ablation of Sox4 expression in Py2T cells lead to a significant reduction in primary tumor growth (Figure 4C). Luciferase activity levels representing the presence of tumor cells in distant organs were found to be decreased in axillary and inguinal lymph nodes and in lungs and livers of mice transplanted with shSox4 Py2T cells as compared to shControl Py2T cells (Figure S2B), even when the luciferase activities were normalized to the decreased primary tumor sizes observed with shSox4 Py2T cells (metastatic index; Figure 4D). The reduced ability of shSox4 Py2T cells to grow as primary tumors was also observed upon subcutaneous implantation into nude mice (Figure 4E).

We next addressed whether Sox4 is required for metastasis formation in other cancer types. Similar to NMuMG cells, in B16-F10 melanoma cells Sox4 expression was increased upon TGF-β treatment (Figure S2C). Sox4 expression was then ablated in B16-F10 melanoma cells by stable expression of shRNA targeting Sox4 (shSox4; Figure S2D), and the cells were injected into the tail vein of C57Bl/6 mice. Quantification of lung metastasis showed that Sox4-depleted B16-F10 cells were significantly impaired in lung colonization as compared to shControl-expressing B16-F10 cells (Figures S2E and S2F). Taken together, these results indicate that the transcription factor Sox4 exerts a critical function in primary tumor growth and metastasis formation.

Ezh2 Is a Direct Transcriptional Target of Sox4

The critical role of Sox4 in EMT and tumor progression motivated us to identify the genes that were directly regulated by Sox4. We first compared the genome-wide gene expression profiles of siControl and siSox4 NMuMG cells in the absence and presence of TGF-β. Genes found to be differentially expressed in dependence on Sox4 function were further analyzed for Sox4 binding motifs within one kilobase of their transcription start sites (Figure S1A). Of the 189 genes fulfilling these criteria (Table S1), 106 also changed in their expression levels during TGF-β-induced EMT. Gene ontology analysis revealed that 32 of these genes have been previously implicated in processes relevant to tumor progression (Table S2). In order to identify the genes directly regulated by Sox4, we assessed the occupancy of Sox4 at the promoters of these 32 genes. NMuMG cells transiently expressing an HA-tagged version of Sox4 (HA-Sox4) were subjected to chromatin immunoprecipitation (ChIP) with HA-specific antibodies followed by quantitative PCR using primers specific for the region spanning the Sox4 motif found in the promoter of these genes. Of the 32 genes analyzed, the promoters of 28 genes—16 in NMuMG cells and 12 in Py2T cells—were directly bound by Sox4, including the key EMT genes *Spred1*, *Edn1*, *Palld*, *Cyr61*, *Ereg*, and *Areg* (Figures S3A–S3H; data not shown). Genes that were not immunoprecipitated by HA-antibody, as well as an intergenic region, served as negative controls (Figure S3I).

Snail, Zeb, and Twist family transcriptional repressors of E-cadherin gene expression are known to play critical roles in the induction of EMT. Hence, we assessed whether Sox4 is epistatic to the expression of these EMT inducers or whether they are regulating Sox4 expression during EMT. Ablation of Sox4 expression in NMuMG cells resulted in the reduced expression of Snail2, Zeb2, and Twist1, but not Snail1 and Zeb1, whereas in Py2T cells the expression of all five repressors was

(C) The expression of E-cadherin, N-cadherin, fibronectin, Ezh2, and HA-tagged Sox4 was determined by immunoblotting of NMuMG cells transiently transfected with empty vector or with HA-Sox4. Actin was used as the loading control.
(D) The localization of E-cadherin was determined by immunofluorescence staining in NMuMG cells transiently transfected with empty-vector control or HA-Sox4. The cells marked with an arrow express Sox4 (HA) and have lost E-cadherin at their cell membranes. Scale bar, 50 μm.
(E) Cell scattering and mesenchymal morphology are observed even in the absence of TGF-β in Py2T cell clones stably expressing HA-Sox4 (Clones 3 and 5), as compared to empty-vector-transfected control cells (Empty). Scale bar, 100 μm.
(F) The expression of E-cadherin (Ecad), N-cadherin (Ncad), fibronectin (Fn1), and Sox4 was determined by quantitative RT-PCR in Py2T cell clones stably transfected with either HA-Sox4 (Clones 3 and 5) or control vector (Empty) in the absence of TGF-β. Fold changes are relative to those of cells transfected with the control vector.
(G) Immunoblotting analysis of E-cadherin, N-cadherin, fibronectin, Ezh2, and HA-Sox4 (HA) during TGF-β-induced EMT in Py2T cell clones either stably transfected with HA-Sox4 (Clones 3 and 5) or with empty control vector (Empty) in the absence of TGF-β. Vinculin was used as loading control.
(H) Cell migration of Py2T cell clones stably transfected with HA-Sox4 (Clones 3 and 5) or transfected with control vector (Empty) was analyzed in a transwell migration assay in the absence of TGF-β after 20 hr of cell seeding with 20% FBS as a chemoattractant.
Statistical values were calculated using an unpaired, two-tailed t test. *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. Error bars indicate the mean ± SD.

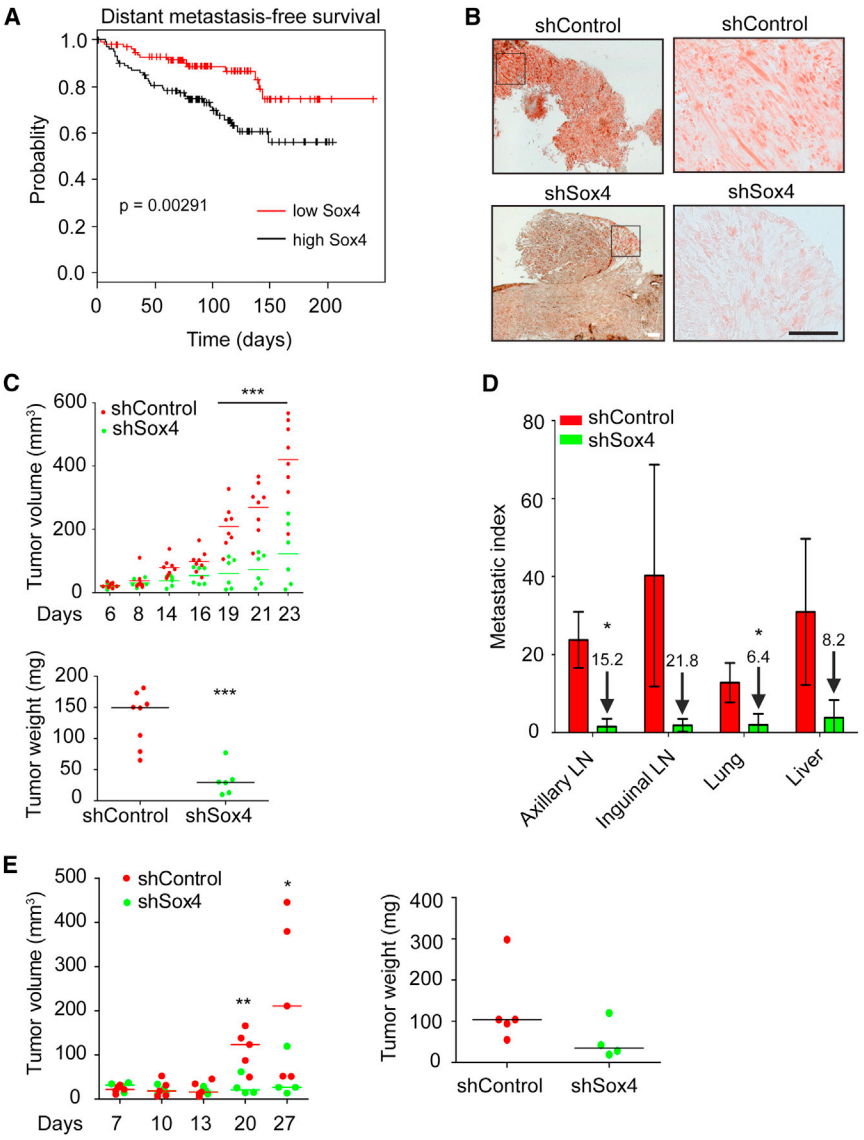

Figure 4. Sox4 Is Critical for Breast Carcinogenesis and Metastasis

(A) Correlation of the levels of Sox4 expression and prognosis in the "Schmidt" array data set of lymph-node-negative breast cancers (Schmidt et al., 2008). Shown is a Kaplan-Meier plot with survival curves for patient samples classified as having high Sox4 expression and those classified as having low Sox4 expression to assess metastasis-free survival.

(B) shSox4-expressing and shControl-expressing Py2T murine breast cancer cells were injected into the mammary fat pads of nude mice, and the resulting tumors were analyzed by immunohistochemistry for Sox4 expression. High magnification, as indicated by boxed areas, shows that shControl-transfected cells localize Sox4 primarily in the nucleus and exhibit spindle-like cell morphology, whereas shSox4-expressing cells lack significant Sox4 expression and exhibit a more differentiated morphology.

(C) shSox4-expressing and shControl-expressing Py2T cells were injected into the mammary fat pads of nude mice, and tumor growth in individual mice was measured over time (top). Tumor weight (bottom) was assessed after sacrificing the tumor-bearing mice 23 days after injection.

(D) Metastatic spread of firefly luciferase-expressing Py2T cells from the primary tumors analyzed in (C) was determined by measuring luciferase activity in extracts of lymph nodes (LN), lungs, and livers of transplanted mice. Luciferase activity levels in the various organs were divided by the primary tumor weights within the same mice to establish the metastatic index. Numbers indicate the fold differences between Sox4-expressing (shControl) and Sox4-depleted (shSox4) Py2T cells. Error bars indicate the mean ± SE.

(E) Primary tumor growth of shControl and shSox4 Py2T cells upon subcutaneous injection into nude mice (left). Tumor weight was assessed 27 days after implantation (right).

Statistical values were calculated using an unpaired, two-tailed t test. *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. Error bars indicate the mean ± SEM. See also Figure S2.

reduced upon Sox4 depletion (Figures S3J–S3N and S3P–S3T). Conversely, when the five repressors were transiently expressed in NMuMG and Py2T cells, the levels of Sox4 mRNA were not affected (Figures S3O and S3U). Together, the data indicate that Sox4 acts upstream of the EMT inducers during EMT, most likely in an indirect manner, since none of the EMT inducers have been detected as a direct transcriptional target of Sox4.

Interestingly, the Sox4-dependent gene expression profiling and ChIP experiments also revealed the promoter of the *Ezh2* gene to be directly bound and regulated by Sox4 in NMuMG and Py2T cells (Figures 5A and 5B). Quantitative RT-PCR and immunoblotting showed that Sox4 depletion led to significantly reduced Ezh2 expression and a global reduction in H3K27me3 levels in the presence of TGF-β in NMuMG and Py2T cells (Figure 5C). Moreover, the forced expression of Sox4 moderately increased Ezh2 mRNA and protein levels (Figures 3B, 3C, and 3G). Finally, *Ezh2* promoter activity was reduced upon Sox4 depletion in NMuMG cells in the presence of TGF-β, as determined by *Ezh2* promoter luciferase reporter assay (Figure 5D). These results indicate that Sox4 is required for *Ezh2* expression during EMT. However, it should be noted that *Ezh2* expression levels were already substantial in epithelial cells, where Sox4 expression was low and did not markedly change with the increasing levels of Sox4 during TGF-β-induced EMT (Figures 5E, 5F, and 6B). Hence, Sox4 is required for the efficient expression of Ezh2 during TGF-β-induced EMT, but not in epithelial cells in the absence of TGF-β, suggesting that factors other than Sox4 are critical for *Ezh2* gene expression in the epithelial state of the cells. Consistent with this notion, the dependence of *Ezh2* expression on Sox4 became already apparent between 2 and 8 hr of TGF-β stimulation in NMuMG and Py2T cells (Figures 5E and 5F).

Loss of Ezh2 Function Impairs EMT and Metastasis

Based on the pleiotropic functions of Ezh2 in mediating the H3K27me3 repressive mark, we speculated whether a major

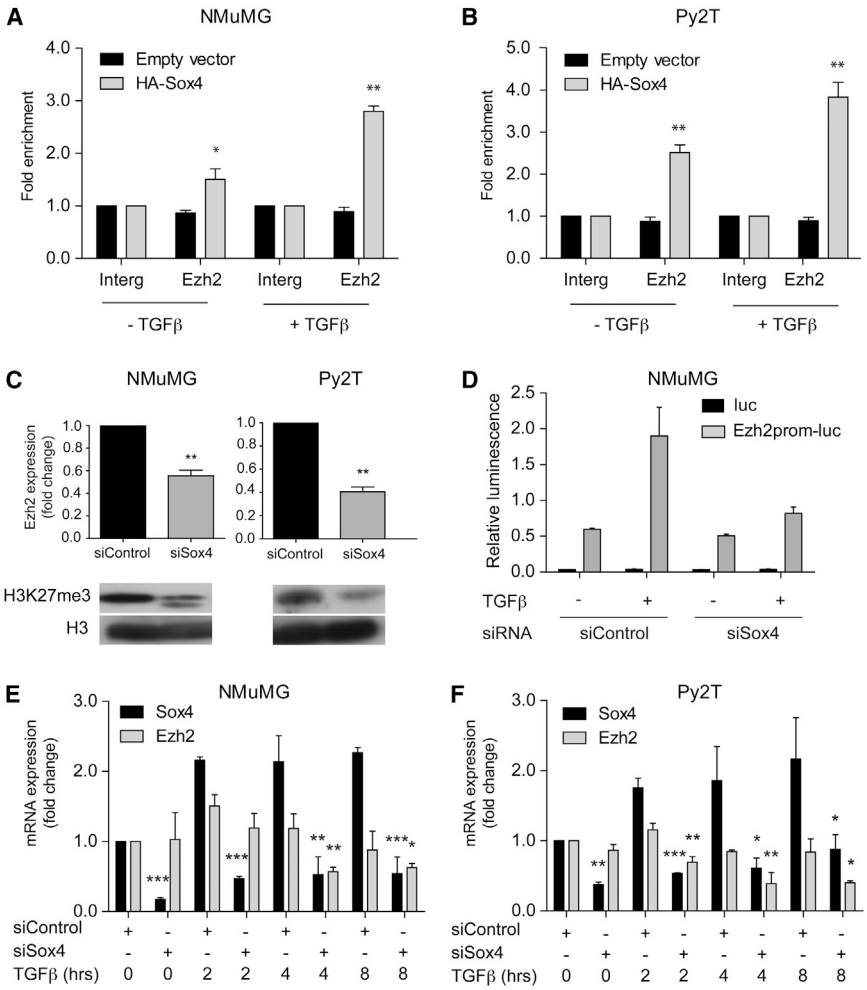

Figure 5. Sox4 Directly Regulates *Ezh2* Gene Expression

(A and B) NMuMG cells (A) and Py2T cells (B) were transiently transfected with a construct encoding HA-Sox4 or empty vector and used for ChIP in the absence and presence of TGF-β for 2 days using HA-specific antibodies. Precipitated DNA fragments corresponding to *Ezh2* promoter sequences were detected by quantitative PCR to assess Sox4 occupancy at this region. Fold changes are relative to those observed for a PCR reaction with primers specific for an intergenic region.

(C) Levels of Ezh2 mRNA were determined by quantitative RT-PCR in NMuMG cells or Py2T cells transfected with control siRNA (siControl) or siRNA against Sox4 (siSox4) in the presence of TGF-β for 2 days. Fold changes are relative to those in siControl-transfected cells (top). Global H3K27me3 levels were determined by immunoblotting in NMuMG cells or Py2T cells transfected with siControl or siSox4 and treated with TGF-β for 2 days. Total histone H3 was used as a loading control (bottom).

(D) Sox4 regulates Ezh2 promoter activity. NMuMG cells were transiently transfected with siControl or siSox4 and with a pGL3 luciferase reporter plasmid containing a 1.9 kb fragment of the Ezh2 promoter controlling the expression of *firefly* luciferase and then treated with or without TGF-β for 2 days. *Firefly* luciferase activity was normalized to cotransfected *renilla* luciferase activity (relative luminescence).

(E and F) Sox4 and Ezh2 expression levels were determined by quantitative RT-PCR after transient transfection of NMuMG (E) and Py2T cells (F) with siControl or siSox4 during the first hours of TGF-β treatment (0, 2, 4, and 8 hr).

Statistical values were calculated using a paired, two-tailed t test. *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. Error bars indicate the mean ± SD.

See also Figure S3 and Tables S1 and S2.

aspect of EMT regulation by Sox4 was dependent on its regulation of *Ezh2* gene expression. Indeed, similar to the ablation of Sox4, siRNA-mediated knockdown of Ezh2 (siEzh2) in NMuMG cells led to an efficient downregulation of Ezh2 and retention of the epithelial phenotype of NMuMG cells upon TGF-β stimulation (Figure 6A; Figure S4A). Immunoblotting and immunofluorescence microscopy analysis demonstrated a failure to lose E-cadherin and ZO-1 expression as well as a reduced gain in the expression of fibronectin and impaired formation of actin stress fibers and focal adhesions in Ezh2-depleted cells compared to siControl-treated cells (Figures 6B and 6C). These effects became even more apparent in NMuMG cells when both methyltransferases, Ezh1 and Ezh2, were concomitantly depleted (Figures 6A and 6C; Figures S4B–S4F). Similar to NMuMG cells, shRNA-mediated stable codepletion of Ezh1 and Ezh2 in Py2T cells also delayed EMT (Figures S4G–S4I). Comparable to the ablation of Sox4 expression, Ezh2-depleted NMuMG cells showed significantly increased apoptosis, a block in G0/G1, and a substantial attenuation in their proliferation during TGF-β-induced EMT (Figures 6D–6F). Finally, similar to the depletion of Sox4 in B16-F10 melanoma cells, shRNA-mediated ablation of Ezh2 or Ezh1/2 together resulted in a significantly reduced ability of these cells to form lung metastasis upon intravenous injection (Figure S4J). Notably, the concomitant high expression of Ezh2 and Sox4 significantly correlated with poor metastasis-free survival of early-stage, lymph-node-negative breast cancer patients (Schmidt et al., 2008; Figure 6G).

To directly assess whether *Ezh2* is the major transcriptional target of Sox4 required for its EMT-inducing function, we transiently ablated Sox4 in NMuMG and Py2T cells and concomitantly expressed Ezh2 by transient transfection of an Ezh2 expression plasmid. Notably, the forced expression of Ezh2 overcame the reduced expression of endogenous Ezh2 upon Sox4 depletion and restored TGF-β-induced EMT and changes in marker expression that otherwise were repressed by Sox4 ablation in NMuMG and in Py2T cells (Figures 7A–7C; Figures S5A and S5B). These results indicate that *Ezh2* is one critical target of Sox4 transcriptional regulation during EMT and metastasis formation.

Ezh2 Regulates EMT Genes via H3K27me3 Modification of Their Promoters

To identify the target genes that are modified by Ezh2-mediated H3K27me3 epigenetic imprint and regulated in their expression

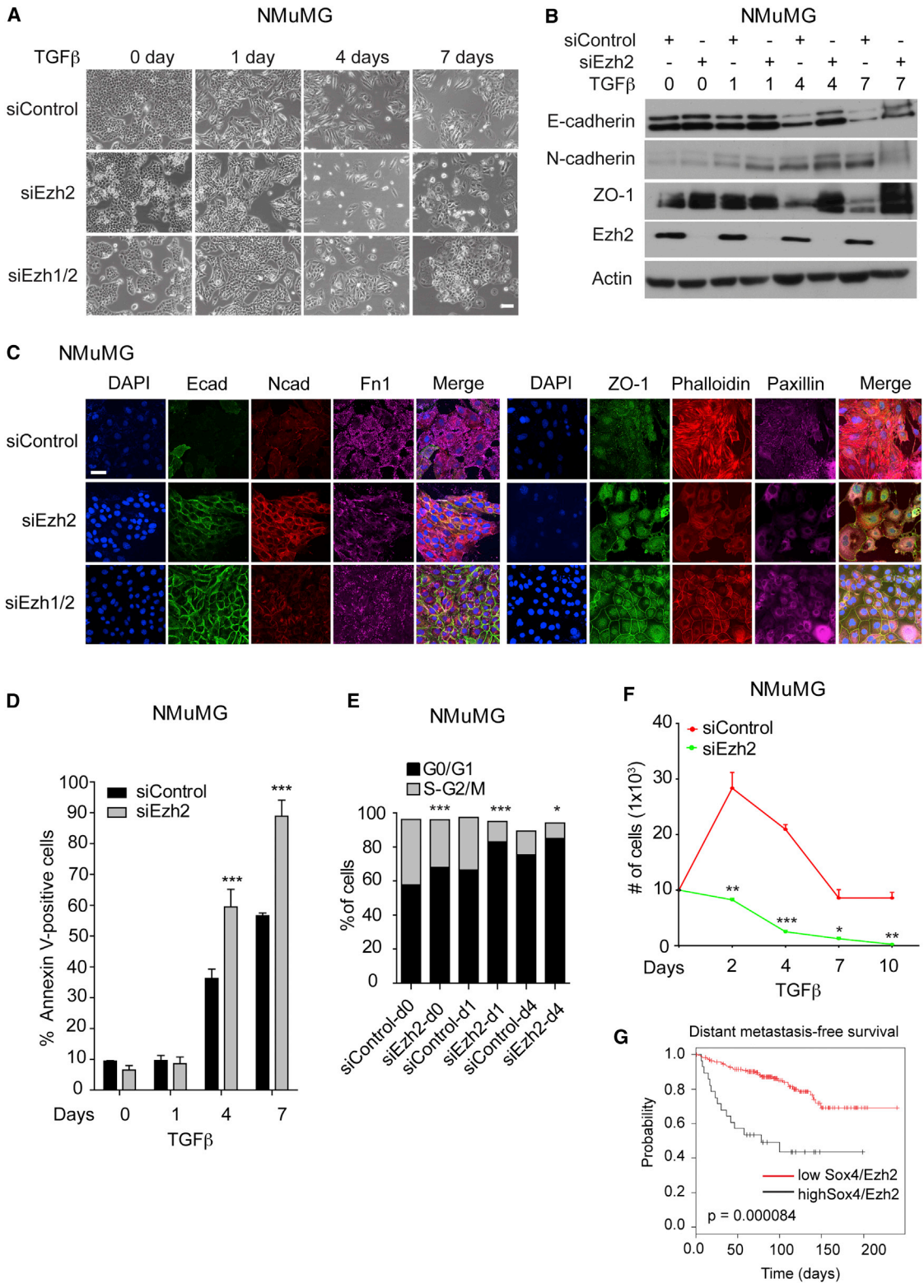

Figure 6. Depletion of Ezh2 Phenocopies the Ablation of Sox4 during TGF-β-Induced EMT

(A) NMuMG cells transfected with either control siRNA (siControl) or with siRNAs against Ezh2 or Ezh1 and Ezh2 together (siEzh1/2) were treated with TGF-β for the days indicated, and their morphology was evaluated by phase-contrast microscopy. Scale bar, 100 μm.

(B) Expression of E-cadherin, N-cadherin, ZO-1, and Ezh2 was determined by immunoblotting analysis during TGF-β-induced EMT in NMuMG cells transfected either with siEzh2 or siControl. Immunoblotting for actin was used as a loading control.

(legend continued on next page)

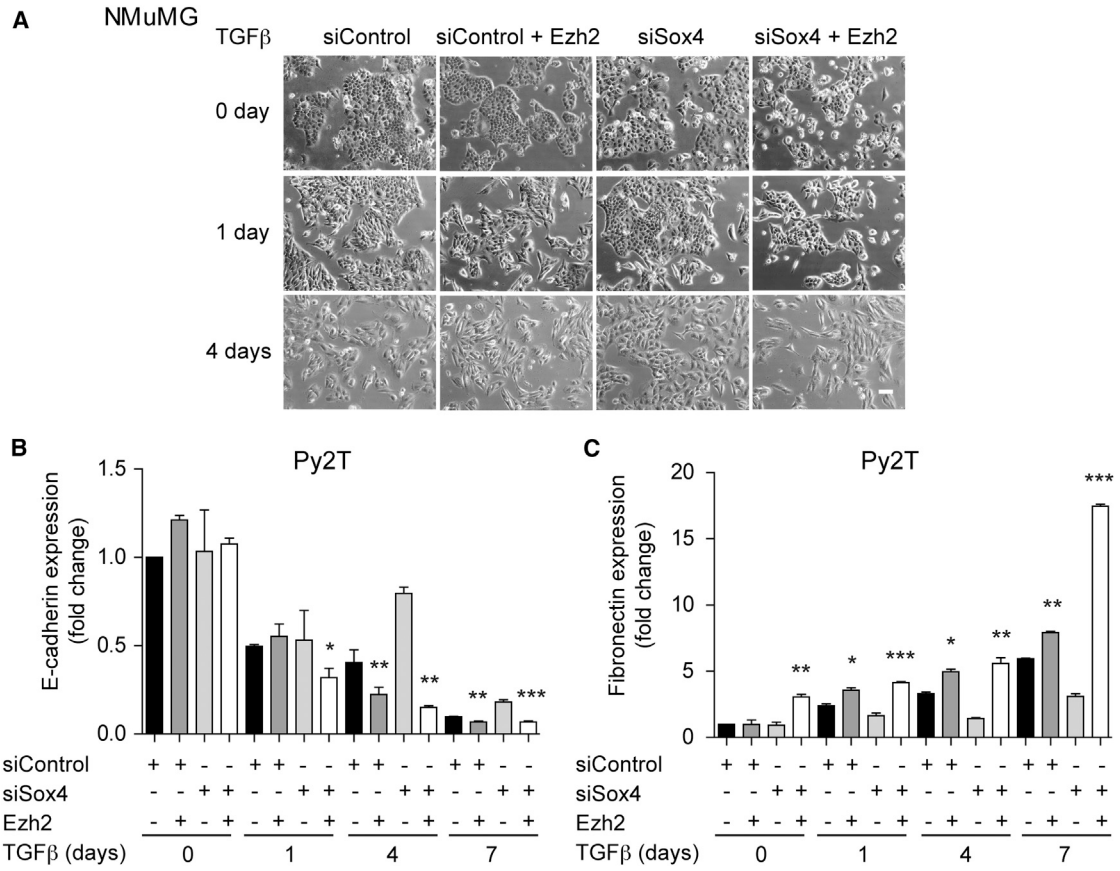

Figure 7. Forced Expression of Ezh2 Overcomes the Lack of Sox4 and Restores TGF-β-Induced EMT

(A) Although ablation of Sox4 expression (siSox4) repressed the TGF-β-induced mesenchymal morphology of NMuMG cells observed in siControl-transfected cells, transient expression of Ezh2 in Sox4-ablated cells (siSox4 + Ezh2) restored EMT and mesenchymal cell morphology. Cells were treated with TGF-β for the indicated times. Scale bar, 100 μm.

(B and C) mRNA expression levels of E-cadherin (C) and fibronectin (D) were determined after Ezh2 overexpression (+Ezh2) in control (shControl) and Sox4-ablated (shSox4) Py2T cells.

Statistical values were calculated using an unpaired, two-tailed t test. *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$.

See also Figure S5.

during EMT, we performed chromatin immunoprecipitation using an H3K27me3-specific antibody upon TGF-β treatment in NMuMG cells for 0, 1, 4, 7, 10, and 20 days followed by next-generation sequencing (ChIP-Seq) in combination with gene expression profiling. Whole-genome scanning revealed 970 regions corresponding to 301 genes with high variance in H3K27me3 levels during EMT and high reproducibility in the biological replicates (Figure 8A). Among the 301 genes, genome-wide gene expression analysis identified 46 genes that were transcriptionally upregulated during EMT and lost their H3K27me3 marks and three genes that were transcriptionally downregulated during EMT and gained H3K27me3 marks (Figure S6A). We validated the changes in H3K27me3 levels at target gene promoters such as *Mcam*, *Pdgfrb*, *Itga5*, *Col4a1*, and *St6galnac4* in NMuMG and Py2T cells (Figures 8B and 8C; Figure S6B) by ChIP followed by quantitative PCR for the promoter

(C) Immunoflorescence microscopy analysis of NMuMG cells transfected with either siControl or with siEzh2 or siEzh1/2. Cells were treated with TGF-β for 7 days and stained with antibodies against E-cadherin, ZO-1, and fibronectin, and with phalloidin to visualize the actin cytoskeleton. Scale bar, 50 μm.

(D) NMuMG cells transfected with either siControl or siEzh2 were treated with TGF-β for the days indicated, and the rates of apoptosis were determined by Annexin-V staining and flow cytometry.

(E) siControl- and siEzh2-transfected NMuMG cells were treated with TGF-β for the days indicated. Cells were stained with propidium iodide (PI), and the percentages of cells in G0/G1 and S-G2/M phases of the cell cycle were determined by flow cytometry.

(F) siRNA-mediated ablation of Ezh2 expression during TGF-β treatment of NMuMG cells results in a significant decrease in cell numbers compared to transfection with control siRNA.

(G) Kaplan-Meier survival curve for patient samples of the "Schmidt" data set of lymph-node-negative breast cancers (Schmidt et al., 2008) classified as having concomitant low Sox4 and Ezh2 expression and high Sox4 and Ezh2 expression, to assess metastasis-free survival.

Statistical values were calculated using an unpaired, two-tailed t test. *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. Error bars indicate the mean ± SD.

See also Figure S4.

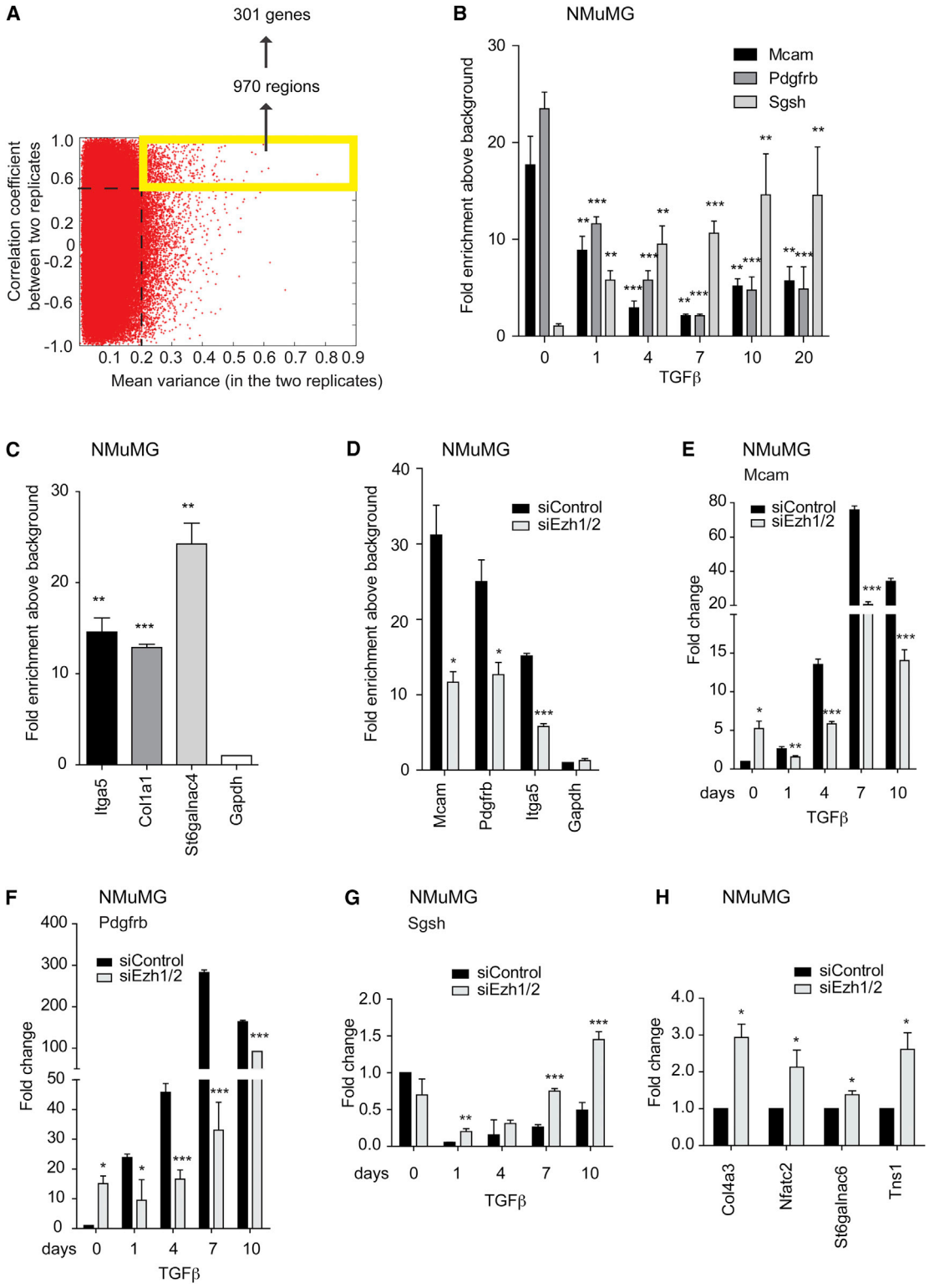

Figure 8. TGF-β-Induced EMT Accompanies Genome-wide Reprogramming of the Polycomb-Associated Mark H3K27me3

(A) TGF-β-induced EMT accompanies genome-wide reprogramming of the Polycomb-associated mark H3K27me3. Genome-wide H3K27me3 enrichment above background was calculated for 2-kb-wide windows overlapping by 1 kb. To extract those H3K27me3-regions that are most relevant for EMT dynamics, for each region both the variance in H3K27me3 levels during the progressive stages of EMT (horizontal axis) and the reproducibility of the H3K27me3 dynamics across biological replicates (vertical axis) were calculated, and 970 genes with high variance and high reproducibility were selected. Comparison of these regions with RefSeq transcripts show that they fall within 1 kb upstream and downstream of 301 distinct target genes.

(legend continued on next page)

regions of the genes. The expression of these target genes appeared highly dependent on the presence of Ezh1/2 and the H3K27me3 repressive imprint. For example, the promoters of the *Mcam*, *Pdgfrb*, and *Itag5* genes were found to be highly enriched in H3K27me3 marks in untreated NMuMG cells. Depletion of Ezh1/2 led to a loss of global H3K27me3 levels (Figure S6C) and of H3K27me3 at the promoters of the *Mcam*, *Pdgfrb*, and *Itag5* genes (Figure 8D; Figure S6D), with a concomitant increase in the expression of these genes (Figures 8E and 8F; Figures S6E and S6F) in NMuMG and Py2T cells. Conversely, the *Sgsh* gene gained H3K27me3 at its promoter during TGF-β-induced EMT, with a concomitant repression of its expression, and upon Ezh1/2 depletion, the lack of H3K27me3 marks on the *Sgsh* promoter during TGF-β-treatment resulted in increased *Sgsh* gene expression (Figure 8G). In addition, we validated a subset of other H3K27me3-enriched genes that changed in their expression profile upon depletion of Ezh1/2 (Figure 8H). Gene ontology analysis revealed a prominent enrichment for gene functions implicated in EMT, malignant tumor progression, and metastasis (Figure S6G). Furthermore, analysis of the "Minn" database of breast cancer (Minn et al., 2005) revealed that the 49 genes signature correlated with reduced bone-metastasis-free survival of patients in highly malignant breast cancers lacking estrogen receptor expression (ER−) or lacking estrogen receptor, progesterone receptor, and ErbB2 expression (triple-negative [TN]; Figures S6H–S6K). Moreover, in the "Schmidt" database of early-stage breast cancer (Schmidt et al., 2008), the 49 genes signature correlated with the probability to develop distant metastasis (Figure S6L). In line with these findings, transwell migration assays revealed significantly lower chemotactic migration of NMuMG and Py2T cells depleted of both Ezh1 and Ezh2 (Ezh1/2) as compared to siControl-transfected cells upon treatment with TGF-β (Figure S6M). Together, these data indicate that Sox4 regulates *Ezh2* gene expression, that Ezh2-mediated H3K27me3 marks are critically involved in the transcriptomic reprogramming of cells undergoing EMT, and that an epigenetic EMT gene signature correlates with survival in breast cancer patients.

DISCUSSION

EMT is accompanied by massive changes in cell morphology and behavior, and transcription factors play a pivotal role in controlling the various cellular functions during EMT, such as cell proliferation, cell survival, cell differentiation, cell migration, and cell adhesion. (Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and Weinberg, 2009; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006). In this study, the transcription factor Sox4 was found to be upregulated in its expression and transcriptional activity during EMT in a number of murine and human nontransformed mammary gland epithelial cells and breast cancer cells. The upregulated expression of Sox4 during TGF-β-induced EMT in NMuMG cells was independent of canonical TGF-β signaling, since shRNA-mediated ablation of Smad4 did not substantially affect Sox4 expression (data not shown). Moreover, treatment of NMuMG and Py2T cells with inhibitors against various signaling pathways revealed that the inhibition of TGF-β receptor and Notch signaling and the activation of Wnt signaling interfered with TGF-β-induced Sox4 expression, whereas the inhibition of EGF receptor increased Sox4 mRNA levels (data not shown). These observations suggest a complex gene regulatory network that drives the increased expression of Sox4 during the early phases of TGF-β-induced EMT, perspectives that warrant further investigation.

We show that Sox4 is required not only for the initiation of EMT but also for its maintenance. Furthermore, Sox4 is crucial for primary tumor growth and metastatic spread of Py2T breast cancer and B16-F10 melanoma cells. Finally, we find a significant positive correlation between increased Sox4 expression and the metastatic potential of early-stage, lymph-node-negative breast cancer in patients. These results are consistent with a recent report demonstrating that Sox4 is critical for EMT and tumor growth of human breast cancer cells and that Sox4 correlates with poor prognosis in cancer patients (Zhang et al., 2012).

Despite the recent implications of Sox4 in EMT, its direct transcriptional target genes have remained elusive. Here, we show that Sox4 exerts its central function in EMT and tumor progression by directly regulating the expression of a number of genes implicated in EMT, cell cycle regulation, cell survival, and cell migration. Notably, one of the direct targets of Sox4 transcriptional control is *Ezh2*, and ablation of Ezh2 function phenocopies the loss of Sox4 during TGF-β-induced EMT. Moreover, the forced expression of Ezh2 overcomes the failure of Sox4-deficient cells to undergo EMT. These results indicate that Ezh2 is a critical downstream effector of Sox4 and that Sox4 mediates epigenetic reprogramming by regulating the expression of epigenetic modulators during EMT. Consistent with this notion, the concomitant high expression of Sox4 and Ezh2 correlates with poor prognosis in a subset of breast cancer patients. These findings are indeed highly relevant, since they unravel a combinatorial

(B and C) ChIP was performed using an H3K27me3-specific antibody in NMuMG cells and the enrichment of H3K27me3 was determined by quantitative PCR at the promoters of the *Mcam*, *Pdgfrb*, and *Sgsh* genes during TGF-β-induced EMT (B) and at the promoters of the *Itga5*, *Col1a1*, and *St6galnac4* genes in the absence of TGF-β (C). Average enrichments from independent assays are plotted on the y axis as the ratio of precipitated DNA relative to the total input DNA and further normalized to the *Gapdh* housekeeping gene promoter.

(D) ChIP was performed using an H3K27me3-specific antibody in NMuMG cells transiently transfected with control siRNA (siControl) or siRNA against Ezh1/2 (siEzh1/2), and quantitative PCR was performed using promoter-specific primers as indicated to determine enrichment of H3K27me3 at the *Mcam*, *Pdgfrb*, and *Itga5* promoters and normalized to the *Gapdh* housekeeping gene promoter.

(E–G) NMuMG cells transiently transfected with control siRNA (siControl) or siRNA against Ezh1/2 (siEzh1/2) were analyzed by quantitative RT-PCR analysis for the mRNA expression levels of *Mcam* (E) and *Pdgfrb* (F), which already carry high H3K27me3 marks in untreated NMuMG cells and lose this mark during EMT, and of *Sgsh* (G), which gains this mark upon TGF-β treatment for 1, 4, 7, and 10 days.

(H) Validation of selected genes enriched in H3K27me3 marks and their transcriptional derepression after Ezh1/2 ablation in NMuMG cells in the absence of TGF-β.

Statistical values were calculated using a paired, two-tailed t test. *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$. Error bars indicate the mean ± SD.

See also Figure S6 and Tables S1 and S2.

action of key transcription factors and epigenetic regulators in driving the transcriptional reprogramming underlying cell-fate changes during morphogenic processes such as EMT.

The methyltransferase Ezh2 is the best-studied component of the Polycomb repressor complex 2 (PRC2). Increased expression or activity of Ezh2 is a marker of advanced and metastatic disease in many solid tumors, including prostate and breast cancer (Chase and Cross, 2011). Our genome-wide study of Ezh2-associated H3K27me3 marks during EMT revealed dynamic changes at a number of genes, of which a subset also show corresponding changes in transcriptional state. We further show that Ezh1 and Ezh2 together determine global as well as promoter-specific H3K27me3 levels. This activity modulates the transcriptional state of target genes, including a number of EMT-relevant genes, and it is a prerequisite for EMT. Notably, this epigenetic EMT signature significantly correlates with poor clinical outcome in patients with highly aggressive and metastatic ER− and triple-negative breast cancer. Thus, Polycomb-mediated transcriptional regulation is a critical contributor to the maintenance of epithelial differentiation and to transcriptomic reprogramming during EMT.

Previously, single-gene studies have suggested a role for Ezh2 in EMT regulation. For example, the PRC2 complex is recruited to the promoter of the *Cdh1* gene for its repression (Cao et al., 2008; Rhodes et al., 2003), and Snail1-mediated repression of *Cdh1* during stem cell differentiation is PRC2-dependent (Herranz et al., 2008). Notably, Ezh2 forms a corepressor complex with HDAC1, HDAC2, and Snail to repress *Cdh1* expression and to promote nasopharyngeal carcinoma malignancy (Tong et al., 2012). Moreover, the loss of integrin β4 expression during EMT correlates with a decrease in the activating histone modifications H3K9Ac and H3K4me3 and an increase in the repressive histone modification H3K27me3 at its promoter (Yang et al., 2009). Thus, previous single-gene studies have clearly demonstrated a role of PcG-mediated mechanisms in carcinogenesis (Mills, 2010; Sauvageau and Sauvageau, 2010; Sparmann and van Lohuizen, 2006). By identifying genome-wide targets of the PcG machinery during the stepwise progression of EMT and showing a direct role of Ezh1 and Ezh2 in the transcriptional regulation of key EMT genes, our study establishes a widespread and crucial role for PcG in EMT. We show that the transcription factor Sox4 acts upstream of epigenetic reprogramming events during EMT by directly regulating expression of the chromatin regulator Ezh2. Once expressed, Ezh2 and its accompanying chromatin regulators need to be recruited to specific target genes during EMT. It is surprising that we have not identified the EMT inducers Snail1/2, Twist1/2, and Zeb1/2 as direct targets of PcG-mediated regulation or Sox4 transcriptional control. Yet, given the critical role of Snail1/2, Twist1/2, and Zeb1/2 in regulating EMT and in associating with chromatin regulators, it will be important to further delineate the crosstalk of these proteins with the epigenetic machinery and to assess whether and how they are targeting epigenetic regulators to specific target genes.

Sox4 has been linked to the etiology of certain cancer types (Penzo-Méndez, 2010). Interestingly, the closely related family member Sox2 is one of the transcriptional factors required to reprogram somatic cells into induced pluripotent stem (iPS) cells (Takahashi and Yamanaka, 2006). In glioma-initiating cells, a crosstalk between Sox4 and Sox2 has been implicated in maintaining the tumorigenicity of the cells by Sox4 binding to the *Sox2* gene enhancer region (Ikushima et al., 2009). Together with previous reports that EMT increases the tumor-initiating (cancer stem cell) potential of cancer cells (Mani et al., 2008; Morel et al., 2008), our observation that Sox4 is required for EMT and for metastasis make Sox4 an exciting target for further investigation. Our experimental results indicate that Sox4 exerts its critical regulatory functions upstream of the Snail, Zeb, and Twist family transcriptional inducers of EMT, but without directly affecting their expression. The promoters of the known EMT regulators such as Snail1/2, Zeb1/2, or Twist1/2 also did not show any changes in H3K27me3 levels during EMT and were not affected following Ezh1/2 depletion (data not shown), suggesting that different regulatory pathways may underlie their expression dynamics during EMT. This suggests that independent regulatory pathways converge to control genes that drive EMT. These findings support our efforts to comprehensively understand the gene regulatory networks that underlie EMT. A recent report demonstrated that Sox9 cooperates with Snail2 to determine mammary stem cells and to promote malignant tumor progression (Guo et al., 2012), yet our study did not find a requirement of Sox9 for TGF-β-induced EMT in NMuMG and Py2T cells. In contrast, another Sox family member, Sox3, was shown to counteract the expression of the EMT-inducer Snail1, and conversely, Snail1 represses Sox3 gene expression during gastrulation and in human breast cancer cells (Acloque et al., 2011). We can only speculate that the differences between these findings may be based on the functional variety of Sox family members, which are classified into subgroups by the differences in their protein moieties outside of the HMG DNA binding domains and their varying biological functions (Chew and Gallo, 2009).

In summary, the data presented here provide critical insights into the regulatory networks by which one transcription factor, Sox4, regulates an important cell-fate determination event, namely, EMT. Notably, the observation that Sox4 regulates a number of EMT-relevant genes and *Ezh2*, and that Ezh2 modifies the expression of a number of genes known to be critical for EMT, exemplifies a critical interplay between transcriptional and epigenetic control during EMT. Our data suggest that the inhibition of Ezh2 function could be an attractive alternative for therapeutic intervention during malignant tumor progression. Indeed, early results with the Ezh2 inhibitor 3-deazaneplanocin (DZNep) are encouraging (Crea et al., 2011), and first results from clinical trials are impatiently awaited.

EXPERIMENTAL PROCEDURES

For more details see Supplemental Information.

Cell Lines

A subclone of NMuMG cells (NMuMG/E9, hereafter NMuMG) and MCF7-shControl and MCF7-shEcad have been described previously (Lehembre et al., 2008; Maeda et al., 2005). B16-F10 melanoma cells, EpRas, and MCF10A cells were commercially available. Py2T cells were derived from a breast tumor of MMTV-PyMT transgenic mice (Waldmeier et al., 2012).

Chromatin Immunoprecipitation

ChIP experiments were performed as previously described (Weber et al., 2005). In brief, crosslinked chromatin was sonicated to achieve an average fragment size of 500 bp. Starting with 100 μg of chromatin and 5 μg

of anti-HA antibody or anti-H3K27me3 antibody, 1 μl of ChIP material and 1 μl of input material were used for quantitative real-time PCR using specific primers covering the motif of Sox4 in the promoter of target genes or covering the 1000 bp promoter region from the transcription start site. Primers covering an intergenic region were used as a control. The efficiencies of PCR amplification were normalized to the PCR product of the intergenic region. Primer sequences are listed in the Supplemental Information.

Next-Generation Sequencing Analysis

The reads were mapped to the mouse genome assembly NCBI37/mm9 using the bowtie algorithm (Langmead et al., 2009), which allows up to two mismatches within the 36-bp-long sequence. Reads that mapped to more than 100 loci were discarded from the subsequent analysis. In total, 24–34 million reads were mapped to the genome in the first replicate and 95–106 million reads to the genome in the second replicate. The mean DNA fragment size was estimated from the mapping profile on both chromosome strands, and the reads were shifted by the half of the size toward the middle of the fragment. These data were exported to wiggle tracks for manual inspection in the UCSC genome browser and served as an input for the enrichment analysis. The genome was tiled with 2 kb windows overlapping by 1 kb, and each window was summarized by the numbers of reads in both replicates at every stage and the number of reads in the input DNA. The H3K27me3-enrichment was calculated assuming binomial statistics of the read counts,

$$z = \frac{\sum_{i=1}^{S} \frac{f_i}{S} - f_b}{\sqrt{\sum_{i=1}^{S} \frac{f_i \times (1 - f_i)}{S^2 N_i} + \frac{f_b \times (1 - f_b)}{N_b}}},$$

where $f_i = n_i / N_i$ is the frequency in the 2 kb window calculated from the number of reads in the window, n_i, and the total number of reads, N_i, in sample i, and $S = 12$ is the number of samples in both replicates. The background frequencies, f_b, are calculated analogously. Neighboring windows with $z \geq 5$ and $n_b \leq 50$ were merged and intersected with RefSeq loci extended by 1 kb upstream and 1 kb downstream. From the list of 96,224 such enriched regions a subset was selected by requiring dynamic (mean variance log-frequencies >0.2) and consistent (Pearson correlation coefficient of log frequencies >0.5) methylation changes. The resulting 970 regions lie within 1 kb of the loci of 301 genes.

Animal Experimentation

All studies involving mice have been approved by the Swiss Federal Veterinary Office (SFVO) and the regulations of the Cantonal Veterinary Office of Basel Stadt (licenses 1878, 1907, and 1908).

Statistical Analysis

Statistical analysis and graphs were generated using the GraphPad Prism software (GraphPad Software, San Diego, CA). All statistical analysis was done by unpaired/paired, two-sided t test.

ACCESSION NUMBERS

Gene expression data for Sox4 knockdown cells in the presence and absence of TGF-β in NMuMG cells and ChIP-seq data of H3K27me3 imprints during TGF-β-induced EMT in NMuMG cells are deposited in the Gene Expression Omnibus (GSE44050 and GSE45579, respectively).

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, two tables, and Supplemental Experimental Procedures can be found with this article online at http://dx.doi.org/10.1016/j.ccr.2013.04.020.

ACKNOWLEDGMENTS

We are grateful to H. Antoniadis, I. Galm, and P. Schmidt for technical support and A. Fantozzi for critical comments. We thank I. Nissen (Laboratory of Quantitative Genomics at D-BSSE, ETH Zürich) and S. Dessus-Babus (FMI, Basel) for Illumina sequencing. We thank A.P. Tsou (National Yang-Ming University, Taipei), A.H. Peters (FMI, Basel), N. Goldfinger and V. Rotter (Weizmann Institute, Rehovot), and G. Berx (VIB, Ghent) for providing important reagents and P. Demougin for gene expression profiling. This research has been supported by the Swiss National Science Foundation, the Swiss Initiative for Systems Biology (RTD Cellplasticity), the EU-FP7 framework program TuMIC 2008-201662, and the Swiss Cancer League. Research in the laboratory of V.K.T. is supported by Deutsche Forschungsgemeinschaft (DFG) grant TI 799/1-1, EpiGeneSys RISE1 program, Marie Curie CIG 322210, and Wilhelm Sander Stiftung 2012.009.1. Research in the laboratory of D.S. is supported by the Novartis Research Foundation, the European Union (NoE ''EpiGeneSys'' FP7- HEALTH-2010-257082 and the ''Blueprint'' consortium FP7-282510), the European Research Council (EpiGePlas), and the SNF Sinergia program.

Received: January 2, 2012
Revised: August 20, 2012
Accepted: April 22, 2013
Published: June 10, 2013

REFERENCES

Aaboe, M., Birkenkamp-Demtroder, K., Wiuf, C., Sørensen, F.B., Thykjaer, T., Sauter, G., Jensen, K.M., Dyrskjøt, L., and Ørntoft, T. (2006). SOX4 expression in bladder carcinoma: clinical aspects and in vitro functional characterization. Cancer Res. *66*, 3434–3442.

Acloque, H., Ocaña, O.H., Matheu, A., Rizzoti, K., Wise, C., Lovell-Badge, R., and Nieto, M.A. (2011). Reciprocal repression between Sox3 and snail transcription factors defines embryonic territories at gastrulation. Dev. Cell *21*, 546–558.

Ahn, S.G., Kim, H.S., Jeong, S.W., Kim, B.E., Rhim, H., Shim, J.Y., Kim, J.W., Lee, J.H., and Kim, I.K. (2002). Sox-4 is a positive regulator of Hep3B and HepG2 cells' apoptosis induced by prostaglandin (PG)A(2) and Δ12-PGJ(2). Exp. Mol. Med. *34*, 243–249.

Brabletz, T., Jung, A., Spaderna, S., Hlubek, F., and Kirchner, T. (2005). Opinion: migrating cancer stem cells—an integrated concept of malignant tumour progression. Nat. Rev. Cancer *5*, 744–749.

Cao, Q., Yu, J., Dhanasekaran, S.M., Kim, J.H., Mani, R.S., Tomlins, S.A., Mehra, R., Laxman, B., Cao, X., Yu, J., et al. (2008). Repression of E-cadherin by the polycomb group protein EZH2 in cancer. Oncogene *27*, 7274–7284.

Chase, A., and Cross, N.C. (2011). Aberrations of EZH2 in cancer. Clin. Cancer Res. *17*, 2613–2618.

Chen, H., Tu, S.W., and Hsieh, J.T. (2005). Down-regulation of human DAB2IP gene expression mediated by polycomb Ezh2 complex and histone deacetylase in prostate cancer. J. Biol. Chem. *280*, 22437–22444.

Cheung, M., Abu-Elmagd, M., Clevers, H., and Scotting, P.J. (2000). Roles of Sox4 in central nervous system development. Brain Res. Mol. Brain Res. *79*, 180–191.

Chew, L.J., and Gallo, V. (2009). The Yin and Yang of Sox proteins: Activation and repression in development and disease. J. Neurosci. Res. *87*, 3277–3287.

Christofori, G. (2006). New signals from the invasive front. Nature *441*, 444–450.

Crea, F., Hurt, E.M., Mathews, L.A., Cabarcas, S.M., Sun, L., Marquez, V.E., Danesi, R., and Farrar, W.L. (2011). Pharmacologic disruption of Polycomb Repressive Complex 2 inhibits tumorigenicity and tumor progression in prostate cancer. Mol. Cancer *10*, 40.

Ezhkova, E., Pasolli, H.A., Parker, J.S., Stokes, N., Su, I.H., Hannon, G., Tarakhovsky, A., and Fuchs, E. (2009). Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. Cell *136*, 1122–1135.

Grünert, S., Jechlinger, M., and Beug, H. (2003). Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. Nat. Rev. Mol. Cell Biol. *4*, 657–665.

Guo, W., Keckesova, Z., Donaher, J.L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zürrer-Härdi, U., Bell, G., et al. (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. Cell *148*, 1015–1028.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Herranz, N., Pasini, D., Díaz, V.M., Francí, C., Gutierrez, A., Dave, N., Escrivà, M., Hernandez-Muñoz, I., Di Croce, L., Helin, K., et al. (2008). Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. Mol. Cell. Biol. *28*, 4772–4781.

Hong, C.S., and Saint-Jeannet, J.P. (2005). Sox proteins and neural crest development. Semin. Cell Dev. Biol. *16*, 694–703.

Huber, M.A., Kraut, N., and Beug, H. (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr. Opin. Cell Biol. *17*, 548–558.

Hur, W., Rhim, H., Jung, C.K., Kim, J.D., Bae, S.H., Jang, J.W., Yang, J.M., Oh, S.T., Kim, D.G., Wang, H.J., et al. (2010). SOX4 overexpression regulates the p53-mediated apoptosis in hepatocellular carcinoma: clinical implication and functional analysis in vitro. Carcinogenesis *31*, 1298–1307.

Ikushima, H., Todo, T., Ino, Y., Takahashi, M., Miyazawa, K., and Miyazono, K. (2009). Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. Cell Stem Cell *5*, 504–514.

Kalluri, R. (2009). EMT: when epithelial cells decide to become mesenchymal-like cells. J. Clin. Invest. *119*, 1417–1419.

Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. J. Clin. Invest. *119*, 1420–1428.

Kang, Y., and Massagué, J. (2004). Epithelial-mesenchymal transitions: twist in development and metastasis. Cell *118*, 277–279.

Kleer, C.G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S.A., Ghosh, D., Sewalt, R.G., Otte, A.P., Hayes, D.F., et al. (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proc. Natl. Acad. Sci. USA *100*, 11606–11611.

Kloeker, S., Major, M.B., Calderwood, D.A., Ginsberg, M.H., Jones, D.A., and Beckerle, M.C. (2004). The Kindler syndrome protein is regulated by transforming growth factor-β and involved in integrin-mediated adhesion. J. Biol. Chem. *279*, 6824–6833.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lehembre, F., Yilmaz, M., Wicki, A., Schomber, T., Strittmatter, K., Ziegler, D., Kren, A., Went, P., Derksen, P.W., Berns, A., et al. (2008). NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin. EMBO J. *27*, 2603–2615.

Liao, Y.L., Sun, Y.M., Chau, G.Y., Chau, Y.P., Lai, T.C., Wang, J.L., Horng, J.T., Hsiao, M., and Tsou, A.P. (2008). Identification of SOX4 target genes using phylogenetic footprinting-based prediction from expression microarrays suggests that overexpression of SOX4 potentiates metastasis in hepatocellular carcinoma. Oncogene *27*, 5578–5589.

Lioubinski, O., Müller, M., Wegner, M., and Sander, M. (2003). Expression of Sox transcription factors in the developing mouse pancreas. Dev. Dyn. *227*, 402–408.

Liu, P., Ramachandran, S., Ali Seyed, M., Scharer, C.D., Laycock, N., Dalton, W.B., Williams, H., Karanam, S., Datta, M.W., Jaye, D.L., and Moreno, C.S. (2006). Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells. Cancer Res. *66*, 4011–4019.

Maeda, M., Johnson, K.R., and Wheelock, M.J. (2005). Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition. J. Cell Sci. *118*, 873–887.

Mani, S.A., Guo, W., Liao, M.J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell *133*, 704–715.

Mills, A.A. (2010). Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. Nat. Rev. Cancer *10*, 669–682.

Min, J., Zaslavsky, A., Fedele, G., McLaughlin, S.K., Reczek, E.E., De Raedt, T., Guney, I., Strochlic, D.E., Macconaill, L.E., Beroukhim, R., et al. (2010). An oncogene-tumor suppressor cascade drives metastatic prostate cancer by coordinately activating Ras and nuclear factor-κB. Nat. Med. *16*, 286–294.

Minn, A.J., Gupta, G.P., Siegel, P.M., Bos, P.D., Shu, W., Giri, D.D., Viale, A., Olshen, A.B., Gerald, W.L., and Massagué, J. (2005). Genes that mediate breast cancer metastasis to lung. Nature *436*, 518–524.

Morel, A.P., Lièvre, M., Thomas, C., Hinkal, G., Ansieau, S., and Puisieux, A. (2008). Generation of breast cancer stem cells through epithelial-mesenchymal transition. PLoS ONE *3*, e2888.

Nieto, M.A. (2010). The ins and outs of the epithelial to mesenchymal transition in health and disease. Annu. Rev. Cell Dev. Biol. *27*, 347–376.

Penzo-Méndez, A.I. (2010). Critical roles for SoxC transcription factors in development and cancer. Int. J. Biochem. Cell Biol. *42*, 425–428.

Polyak, K., and Weinberg, R.A. (2009). Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat. Rev. Cancer *9*, 265–273.

Pramoonjago, P., Baras, A.S., and Moskaluk, C.A. (2006). Knockdown of Sox4 expression by RNAi induces apoptosis in ACC3 cells. Oncogene *25*, 5626–5639.

Restivo, A., Piacentini, G., Placidi, S., Saffirio, C., and Marino, B. (2006). Cardiac outflow tract: a review of some embryogenetic aspects of the conotruncal region of the heart. Anat. Rec. A Discov. Mol. Cell. Evol. Biol. *288*, 936–943.

Rhodes, D.R., Sanda, M.G., Otte, A.P., Chinnaiyan, A.M., and Rubin, M.A. (2003). Multiplex biomarker approach for determining risk of prostate-specific antigen-defined recurrence of prostate cancer. J. Natl. Cancer Inst. *95*, 661–668.

Sauvageau, M., and Sauvageau, G. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell *7*, 299–313.

Scharer, C.D., McCabe, C.D., Ali-Seyed, M., Berger, M.F., Bulyk, M.L., and Moreno, C.S. (2009). Genome-wide promoter analysis of the SOX4 transcriptional network in prostate cancer cells. Cancer Res. *69*, 709–717.

Schilham, M.W., Oosterwegel, M.A., Moerer, P., Ya, J., de Boer, P.A., van de Wetering, M., Verbeek, S., Lamers, W.H., Kruisbeek, A.M., Cumano, A., and Clevers, H. (1996). Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. Nature *380*, 711–714.

Schilham, M.W., Moerer, P., Cumano, A., and Clevers, H.C. (1997). Sox-4 facilitates thymocyte differentiation. Eur. J. Immunol. *27*, 1292–1295.

Schmidt, M., Böhm, D., von Törne, C., Steiner, E., Puhl, A., Pilch, H., Lehr, H.A., Hengstler, J.G., Kölbl, H., and Gehrmann, M. (2008). The humoral immune system has a key prognostic impact in node-negative breast cancer. Cancer Res. *68*, 5405–5413.

Shen, R., Pan, S., Qi, S., Lin, X., and Cheng, S. (2010). Epigenetic repression of microRNA-129-2 leads to overexpression of SOX4 in gastric cancer. Biochem. Biophys. Res. Commun. *394*, 1047–1052.

Shin, M.S., Fredrickson, T.N., Hartley, J.W., Suzuki, T., Akagi, K., and Morse, H.C., 3rd. (2004). High-throughput retroviral tagging for identification of genes involved in initiation and progression of mouse splenic marginal zone lymphomas. Cancer Res. *64*, 4419–4427.

Sinner, D., Kordich, J.J., Spence, J.R., Opoka, R., Rankin, S., Lin, S.C., Jonatan, D., Zorn, A.M., and Wells, J.M. (2007). Sox17 and Sox4 differentially regulate β-catenin/T-cell factor activity and proliferation of colon carcinoma cells. Mol. Cell. Biol. *27*, 7802–7815.

Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. Nat. Rev. Cancer *6*, 846–856.

Su, I.H., Basavaraj, A., Krutchinsky, A.N., Hobert, O., Ullrich, A., Chait, B.T., and Tarakhovsky, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. Nat. Immunol. *4*, 124–131.

Suzuki, H., Forrest, A.R., van Nimwegen, E., Daub, C.O., Balwierz, P.J., Irvine, K.M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M.J., et al.; FANTOM Consortium; Riken Omics Science Center. (2009). The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat. Genet. *41*, 553–562.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Tavazoie, S.F., Alarcón, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P.D., Gerald, W.L., and Massagué, J. (2008). Endogenous human microRNAs that suppress breast cancer metastasis. Nature *451*, 147–152.

Thiery, J.P., and Morgan, M. (2004). Breast cancer progression with a Twist. Nat. Med. *10*, 777–778.

Thiery, J.P., and Sleeman, J.P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. Nat. Rev. Mol. Cell Biol. *7*, 131–142.

Tong, Z.T., Cai, M.Y., Wang, X.G., Kong, L.L., Mai, S.J., Liu, Y.H., Zhang, H.B., Liao, Y.J., Zheng, F., Zhu, W., et al. (2012). EZH2 supports nasopharyngeal carcinoma cell aggressiveness by forming a co-repressor complex with HDAC1/HDAC2 and Snail to inhibit E-cadherin. Oncogene *31*, 583–594.

van de Wetering, M., Oosterwegel, M., van Norren, K., and Clevers, H. (1993). Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. EMBO J. *12*, 3847–3854.

Varambally, S., Cao, Q., Mani, R.S., Shankar, S., Wang, X., Ateeq, B., Laxman, B., Cao, X., Jing, X., Ramnarayanan, K., et al. (2008). Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science *322*, 1695–1699.

Waldmeier, L., Meyer-Schaller, N., Diepenbruck, M., and Christofori, G. (2012). Py2T murine breast cancer cells, a versatile model of TGFβ-induced EMT in vitro and in vivo. PLoS ONE *7*, e48651.

Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L., and Schübeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat. Genet. *37*, 853–862.

Wilson, M.E., Yang, K.Y., Kalousova, A., Lau, J., Kosaka, Y., Lynn, F.C., Wang, J., Mrejen, C., Episkopou, V., Clevers, H.C., and German, M.S. (2005). The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. Diabetes *54*, 3402–3409.

Yang, X., Pursell, B., Lu, S., Chang, T.K., and Mercurio, A.M. (2009). Regulation of β4-integrin expression by epigenetic modifications in the mammary gland and during the epithelial-to-mesenchymal transition. J. Cell Sci. *122*, 2473–2480.

Zhang, J., Liang, Q., Lei, Y., Yao, M., Li, L., Gao, X., Feng, J., Zhang, Y., Gao, H., Liu, D.X., et al. (2012). SOX4 induces epithelial-mesenchymal transition and contributes to breast cancer progression. Cancer Res. *72*, 4597–4608.