

ZEB1 drives epithelial-to-mesenchymal transition in lung cancer

Jill E. Larsen,^{1,2,3} Vaishnavi Nathan,² Jihan K. Osborne,⁴ Rebecca K. Farrow,² Dhruva Deb,¹ James P. Sullivan,¹ Patrick D. Dospoy,¹ Alexander Augustyn,¹ Suzie K. Hight,¹ Mitsuo Sato,⁵ Luc Girard,^{1,4} Carmen Behrens,⁶ Ignacio I. Wistuba,⁶ Adi F. Gazdar,^{1,7} Nicholas K. Hayward,² and John D. Minna^{1,4,8}

¹Hamon Center for Therapeutic Oncology Research and the Simmons Comprehensive Cancer Center, The University of Texas Southwestern Medical Center, Dallas, Texas, USA. ²QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ³School of Medicine, The University of Queensland, Brisbane, Queensland, Australia. ⁴Department of Pharmacology, The University of Texas Southwestern Medical Center, Dallas, Texas, USA. ⁵Department of Respiratory Medicine, Nagoya University Graduate School of Medicine, Nagoya, Japan. ⁶Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA. ⁷Department of Pathology, and ⁸Department of Internal Medicine, The University of Texas Southwestern Medical Center, Dallas, Texas, USA.

Increased expression of zinc finger E-box binding homeobox 1 (ZEB1) is associated with tumor grade and metastasis in lung cancer, likely due to its role as a transcription factor in epithelial-to-mesenchymal transition (EMT). Here, we modeled malignant transformation in human bronchial epithelial cells (HBECs) and determined that EMT and ZEB1 expression are early, critical events in lung cancer pathogenesis. Specific oncogenic mutations in TP53 and KRAS were required for HBECs to engage EMT machinery in response to microenvironmental (serum/TGF- β) or oncogenetic (MYC) factors. Both TGF- β - and MYC-induced EMT required ZEB1, but engaged distinct TGF- β -dependent and vitamin D receptor-dependent (VDR-dependent) pathways, respectively. Functionally, we found that ZEB1 causally promotes malignant progression of HBECs and tumorigenicity, invasion, and metastases in non-small cell lung cancer (NSCLC) lines. Mechanistically, ZEB1 expression in HBECs directly repressed epithelial splicing regulatory protein 1 (ESRP1), leading to increased expression of a mesenchymal splice variant of CD44 and a more invasive phenotype. In addition, ZEB1 expression in early stage IB primary NSCLC correlated with tumor-node-metastasis stage. These findings indicate that ZEB1-induced EMT and associated molecular changes in ESRP1 and CD44 contribute to early pathogenesis and metastatic potential in established lung cancer. Moreover, TGF- β and VDR signaling and CD44 splicing pathways associated with ZEB1 are potential EMT chemoprevention and therapeutic targets in NSCLC.

Introduction

Human lung cancer is the most common cause of cancer-related death worldwide (1). It develops through a multistep process, usually after prolonged smoke-related tobacco exposure resulting in oncogenic mutations in lung epithelial cells (2). We have previously modeled the step-wise progression of lung cancer pathogenesis in vitro by introducing common lung cancer “driver” mutations into CDK4/TERT-immortalized human bronchial epithelial cells (HBECs) and progressing them to full malignancy (3, 4). While immortalized HBECs (CDK4 and TERT only) do not exhibit any in vitro or in vivo transformation, loss of the tumor-suppressor p53 and overexpression of mutant KRAS^{V12} results in partial, but not full, oncogenic transformation. Full transformation, defined as growth of tumor xenografts in immunocompromised mice, occurs with the subsequent addition of MYC overexpression or growth in serum-containing media (4).

Interestingly, both the genetic-induced (MYC overexpression) and microenvironmental-induced (growth in serum-containing media) transformation of sh-p53+KRAS^{V12}-manipulated HBECs resulted in an epithelial-to-mesenchymal transition (EMT). EMT

and the reverse process of mesenchymal-to-epithelial transition (MET) are critical developmental processes (5). EMT, however, can be reactivated in cancer where it promotes tumorigenic progression of epithelial cells, such as increasing migration and invasion, “stemness,” and inhibiting apoptosis and senescence (5, 6). A hallmark of EMT is the functional loss of adherens junction protein E-cadherin, resulting in loss of cell polarity and tissue organization. Additionally, it is regulated by several factors, including important EMT-transcription factors (EMT-TFs) comprising the ZEB, Snail, and Twist families (5). TGF- β signaling can act as a key inducer of EMT through, in part, its regulation of EMT-TFs (5). While functioning as a tumor suppressor in normal cells and early stage cancers, TGF- β can serve as a tumor promoter in later-stage cancers. These divergent tumor-suppressive and tumor-progressing roles are termed the “TGF- β paradox” (7).

The EMT-TF zinc finger/homeodomain proteins ZEB1 and ZEB2 can act as both transcriptional activators (by binding to histone acetyl-transferases p300/pCAF) and repressors (by binding to CtBP corepressors, histone acetyl-transferase TIP60, chromatin remodeling ATPase BRG1, and histone deacetylase SIRT1) (6). ZEB proteins inhibit epithelial differentiation, in part, by repressing miRNA-200 (miR-200) family members (*miR-200a*, *miR-200b*, *miR-200c*, and *miR-141*). The miR-200 family in turn can repress ZEB proteins to form an important negative feedback

Conflict of interest: The authors have declared that no conflict of interest exists.

Submitted: April 22, 2014; **Accepted:** June 9, 2016.

Reference information: *J Clin Invest*. 2016;126(9):3219–3235. doi:10.1172/JCI67625.

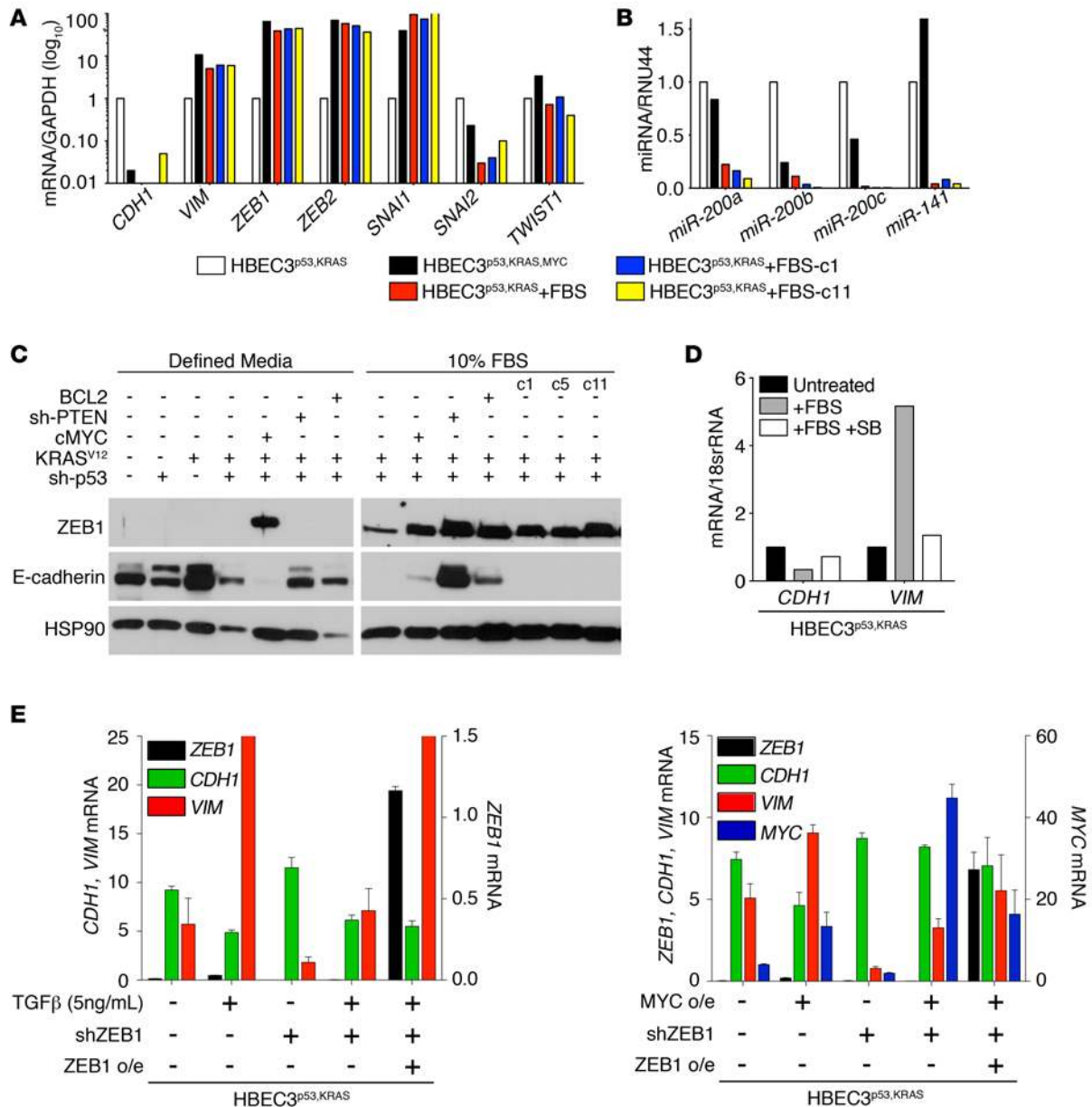


Figure 1. Genetic (MYC) and microenvironmental (TGF-β) oncogenic cues can induce ZEB1-dependent EMT and increase transformation in partially transformed HBEC3^{p53,KRAS} cells, but not in nontransformed HBECs. (A and B) Expression of (A) EMT markers and EMT-TFs and (B) miR-200 family members in HBEC3^{p53,KRAS}, HBEC3^{p53,KRAS,MYC}, and HBEC3^{p53,KRAS,+FBS} (whole population and 2 clonal populations, c1 and c11) (mean ± SD). (C) Immunoblot of ZEB1 and E-cadherin across oncogenically manipulated HBECs grown in defined (serum-free) or 10% FBS media with 3 clonal populations of HBEC3^{p53,KRAS,+FBS}: c1, c5, and c11. (D) mRNA expression of EMT markers in HBEC3^{p53,KRAS} following 72-hour treatment with 10% FBS with or without 10 μM SB431542 (mean ± SD). (E) mRNA expression of EMT markers in HBEC3^{p53,KRAS} and HBEC3^{p53,KRAS,shZEB1} following 5 ng/ml TGF-β treatment (left) or MYC overexpression (o/e, right), followed with ZEB1 overexpression rescue (mean ± SD). Data are representative of at least 3 independent experiments.

loop regulating epithelial differentiation (8). ZEB1 and ZEB2 overexpression have been found in several human cancers, including ZEB1 in non-small cell lung cancer (NSCLC) (9–11). Increased ZEB1 has been associated with tumor grade in lung adenocarcinomas (9) and lymphatic or distant metastasis in lung squamous cell carcinomas (12). Similarly, ZEB1 can also promote colorectal and breast cancer metastasis (13). In lung cancer cell lines, ZEB1 is inversely correlated with expression of E-cadherin (10, 14, 15) and promotes anchorage-independent colony formation (15).

EMT and metastasis are generally considered late events in tumorigenesis, where acquiring a mesenchymal phenotype allows a malignant epithelial cell to detach from the primary tumor. However, EMT and early metastatic processes have also been shown to occur in early “preinvasive” stages of pancreatic adenocarcinoma (16). Thus, activation of an EMT program with its associated potential for metastasis may be a much earlier event in cancer progression than previously thought. In this study, using an in vitro model of pre- and early neoplastic HBECs, we found ZEB1-induced EMT

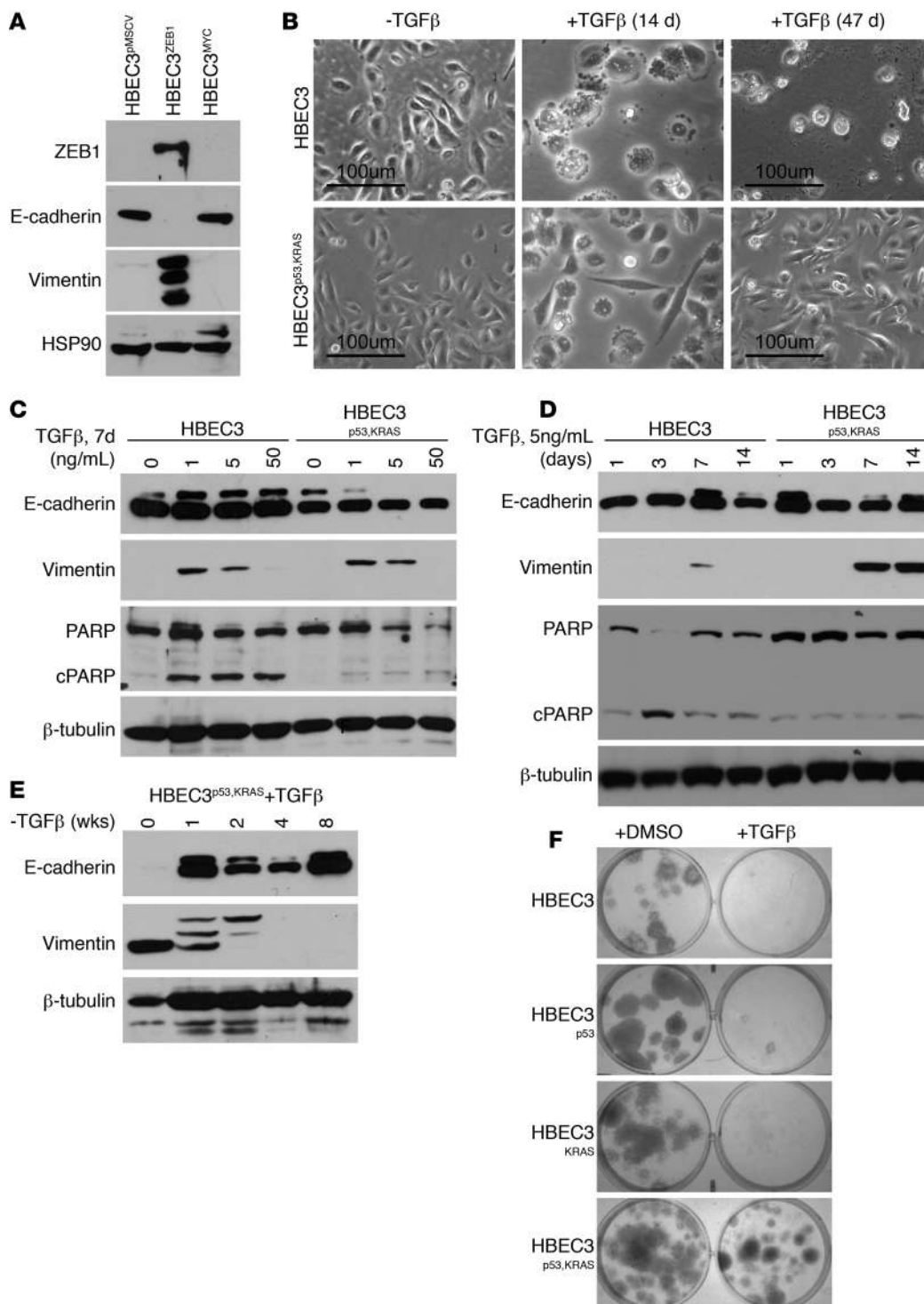


Figure 2. Genetic (MYC) and microenvironmental (TGF-β) induction of EMT in HBECs is cell context dependent and illustrates the paradoxical role of TGF-β in tumorigenesis. (A) Immunoblot of EMT markers in HBEC3 (grown in serum-free media) with and without MYC or ZEB1 overexpression, grown in serum-free media (pMSCV, vector control). (B) Phase micrographs of HBEC3 and HBEC3^{p53,KRAS} following TGF-β treatment. Scale bars 100 μm. (C and D) Immunoblots showing acute dose-dependent (C) and time-dependent (D) response of HBEC3 and HBEC3^{p53,KRAS} following TGF-β treatment. (E) Immunoblot of EMT markers in long-term TGF-β-treated HBEC3^{p53,KRAS} (45 days, 5 ng/ml TGF-β) following cessation of TGF-β treatment. (F) Anchorage-dependent (liquid) colony forming ability showing acute response of HBEC3, HBEC3^{p53}, HBEC3^{KRAS}, and HBEC3^{p53,KRAS} to TGF-β treatment. HSP90 and β-tubulin were used as loading controls. Data are representative of at least 3 independent experiments.

to be a critical early event in the progression of HBECs to malignancy. Importantly, we found that engagement of EMT-activating machinery in response to genetic (such as MYC) and microenvironmental stimuli (such as TGF-β) requires bronchial epithelial cells to harbor premalignant oncogenic mutations (such as in *TP53* and *KRAS*). Thus, a ZEB1⁺ EMT phenotype is an important biomarker of lung cancer pathogenesis of potential importance for early detection of lung epithelial cells harboring oncogenic changes and also as a potential chemoprevention and/or therapeutic target.

Results

Genetic and microenvironmental factors can induce EMT in HBECs partially transformed with defined oncogenic changes. HBECs oncogenically manipulated with p53 knockdown and mutant KRAS^{V12} (hereafter referred to as HBEC3^{p53,KRAS}) are partially transformed (form colonies in soft agar but do not grow tumor xenografts in immunocompromised mice) (4). They can be fully transformed following either exogenous expression of MYC (referred to as HBEC3^{p53,KRAS,MYC}) or switching the cells from serum-free to serum-

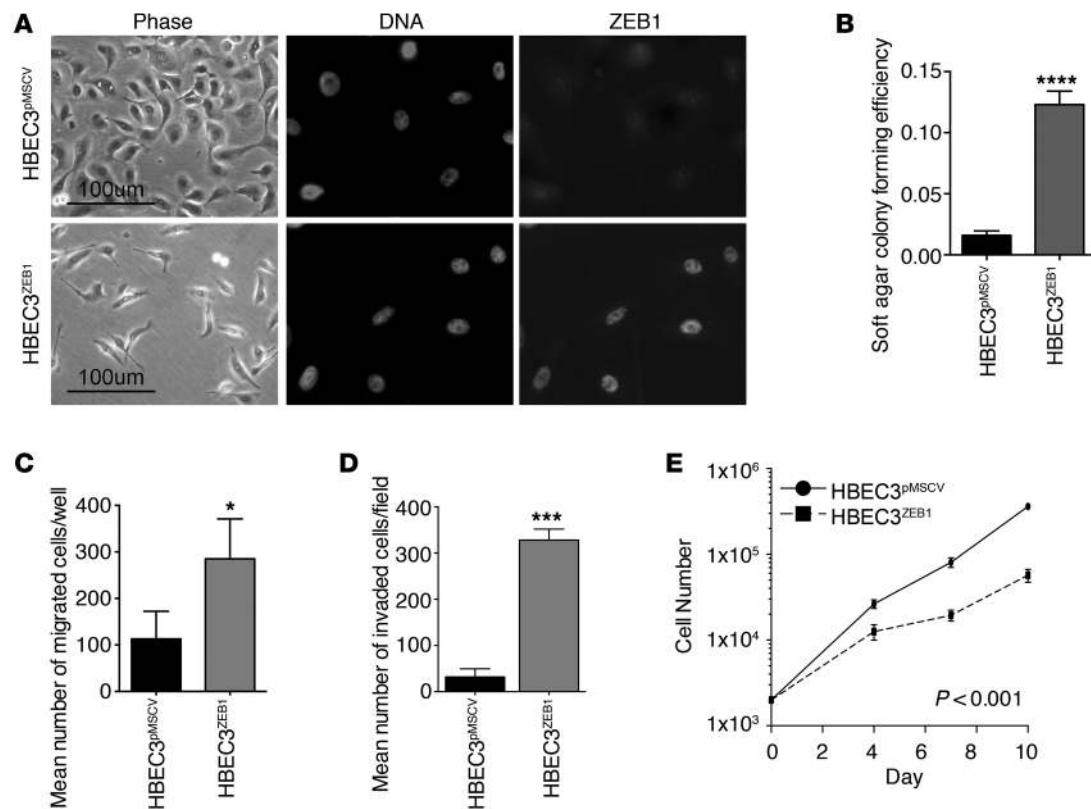


Figure 3. ZEB1 drives EMT in parental HBECs. (A) Phase micrographs and immunofluorescent ($500 \times$) staining of DNA (DAPI) and ZEB1 (FITC) in HBEC3 and HBEC3^{ZEB1}. Scale bars 100 μ m. (B–E) Effect of ZEB1 overexpression in HBEC3 in terms of anchorage-independent (soft agar) colony formation (B), migration (scratch assay) (C), invasion (Matrigel invasion assay) (D), and growth rate (proliferation assay) (E) (mean \pm SD). *P* values were obtained by 1-way ANOVA (B–D) and a nonlinear regression model (E). Data are presented as mean \pm SD and are representative of at least 3 independent experiments. *n* = 3. **P* < 0.05; ****P* < 0.001; *****P* < 0.0001.

containing growth media (RPMI1640 + 10% FBS) (referred to as HBEC3^{p53,KRAS}+FBS) (4). This is accompanied by the cells undergoing an EMT (4). To better understand this mechanism, we measured the mRNA expression of EMT-TFs (*ZEB1*, *ZEB2*, *SNAI1*, *SNAI2*, and *TWIST1*) (Figure 1A). MYC- or serum-induced EMT were both characterized by increased expression of *ZEB1*, *ZEB2*, and *SNAI1*. *ZEB1* had the most significant correlation with expression of the mesenchymal marker vimentin (*VIM*) in both our isogenic series of oncogenically manipulated HBEC3 and a panel of 10 NSCLC cell lines (Spearman $r = 0.82$, $P = 0.0001$; and Spearman $r = 0.77$, $P = 0.013$, respectively) (Supplemental Tables 1 and 2; supplemental material available online with this article; doi:10.1172/JCI176725DS1). Analysis of negative regulators of *ZEB1* and *ZEB2*, the miR-200 family, showed MYC-induced EMT resulted in a decrease in miR-200b and miR-200c, while serum-induced EMT resulted in a decrease in all miR-200 family members (*miR-200a*, *miR-200b*, *miR-200c*, and *miR-141*) (Figure 1B). Immunoblotting confirmed strong induction of *ZEB1* protein expression with both MYC- and serum-induced EMT and loss of expression of the epithelial marker E-cadherin (Figure 1C). Interestingly, introduction of PTEN knockdown or BCL2 overexpression to HBEC3^{p53,KRAS} did not induce an EMT (Figure 1C), suggesting specific oncogenotypes play an important role in EMT induction. Analysis of oncogenically manipulated HBEC lines derived from 2 other individuals besides HBEC3, HBEC2^{p53,KRAS}, and HBEC17^{p53,KRAS} found EMT

was similarly induced with MYC overexpression or growth in serum (Supplemental Figure 1).

ZEB1 is required for MYC- and serum/TGF- β -induced EMT in HBEC3^{p53,KRAS}. FBS contains high levels of TGF- β 1 with 10% FBS-supplemented media reported to contain 1–2 ng/ml of latent TGF- β (17, 18). As TGF- β signaling is a key inducer of EMT (5), we predicted that the serum component inducing EMT in HBEC3^{p53,KRAS}+FBS was TGF- β . This was confirmed by inhibiting serum-induced EMT in HBEC3^{p53,KRAS} with a TGF- β R kinase inhibitor (SB431542) (Figure 1D). As a result, TGF- β was used to represent serum-induced EMT in subsequent experiments. Stable knockdown of *ZEB1* in HBEC3^{p53,KRAS} (referred to as HBEC3^{p53,KRAS,shZEB1}) found *ZEB1* was required for MYC- and TGF- β -induced EMT, where neither factor could induce EMT markers without *ZEB1* expression (Figure 1E). The specificity of the *ZEB1* knockdown was validated with subsequent overexpression of *ZEB1*, which rescued the induced EMT phenotype (Figure 1E).

Genetic and microenvironmental induction of EMT in HBECs is cell context dependent and illustrates the paradoxical role of TGF- β in tumorigenesis. Exogenous expression of MYC (in serum-free media) in nontransformed parental HBEC3 (i.e., without p53 and KRAS manipulations) does not induce EMT (Figure 2A). In addition, we have previously shown that growth of parental HBEC3 cells in serum-containing media does not induce EMT, but rather cellular differentiation and growth arrest (4). Recombinant TGF- β

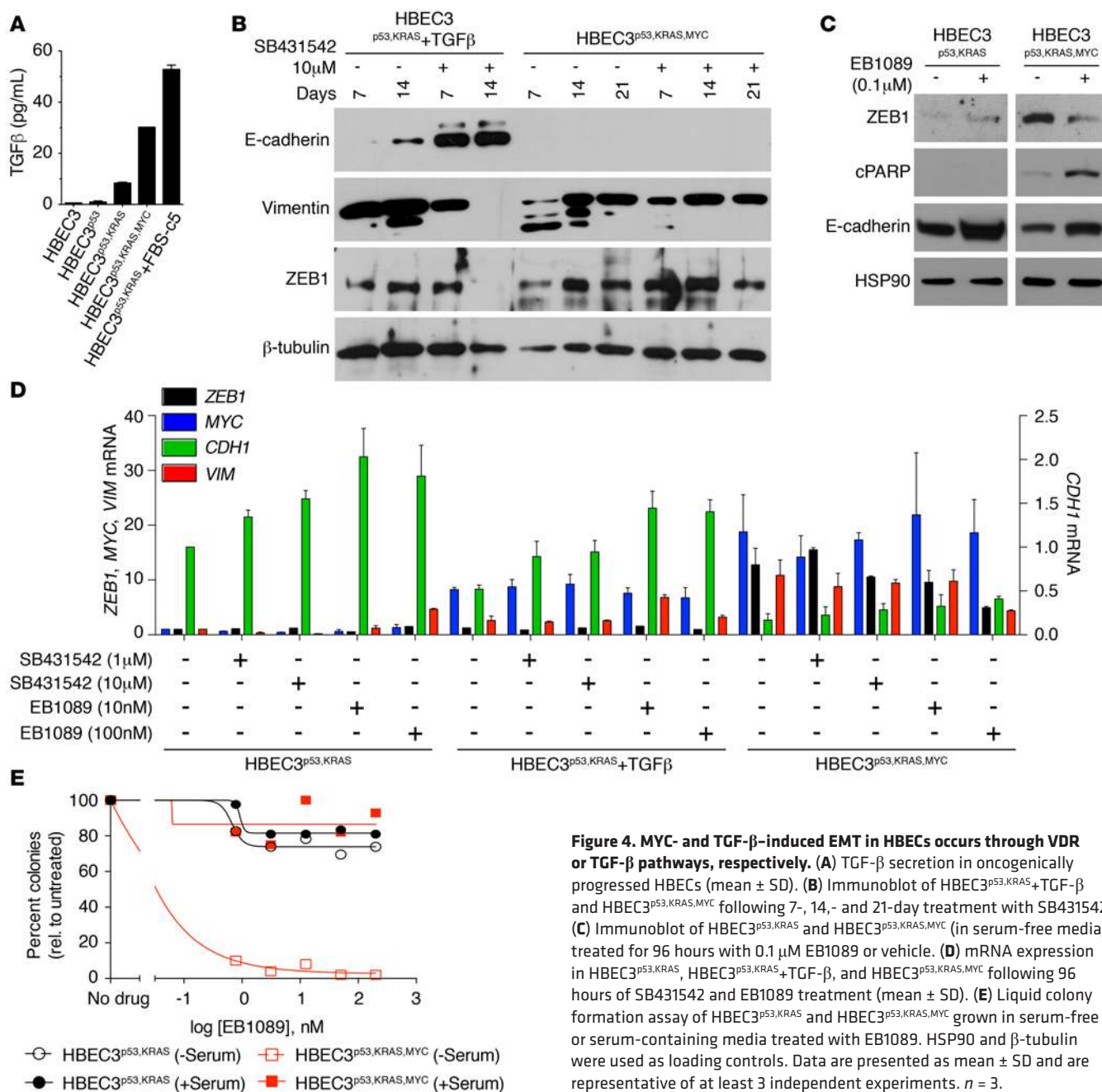


Figure 4. MYC- and TGF-β-induced EMT in HBECs occurs through VDR or TGF-β pathways, respectively. (A) TGF-β secretion in oncogenically progressed HBECs (mean ± SD). (B) Immunoblot of HBEC3^{p53,KRAS}+TGF-β and HBEC3^{p53,KRAS,MYC} following 7-, 14-, and 21-day treatment with SB431542. (C) Immunoblot of HBEC3^{p53,KRAS} and HBEC3^{p53,KRAS,MYC} (in serum-free media) treated for 96 hours with 0.1 μM EB1089 or vehicle. (D) mRNA expression in HBEC3^{p53,KRAS}, HBEC3^{p53,KRAS}+TGF-β, and HBEC3^{p53,KRAS,MYC} following 96 hours of SB431542 and EB1089 treatment (mean ± SD). (E) Liquid colony formation assay of HBEC3^{p53,KRAS} and HBEC3^{p53,KRAS,MYC} grown in serum-free or serum-containing media treated with EB1089. HSP90 and β-tubulin were used as loading controls. Data are presented as mean ± SD and are representative of at least 3 independent experiments. n = 3.

treatment had a similar effect, only inducing growth arrest and apoptosis in parental HBEC3 while inducing EMT in partially transformed HBEC3^{p53,KRAS} (Figure 2B) in a dose- and time-dependent manner (Figures 2, C and D). EMT was transient, where long-term (45 days) TGF-β-treated HBEC3^{p53,KRAS} cells returned to an epithelial-like state (E-cadherin positive, vimentin negative) after 4 weeks of treatment cessation (Figure 2E) during which the cells passed through an intermediate phase where both E-cadherin and vimentin were expressed. Thus, MYC- and serum/TGF-β-induced EMT in HBECs is context dependent, requiring cells to harbor other premalignant oncogenic changes.

TGF-β treatment of our defined HBEC series further delineated the role of genetic context. TGF-β exhibits a growth-inhibitory effect in HBECs with single oncogenic manipulations, HBEC^{p53}

(p53 knockdown) or HBEC3^{KRAS} (mutant KRAS), which is similar to the growth arrest found in parental HBEC3 (Figure 2F), indicating that oncogenic manipulation of both p53 and KRAS are required for HBECs to exhibit a protumorigenic response to TGF-β signaling. Similar findings were made in independent HBEC lines from multiple patients (Supplemental Figure 2). These findings show how our isogenic series of oncogenically manipulated HBECs can help unravel the TGF-β paradox by demonstrating the number of oncogenic changes needed for TGF-β to induce EMT rather than induce growth-inhibitory effects.

ZEB1 can induce EMT in immortalized HBEC3 without oncogenic manipulations, resulting in increased motility and invasiveness despite decreased proliferation. Exogenous overexpression of ZEB1 in parental HBEC3 was sufficient to induce EMT (Figure 2A),

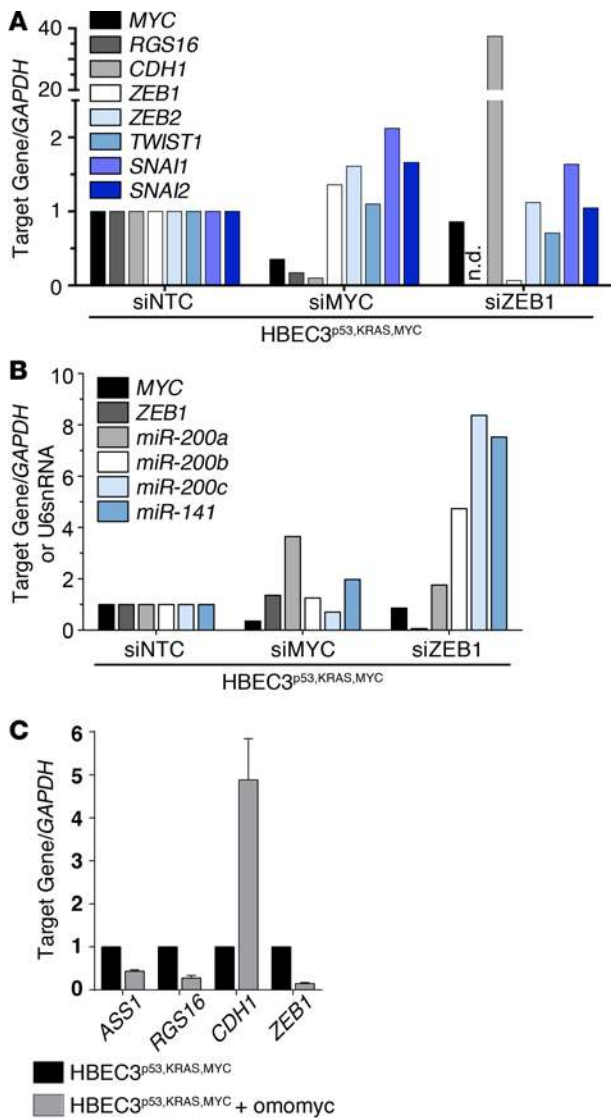


Figure 5. ZEB1 is an indirect target of MYC. (A–B) Expression of EMT markers and EMT-TFs (A) and miR-200 family members (B) in HBE3^{p53,KRAS,MYC} following siRNA-mediated MYC knockdown (mean ± SD). (C) mRNA expression of MYC targets and EMT markers in HBE3^{p53,KRAS,MYC} with or without Omomyc-mediated MYC repression (mean ± SD). Data are representative of at least 3 independent experiments.

ZEB1 expression or EMT phenotype (Figure 4B and Supplemental Figure 4, A–C), indicating MYC-induced EMT was independent of TGF-β signaling despite the increase in TGF-β secretion.

Comparison of mRNA expression microarray data sets of the different oncogenically manipulated HBE3Cs showed a significant overrepresentation of genes involved in the vitamin D receptor (VDR) activation pathway in HBE3^{p53,KRAS,MYC} compared with HBE3^{p53,KRAS} (Supplemental Figure 4D). The VDR pathway promotes epithelial differentiation by VDR binding β-catenin and preventing its nuclear localization and activation of transcriptional targets (19). Activation of the VDR pathway with EB1089 and calcitriol, which are analogues for the VDR ligand 1α,25-dihydroxyvitamin D₃ (1α,25(OH)₂D₃), resulted in increased E-cadherin and decreased VIM and ZEB1 expression in HBE3^{p53,KRAS,MYC} cells (Figure 4, C and D). EMT was also reversed in HBE3^{p53,KRAS}+TGF-β with EB1089 treatment, where 1α,25(OH)₂D₃ is known to modulate the TGF-β signaling pathway (20). EB1089 treatment of HBE3^{p53,KRAS,MYC} also resulted in induction of cleaved poly ADP ribose polymerase (PARP) (Figure 4C), an indicator of cellular toxicity, which was not observed in EB1089 treatment of HBE3^{p53,KRAS}. These effects were further demonstrated with a colony formation assay (Figure 4E), with no colony growth observed in serum-free, mesenchymal HBE3^{p53,KRAS,MYC} cells while no effect upon colony growth was observed in serum-free, epithelial HBE3^{p53,KRAS} cells or serum-induced, mesenchymal HBE3^{p53,KRAS}+FBS or HBE3^{p53,KRAS,MYC}+FBS cells. These data indicate that under serum-free conditions, MYC-induced EMT in HBE3^{p53,KRAS} likely occurs through inhibition of the VDR pathway. However, the VDR pathway is poised to respond and EMT status can be overcome by providing a VDR ligand.

ZEB1 is an indirect target of MYC. To determine whether ZEB1 is a direct transcriptional target of MYC, we transiently knocked down MYC in HBE3^{p53,KRAS,MYC}. RGS16, a known MYC target, was significantly repressed in response to siMYC; however, the expression of ZEB1 and other EMT-TFs was unaltered (Figure 5A). We previously showed that MYC-induced EMT resulted in a decrease in miR-200b and miR-200c (Figure 1B), but knockdown of MYC in HBE3^{p53,KRAS,MYC} cells did not result in any increase in these genes and only a modest increase in miR-200a expression (4-fold). A much greater effect in miR-200 gene expression was observed following ZEB1 knockdown, with a 5- to 9-fold increase in miR-200b, miR-200c, and miR-141 (Figure 5B). To assess the effect of long-term inhibition of MYC on ZEB1 expression (which could occur if MYC indirectly activated ZEB1), we introduced Omomyc into HBE3^{p53,KRAS,MYC}. Omomyc is a dominant-negative mini-protein derived from the bHLH region of MYC that has been mutated to resemble Max, the binding partner of MYC, allowing for dimerization of Omomyc and MYC, resulting in subsequent repression of MYC transcriptional activity (21). In fact, Omomyc-mediated stable repression of MYC in HBE3^{p53,KRAS,MYC} led to decreased ZEB1

indicating the failure of either MYC overexpression or serum/TGF-β treatment alone to induce EMT in parental HBE3 was due to an inability to activate EMT-TFs rather than inability of the cells to undergo EMT. HBE3^{ZEB1} cells displayed an elongated cellular morphology (Figure 3A), increased soft agar colony-forming ability, motility (scratch assay), and invasiveness (Matrigel invasion assay) (Figures 3, B–D, and Supplemental Figure 3) and a decreased proliferation rate (Figure 3E), phenotypes that align with the mesenchymal-like state of promoting motility and survival rather than proliferation (5).

Activation of ZEB1 and EMT in HBE3Cs by MYC occurs through inhibition of the VDR pathway, while that for microenvironmental effects occurs through the TGF-β pathway. TGF-β secretion increased concordantly with oncogenic manipulation in HBE3, including in HBE3^{p53,KRAS,MYC} cells (Figure 4A). Thus, it was possible that this increase in TGF-β secretion in HBE3^{p53,KRAS,MYC} cells was inducing EMT through autocrine activation of TGF-β signaling. However, in contrast with HBE3^{p53,KRAS}+TGF-β cells, inhibition of TGF-β signaling in HBE3^{p53,KRAS,MYC} with a TGF-βR kinase inhibitor (SB431542) or TGF-β monoclonal antibody (1D11) did not affect

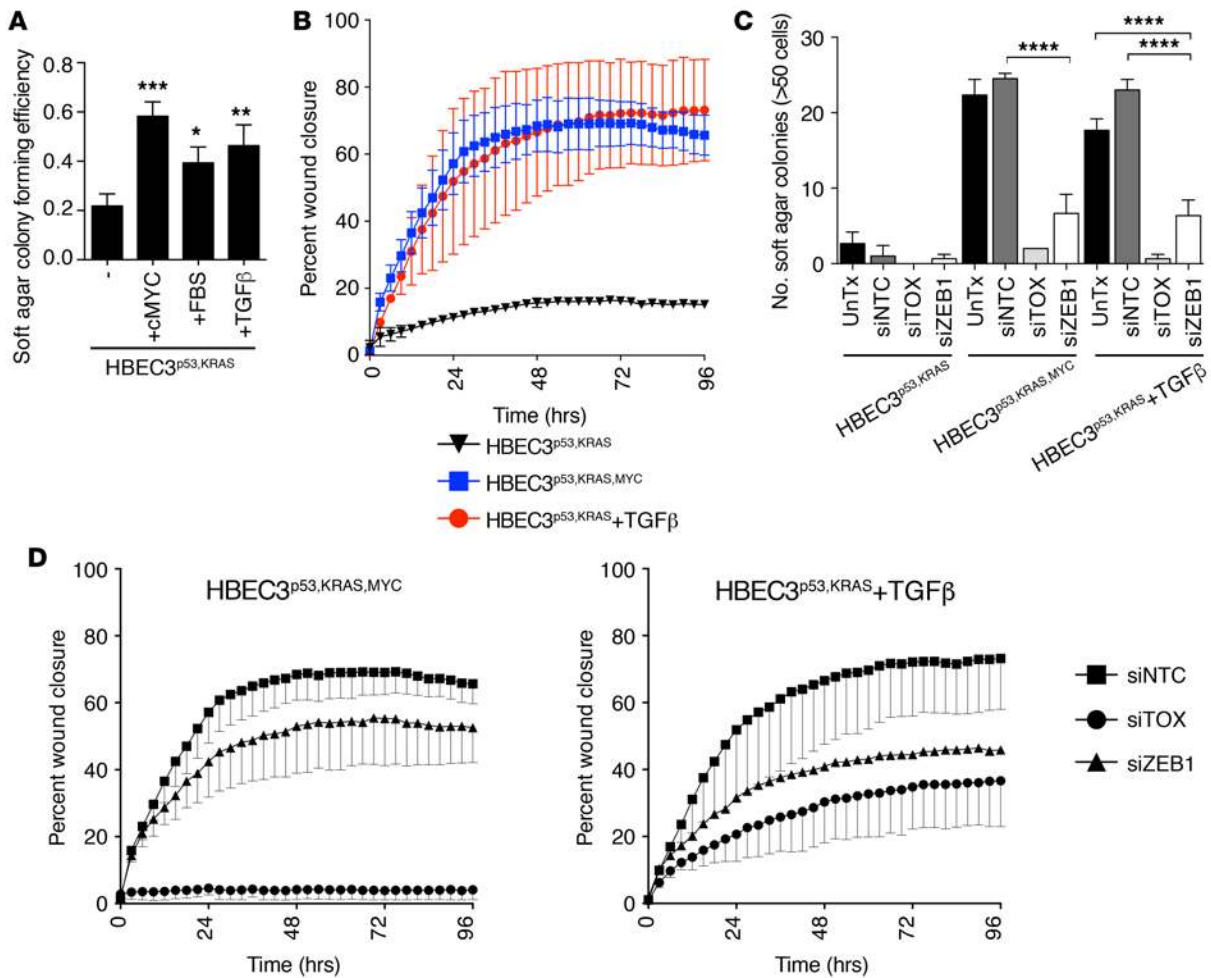


Figure 6. ZEB1 promotes in vitro tumorigenic phenotypes in oncogenically manipulated HBECs. (A and B) Anchorage-independent (soft agar) growth (A) and invasion (B) in HBEC3^{p53,KRAS} transformed with MYC, serum (FBS), or TGF- β . (C and D) Anchorage-independent (soft agar) growth (C) and invasion (D) in HBEC3^{p53,KRAS}, HBEC3^{p53,KRAS,MYC}, and HBEC3^{p53,KRAS}+TGF- β following ZEB1 knockdown (mean \pm SD). UnTx, untreated; siTOX, toxic siRNA. *P* values were obtained by 1-way ANOVA (A and C) and a nonlinear regression model (D). Data are presented as mean \pm SD and are representative of at least 3 independent experiments. *n* = 3. **P* < 0.05; ***P* < 0.01; ****P* < 0.001; *****P* < 0.0001.

and increased *CDH1* expression as well as decreased expression of MYC target genes *ASS1* and *RGS1* (Figure 5C). These data support ZEB1 being an indirect transcriptional target of MYC. Such an indirect effect agrees with publically available ZEB1 ChIP-sequencing (ChIP-Seq) (ENCODE) that shows ZEB1 is not a predicted binding target of MYC (22, 23) (Supplemental Table 3).

ZEB1 causally promotes in vitro transformation and invasiveness in tumorigenic variants of HBEC3. Induction of EMT in partially transformed HBEC3^{p53,KRAS} cells with MYC, serum, or TGF- β leads to significantly increased in vitro transformation, as measured by anchorage-independent (soft agar) growth (Figure 6A) and invasion through Matrigel (Figure 6B). To determine whether the tumorigenic phenotypes were dependent on the high levels of ZEB1, we transiently knocked down ZEB1 (Supplemental Figure 5, A-B). Compared with a nontargeting control siRNA (siNTC), siZEB1 resulted in a significant decrease in anchorage-independent (soft agar) colony formation (Figure 6C) and Matrigel invasion (Figure 6D and Supplemental Figure 5C), demonstrating that ZEB1 is causally involved in the tumorigenic capability of oncogenically manipulated HBECs.

ZEB1 expression varies widely in NSCLC lines and is highly expressed in SCLC lines. To characterize the prevalence and tumorigenic capacity of ZEB1 in human lung cancers, we profiled a panel of 79 NSCLC and 19 small cell lung carcinoma (SCLC) cell lines by quantitative reverse-transcription PCR (qRT-PCR) (Figure 7A), with protein expression confirmed in a subset of the lines (Figure 7D). *ZEB1* expression varied more than 4 logs across the 98 lung cancer cell lines, with significantly higher expression found in SCLC compared with NSCLC lines (*P* < 0.0001), and *ZEB1* expression was correlated with mesenchymal-like status of NSCLC lines, as determined by the expression of E-cadherin and vimentin (Figure 7A). In contrast, there was no significant difference in *ZEB1* expression between cell lines derived from primary or metastatic tumors, NSCLC histological subtypes, or the mutation status of relevant driver lung cancer genes (Supplemental Figure 6, A and B).

ZEB1 promotes in vitro and in vivo tumorigenic phenotypes in lung cancer cell lines. To determine whether ZEB1 drives tumorigenic phenotypes in lung cancer cells similarly to oncogenic HBECs, ZEB1 was stably knocked down in a subset of NSCLC

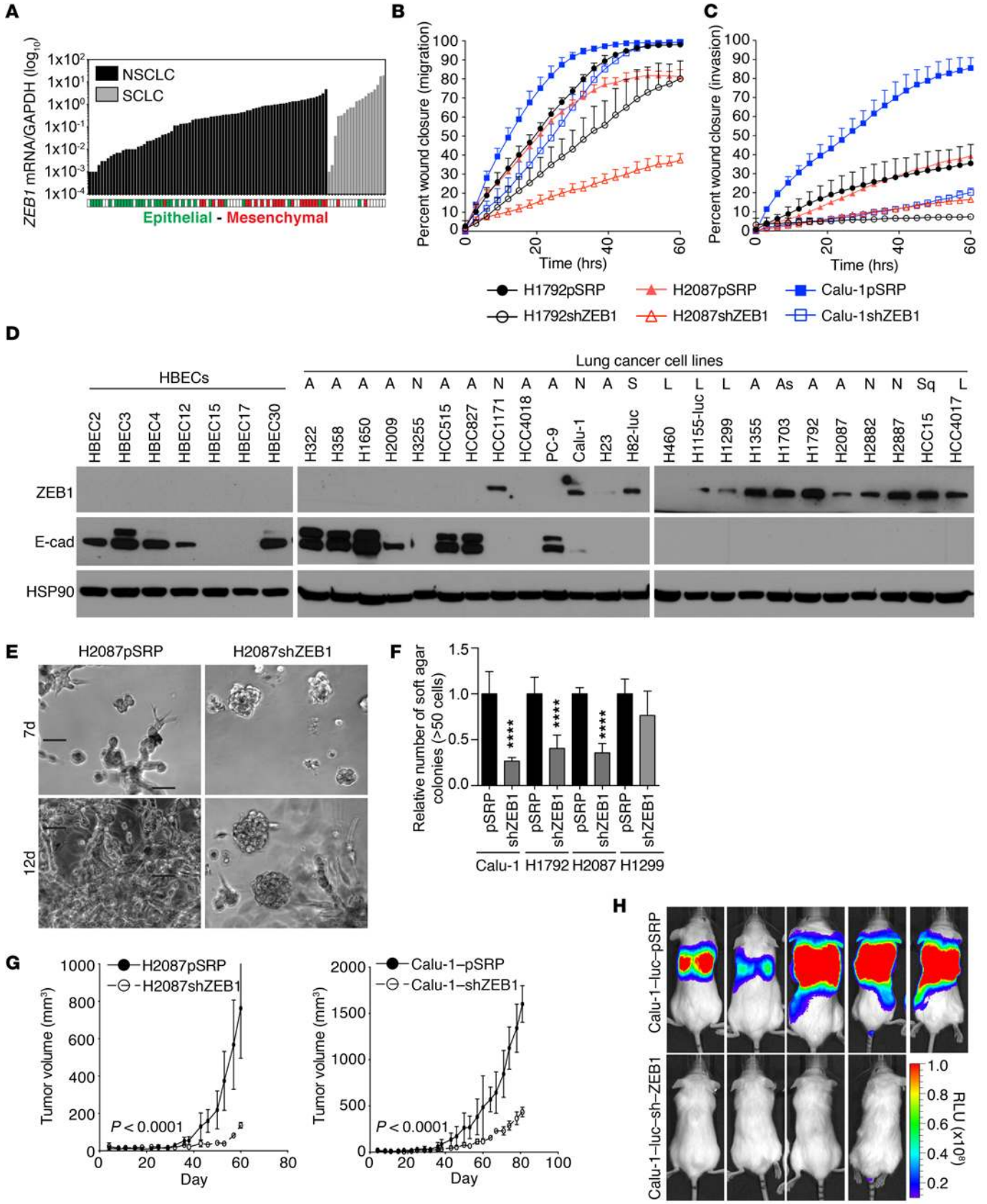


Figure 7. ZEB1 drives tumorigenic phenotypes in lung cancer cell lines in vitro and in vivo. (A) mRNA expression of *ZEB1* using qRT-PCR in a panel of NSCLC (black) and SCLC (gray) cell lines. Cell lines were classified as epithelial-like (green) or mesenchymal-like (red) based upon the expression of E-cadherin and vimentin. (B and C) Cellular migration (scratch assay) (B) and invasion (C) in NSCLC cell lines following stable knockdown of *ZEB1* (mean \pm SD). (D) Immunoblot of *ZEB1* and E-cadherin expression in HBECs, NSCLC, and SCLC cell lines, loaded in order of increasing *ZEB1* mRNA expression (left to right). A, adenocarcinoma; Sq, squamous cell carcinoma; As, adenosquamous cell carcinoma; L, large cell carcinoma; N, NSCLC; S, SCLC. (E) 3D organotypic growth in 2% Matrigel in NCI-H2087 following *ZEB1* knockdown. Scale bars: 10 μ m. (F) Anchorage-independent (soft agar) colony formation in Calu-1, NCI-H1792, NCI-H2087, and NCI-H1299 following *ZEB1* knockdown (mean \pm SD). (G) Subcutaneous xenograft growth in NOD/SCID mice of Calu-1 and NCI-H2087 following *ZEB1* knockdown (mean \pm SD). (H) Bioluminescence imaging of colonizing ability of Calu-1-Luc following *ZEB1* knockdown following intravenous injection into NOD/SCID mice. HSP90 was used as a loading control. pSRP, vector control; RLU, relative light units. *P* values were obtained by 1-way ANOVA (F) and a nonlinear regression model (B, C, G). Data are presented as mean \pm SD. *n* = 3. B, C, E-H show representative data of at least 3 independent experiments. *****P* < 0.0001.

and SCLC lines (NSCLCs, Calu-1, NCI-H1299, NCI-H1792, NCI-H2087; NSCLC neuroendocrine, NCI-H1155; and SCLC, NCI-H82) representing a range of histologies and oncogenotypes (Supplemental Table 4 and Supplemental Figure 6C). *ZEB1* knockdown resulted in an increase in *CDH1* (E-cadherin) expression (Supplemental Figure 6D) and a significant decrease in cell migration (Figure 7B) and invasion (Figure 7C), but did not affect the proliferation rate (Supplemental Figure 6E). Loss of *ZEB1* also decreased 3D growth in a Matrigel basement membrane matrix (Figure 7E) and anchorage-independent soft agar growth (in NCI-H1792, NCI-H2087, and Calu-1, but not NCI-H1299) (Figure 7F). To determine whether these in vitro phenotypes translated to loss of tumorigenicity as an in vivo phenotype, *ZEB1*-manipulated tumor cells were injected subcutaneously and intravenously via the tail vein into NOD/SCID mice and monitored for tumor growth. sh*ZEB1* cells exhibited significantly reduced subcutaneous tumor growth (NCI-H2087, Calu-1, Figure 7G; NCI-H1792, Supplemental Figure 6F). *ZEB1* knockdown also inhibited in vivo metastatic capacity in Calu-1-Luc (Figure 7H), a luciferase-expressing cell line that reproducibly colonizes to the lung following intravenous injection (Supplemental Figure 6G). Intravenous injection models later metastatic events, including survival and transport in circulation, arrest and extravasation from circulation, and colonization and proliferation. Calu-1-Luc vector colonized to the lung in 5/5 NOD/SCID mice, whereas Calu-1-Luc sh*ZEB1* did not exhibit any lung-colonized growth in 4/4 NOD/SCID mice. Collectively, these data show a causal role for *ZEB1* in promoting tumorigenesis in *ZEB1*⁺ lung cancer cell lines, specifically promoting both growth (in soft agar in vitro and subcutaneous in vivo xenografts) and metastatic phenotypes (by in vitro cellular migration and invasion and in vivo colonization).

Identification of candidate ZEB1-regulated genes including downregulation of ESRP1 and upregulation of PMP22 through a multi-data set mRNA microarray analysis and validation of candidate ZEB1-repression targets. To better understand how *ZEB1* promotes tumorigenesis, we generated a list of *ZEB1*-associated genes whose mRNA expression correlated with *ZEB1* across 7 independent whole-genome mRNA microarray data sets comprising 4 isogenic HBECs engineered to have high or low *ZEB1* expression, 157 NSCLC cell lines, and 1,480 primary NSCLC tumors that varied in *ZEB1* expression (10, 24–26) (Supplemental Table 5). These large data sets identified 110 genes (63 positively, 47 negatively) that were associated with *ZEB1* expression in at least 4 of the data sets and included downregulation of epithelial (*CDH1*, *CLDN7*, and *EPHA1*) and upregulation of mesenchymal (*VIM* and *ZEB2*) genes (Supplemental Figure 7A). Peripheral

myelin protein 22 (*PMP22*) and epithelial splicing regulatory protein 1 (*ESRP1*) were the only genes up- or downregulated, respectively, in all 7 data sets. Previously *ESRP1* has been inversely correlated with *ZEB1* expression in lung cancer (10), but there have been no reported associations between *PMP22* and *ZEB1* or EMT. Clustering of the parental and oncogenically manipulated or serum/TGF- β -treated HBEC lines by the 110-gene *ZEB1* signature cleanly separated epithelial-like from mesenchymal-like lines (Supplemental Figure 7B). It also revealed further differences between MYC-induced and TGF- β /serum-induced EMT in HBEC3^{p53,KRAS} cells. For instance, HBEC3^{p53,KRAS,MYC} cells had high expression of *DARC*, *COL6A1/2/3*, *GAS1*, and *ENPP2*, while HBEC3^{p53,KRAS}+TGF- β had high levels of *RGS4* and *LAMA4*. Negatively correlated *ZEB1*-associated genes were downregulated more in HBEC3^{p53,KRAS}+TGF- β compared with HBEC3^{p53,KRAS,MYC}.

A subset of downregulated genes (*ESRP1*, *EPHA1*, *CLDN7*, and *SYK*) was validated by qRT-PCR in *ZEB1*-manipulated cell lines (Figure 8). *ESRP1* and *EPHA1* expression inversely responded to *ZEB1* manipulation, suggesting they are targets of *ZEB1* repression. In contrast, *CLDN7* expression strongly correlated with epithelial-like differentiation, but did not change in response to *ZEB1* manipulation; therefore, it is unlikely that it is regulated by *ZEB1*. Spleen tyrosine kinase *SYK* responded to *ZEB1* in a cell context-specific manner: responding inversely to exogenous *ZEB1* expression in HBEC3 and knockdown in NCI-H1792, but unaltered (NCI-H1155 and NCI-H82) or undetected in other cell lines.

ZEB1, ZEB2, CDH1, ESRP1, and SNAIL1 expression in primary NSCLC tumors. To confirm that the relationships we observed between *ZEB1* and its target genes in lung cancer cell lines were also present in lung tumors, we analyzed gene expression by qRT-PCR in our Specialized Program of Research Excellence (SPORE) cohort of (*n* = 267) clinically annotated primary resected lung adenocarcinomas and squamous cell carcinomas. Spearman correlation analysis found *ZEB1* expression was positively correlated with *ZEB2*, negatively correlated with *CDH1* and *ESRP1* (*P* < 0.0001), and had no correlation with *SNAIL1*, recapitulating our observations in lung cancer cell lines (Supplemental Table 6). Correlation of mRNA expression with clinicopathological variables (histology, tumor stage, sex, age, and neoadjuvant treatment) and mutation status (of *KRAS* and *EGFR*) found *ZEB1* was positively associated with Tumor-Node-Metastasis (TNM) stage (ANOVA Kruskal-Wallis test, *P* = 0.0002) (Figure 9A) and Primary Tumor (T) stage (Pearson χ^2 , *P* = 0.006). This correlation also extended to *ZEB2*, *CDH1*, and *ESRP1* (TNM stage, ANOVA Kruskal-Wallis test, all *P* < 0.0001; T stage: *P* = 0.002, *P* < 0.001, and *P* < 0.022; respectively) (Figures 9, B-D).

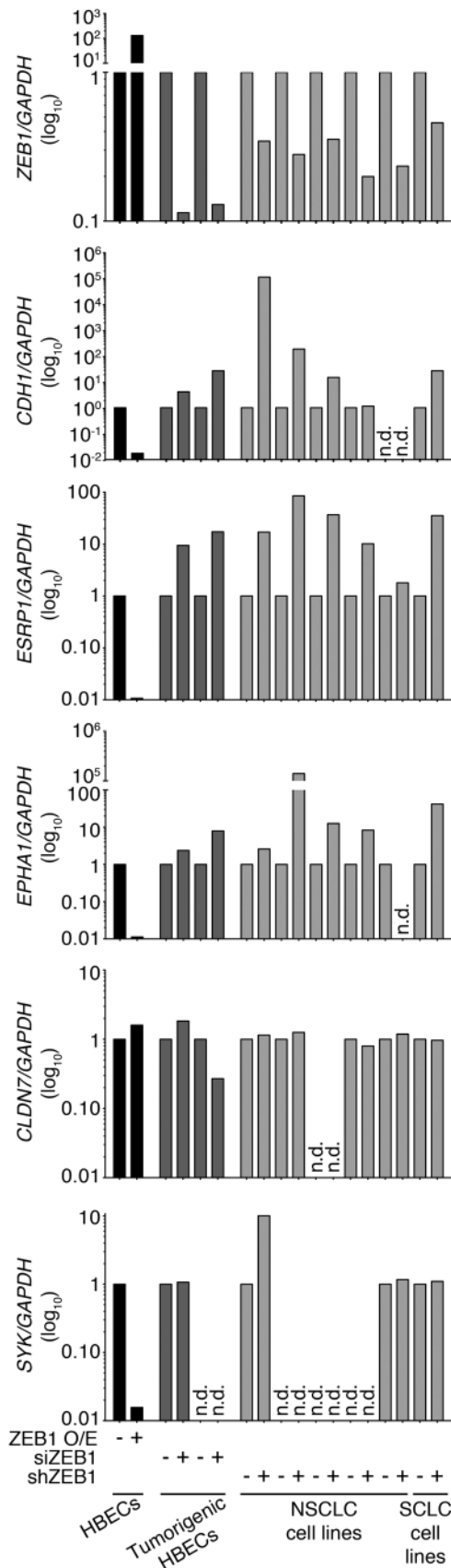


Figure 8. Identification and validation of ZEB1-associated genes through analysis of 7 independent mRNA microarray data sets. Validation of mRNA expression of candidate ZEB1-associated genes in HBEC, NSCLC, and SCLC cell lines manipulated with either ZEB1 overexpression (ZEB1) or ZEB1 knockdown (siZEB1 and shZEB1), relative to expression in control vector cells. ND, not detected. Data are presented as mean ± SD. n = 3.

Stage IV tumors were excluded from the analysis due to small sample size (n = 5). Stratification by smoking history found ZEB1 retained an association with TNM stage (ANOVA Kruskal-Wallis test, P = 0.0023) in smokers (n = 247) (Supplemental Figure 8A). Nonsmokers were not analyzed due to small sample (n = 28). While neither ZEB1 expression nor its closely associated genes ZEB2, CDH1, and ESRP1 were associated with patient outcome across all patients (Supplemental Figure 8, B-I), high expression of ZEB1 was associated with shorter survival within the earliest stage IA tumors (log-rank [Mantel-Cox], P = 0.047) (Figure 9E). SNAI1 did not correlate with tumor TNM stage, but high SNAI1 expression was associated with shorter overall and cancer-free survival (log-rank [Mantel-Cox], P = 0.0251 and P = 0.0052, respectively) (Figure 9, F and G), independently of established clinical prognostic markers such as TNM stage and age. Stratification of tumor samples by histology found that high SNAI1 expression was associated with poorer outcome in adenocarcinomas, but not squamous cell carcinomas (Supplemental Figure 8, J and K), confirming a previous report (27).

ZEB1 inhibits epithelial splicing of CD44 via direct repression of ESRP1, promoting a tumor cell surface CD44^{hi} profile. ZEB1 interacts with either E-box elements (5'-CAGGTG-3', 5'-CACCTG-3', or 5'-CATGTG-3') or Z-box elements (5'-TACCTG-3' or 5'-CAGGTA-3') in the promoter region of target genes, including CDH1 (E-cadherin) (28). Analysis of the upstream promoter sequence of ESRP1 found 3 E-box and 2 Z-box binding sites (Supplemental Figure 9A). ChIP of cell lines with high exogenous (HBEC3^{ZEB1}) or endogenous (HBEC3^{p53, KRAS}+FBS-c11) ZEB1 expression compared with parental HBEC3 found enhanced ZEB1 binding at 2 ESRP1 sites (-738 and -1858) (Figure 10A). ZEB1 repression of ESRP1 is therefore likely to occur through direct binding of ZEB1 to the promoter of ESRP1.

ESRP1 promotes epithelial-type splicing of CD44, a transmembrane glycoprotein, by regulating alternate splicing of the 10 inner variant exons of CD44 pre-mRNA (29). Epithelial (variant) isoforms of CD44 (CD44v) contain combinations of the 10 inner variant exons, whereas the mesenchymal (standard) isoform (CD44s) is devoid of all 10 variant exons. Variant isoforms of CD44 enlarge the stem structure that separates the extracellular domain from the plasma membrane, thereby increasing binding-site exposure (30). RT-PCR analysis of CD44 isoforms (outlined in Supplemental Figure 9B) found that exogenous expression of ZEB1 in HBEC3 leads to loss of epithelial-type CD44v isoforms and increased expression of mesenchymal-type CD44s (Figure 10B). Immunofluorescence analysis confirmed that parental HBEC3 has strong nuclear expression of ESRP1 and low expression of CD44s, but HBEC3^{ZEB1} cells (stable overexpression of ZEB1) lose ESRP1 expression and gain strong membranous expression of CD44s (Figure 10D).

To determine whether CD44 mRNA isoform expression correlated with CD44 cell-surface expression, HBEC3 and HBEC3^{ZEB1}

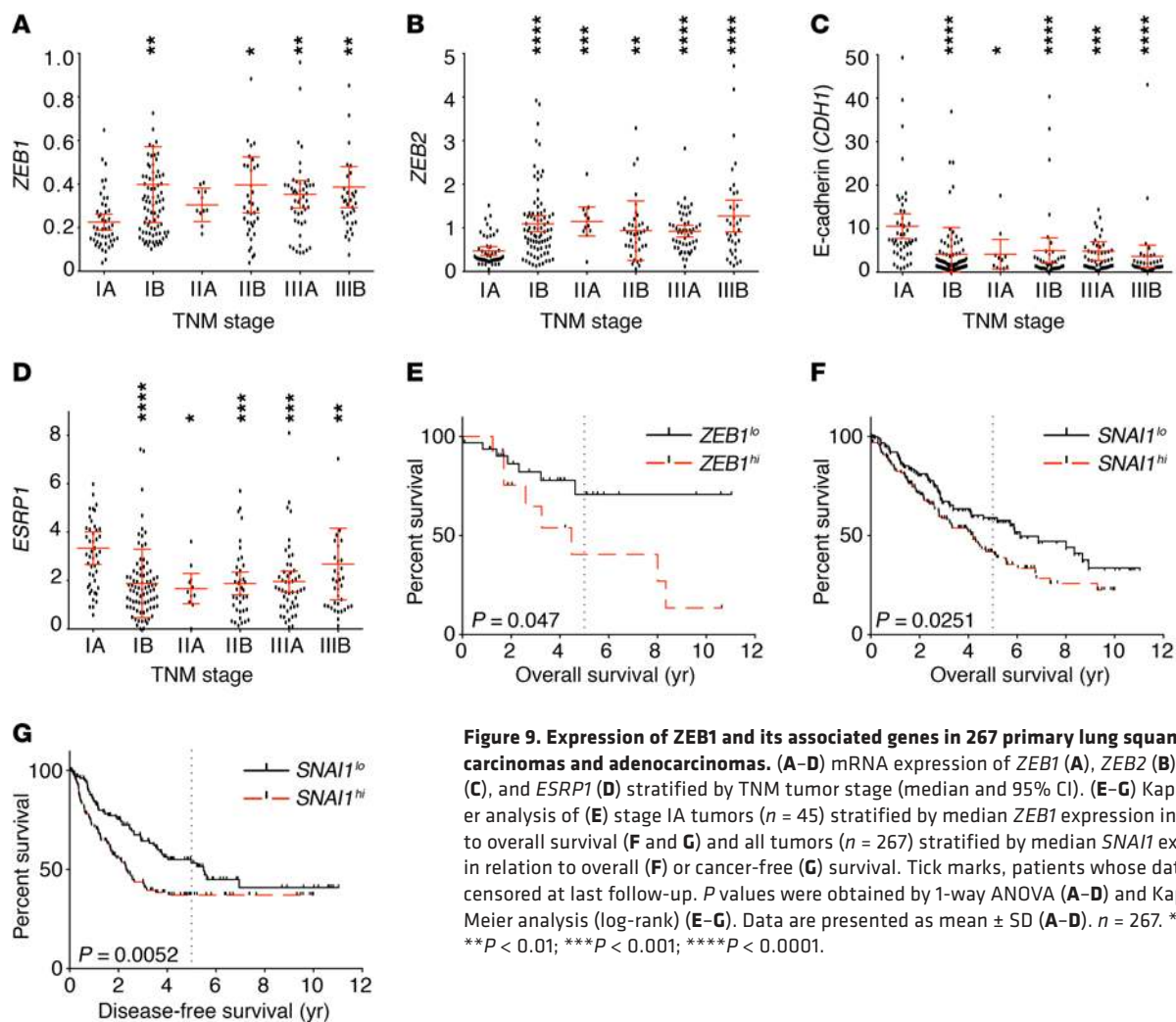


Figure 9. Expression of ZEB1 and its associated genes in 267 primary lung squamous cell carcinomas and adenocarcinomas. (A–D) mRNA expression of ZEB1 (A), ZEB2 (B), CDH1 (C), and ESRP1 (D) stratified by TNM tumor stage (median and 95% CI). (E–G) Kaplan-Meier analysis of (E) stage IA tumors ($n = 45$) stratified by median ZEB1 expression in relation to overall survival (F and G) and all tumors ($n = 267$) stratified by median SNAI1 expression in relation to overall (F) or cancer-free (G) survival. Tick marks, patients whose data were censored at last follow-up. P values were obtained by 1-way ANOVA (A–D) and Kaplan-Meier analysis (log-rank) (E–G). Data are presented as mean \pm SD (A–D). $n = 267$. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

were profiled with flow cytometric analysis. Exogenous expression of ZEB1 results in a 10-fold increase in CD44 cell-surface expression (Figure 10C), switching cells from a CD44^{lo} to a CD44^{hi} phenotype. Thus, we found that high expression of ZEB1 in HBECs leads to repression of ESRP1, which inhibits epithelial-specific splicing of CD44, leading to increased expression of the mesenchymal isoform, CD44s, and switching cells to a CD44^{hi} profile.

ZEB1 causes HBEC3 to acquire a CD24^{lo}CD44^{hi} phenotype with tumorigenic progression. CD24 is a cell-surface protein anchor that is thought to have an essential role in cell differentiation. It has been used to define a subset of breast cancer stem cells (CSCs) in conjunction with CD44 (CD24^{lo}CD44^{hi}) (31). There is a close link between EMT and tumor-initiating cells/CSCs (32), and we found that, in addition to ZEB1 promoting a CD44^{hi} profile (Figure 10C), CD24 negatively correlated with ZEB1 expression in 3 mRNA data sets (Supplemental Table 5). Profiling parental HBEC3^{pMSCV} (control vector) and HBEC3^{ZEB1} cells for CD24/CD44 cell surface expression found overexpression of ZEB1 switched cells from a CD24^{lo}CD44^{lo} to a CD24^{lo}CD44^{hi} profile (Figure 11A and Supplemental Figure 10).

Profiling our series of oncogenically progressed HBEC3 for CD44 splicing and CD24/CD44 cell surface expression found oncogenic progression resulted in a CD44 isoform switch with epithelial derivatives (HBEC3, HBEC3^{p53}, and HBEC3^{p53,KRAS}) express-

ing CD44v and mesenchymal derivatives (HBEC3^{p53,KRAS,MYC} and HBEC3^{p53,KRAS+FB}) expressing only CD44s (Figure 11B). This correlated with a gradual increase in CD44^{hi} cells where each additional manipulation increased the population (~1% in parental HBEC3 to 91% in HBEC3^{p53,KRAS,MYC} and 94% in HBEC3^{p53,KRAS+TGF- β}) (Figure 11C). Thus, ZEB1 strongly drives the CD44 but not CD24 profile in HBECs and increasing oncogenic transformation of HBECs correlates with an increasing population of CD44^{hi} cells.

CD24^{lo}CD44^{hi} selects for a highly transformed subpopulation in HBEC3^{p53,KRAS}. To determine whether CD24^{lo}CD44^{hi} HBEC populations differ in tumorigenicity, we isolated 3 distinct cell populations observed among HBEC3^{p53,KRAS} cells (CD24^{lo}CD44^{lo}, CD24^{hi}CD44^{lo}, and CD24^{lo}CD44^{hi}) by FACS (Figure 12A). Long-term (4 weeks) culturing in vitro found CD24^{lo}CD44^{lo} and CD24^{hi}CD44^{lo} populations maintained a stable CD24/CD44 profile; however, only approximately 50% of CD24^{lo}CD44^{hi} cells retained their original profile, with the remainder becoming CD44^{lo} (Supplemental Figure 11A). This suggests either a potential CD44^{lo} population outgrowth (post-sort purity > 98% CD44^{hi}) or that CD24^{lo}CD44^{hi} cells possess regenerative capability. Morphologically, CD24^{lo}CD44^{hi} cells are mesenchymal like, while CD24^{lo}CD44^{lo} and CD24^{hi}CD44^{lo} cells are epithelial like (Supplemental Figure 11B), confirmed by expression

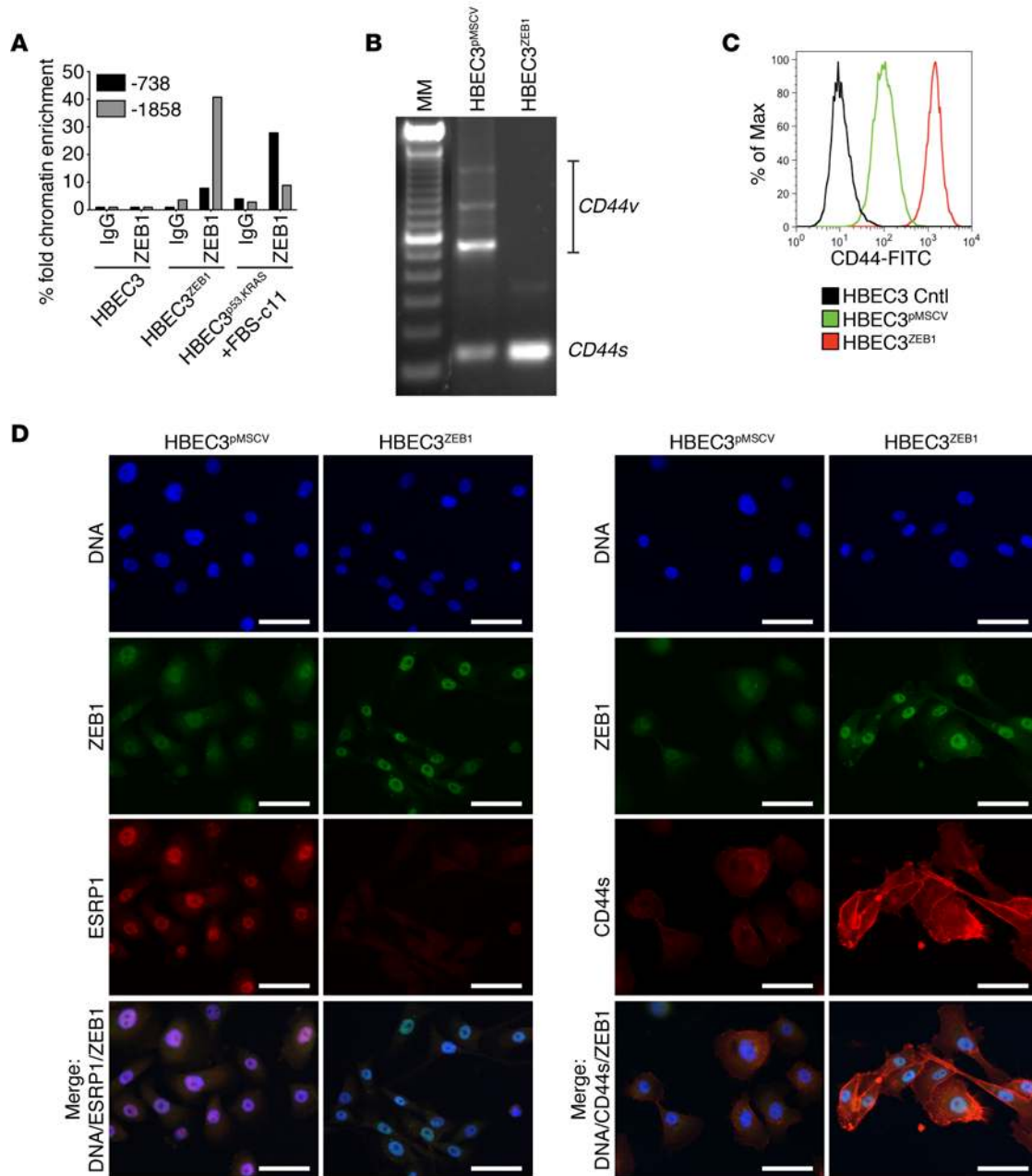


Figure 10. ZEB1 promotes mesenchymal splicing of CD44 to CD44s by directly repressing ESRP. (A) ChIP assay of -738 and -1858 putative ZEB1-binding sites in parental HBEC3 cells (no ZEB1 expression) and HBEC3^{ZEB1} and HBEC3^{p53,KRAS}-c11 cells (high ZEB1 expression) (mean ± SD). (B) Expression of CD44 isoforms in HBEC3^{pMSCV} and HBEC3^{ZEB1}. (C) FACS analysis of cell-surface expression of CD44 following overexpression of ZEB1 in HBEC3. (D) Immunofluorescent (500 s⁻¹) staining of DNA (DAPI, blue), ZEB1 (FITC, green), ESRP1, and CD44s (Alexa Fluor 594, red) in HBEC3^{pMSCV} and HBEC3^{ZEB1}. Scale bar 50 μM. pMSCV, vector control. Data in A are shown as mean ± SD and are representative of at least 3 independent experiments. n = 3.

of vimentin and E-cadherin (Figure 12B), with CD24^{hi}CD44^{lo} cells in an intermediate EMT phase lacking expression of both E-cadherin and vimentin. ZEB1 is not detectable in parental, unsorted HBEC3^{p53,KRAS}, or in CD24^{lo}CD44^{lo} or CD24^{hi}CD44^{lo} populations, yet is highly expressed in CD24^{lo}CD44^{hi} cells (Figure 12B). This corresponds with CD24^{lo}CD44^{hi} cells primarily expressing mesenchymal CD44s while the parental and other sorted populations express epithelial CD44v (Figure 12C). Anchorage-dependent colony formation assays show mesenchymal CD24^{lo}CD44^{hi} cells proliferate more slowly than CD24^{lo}CD44^{lo}

and CD24^{hi}CD44^{lo} cells (Supplemental Figure 11C) and that this phenotype remained stable with long-term culture. However, they generate a significantly greater number of anchorage-independent (soft agar) colonies compared with CD24^{lo}CD44^{lo} and CD24^{hi}CD44^{lo} cells (Figure 12D). CD24^{lo}CD44^{hi} cells also form very large, macroscopic colonies (Figure 12D), which is an HBEC phenotype we have previously correlated with in vivo tumorigenicity (4). CD24^{lo}CD44^{hi} cells were also significantly more invasive in vitro compared with CD24^{lo}CD44^{lo} and CD24^{hi}CD44^{hi} populations (Figure 12E). Subcutaneous injection of the 3 CD24/CD44

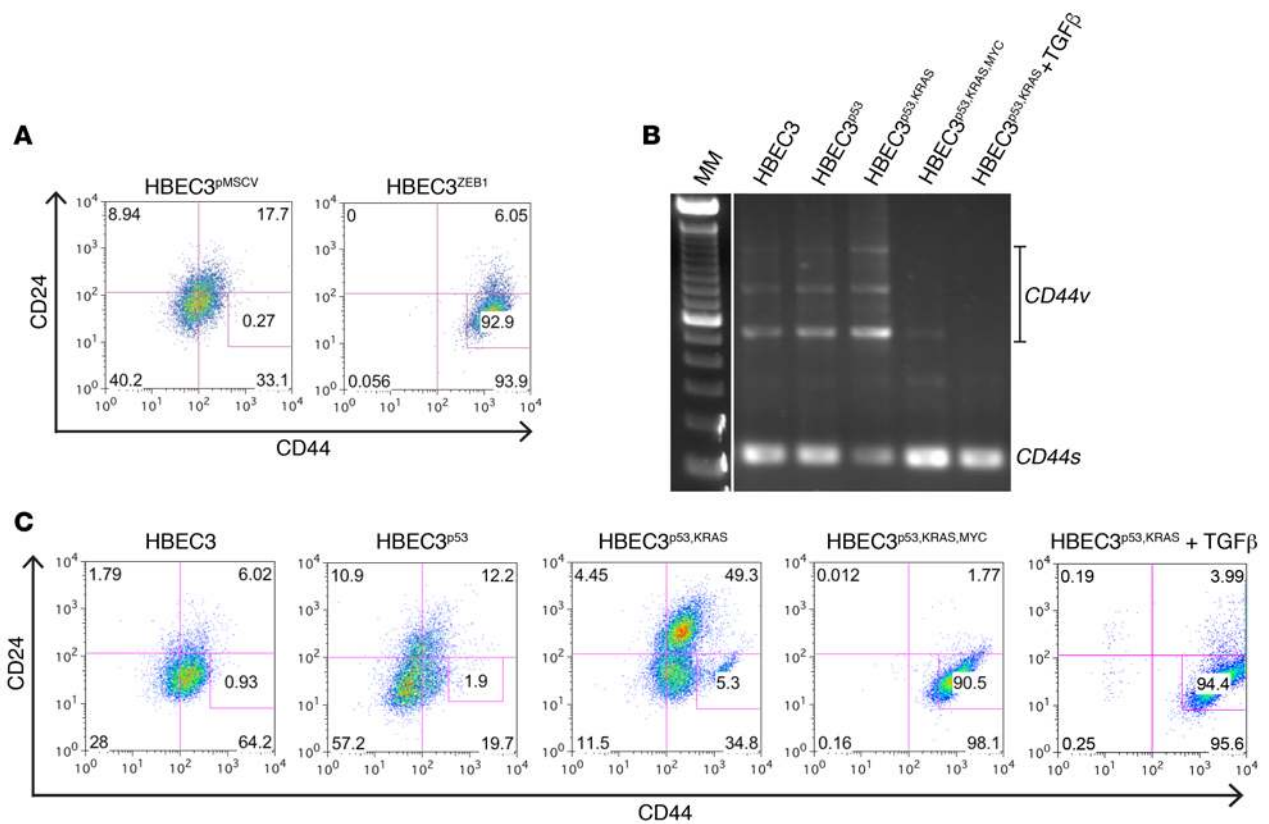


Figure 11. ZEB1 promotes mesenchymal splicing of CD44 to CD44s by directly repressing ESRP1, correlating with a CD44^{hi} profile. (A) FACS analysis of cell-surface expression of CD24/CD44 following overexpression of ZEB1 in HBEC3. (B) CD44 isoform expression in the HBEC3 oncogenic progression series. Line indicates lanes were run on the same gel but were noncontiguous. (C) CD24/CD44 profiles of the HBEC3 oncogenic progression series. In A and C, numbers in each corner represent the percentage of cells within that quarter. Gates were drawn in control cells to represent CD24^{lo}CD44^{hi} cells, and numbers within the boxed regions represent percentage of CD24^{lo}CD44^{hi} cells. pMSCV, vector control. Data are representative of at least 3 independent experiments.

sorted populations of HBEC3^{p53,KRAS} into NOD/SCID mice did not, however, result in any tumor xenograft growth (data not shown). Importantly, transient knockdown of ZEB1 in the CD24^{lo}CD44^{hi} population resulted in significantly decreased anchorage-independent growth (Figure 12F) and cellular invasion (Figure 12G), showing that ZEB1 was integral to driving these transformed phenotypes in HBEC3^{p53,KRAS}-CD24^{lo}CD44^{hi} cells.

Aldehyde dehydrogenase (ALDH) activity has been shown to identify lung CSCs/tumor propagating cells (TPCs) (33, 34). Non-transformed HBECs are enriched for ALDH⁺ cells (~50%), consistent with HBECs representing a basal lung stem cell population, but the ALDH⁺ population decreases with each successive oncogenic manipulation to where tumorigenic HBEC3 derivatives have only 10%–15% ALDH⁺ cells (Supplemental Figure 11D). Thus, with each defined increase in malignant transformation in HBECs, we see a decrease in the ALDH⁺ population and a concomitant increase in the CD24^{lo}CD44^{hi} population. These data indicate that within epithelial HBECs, the subpopulation of CD24^{lo}CD44^{hi} cells are mesenchymal and strongly ZEB1 positive, with enhanced anchorage-independent growth, and while they may represent tumor-initiating cells, they are not yet fully tumorigenic. These findings are other further indications of the important early role ZEB1 plays in the premalignant phenotype.

Discussion

In the current work, we have functionally and mechanistically studied the EMT phenotype in a series of isogenic, oncogenically manipulated HBECs, a panel of lung cancer lines, and large molecularly profiled, clinically annotated, resected NSCLC tumor data sets. From these results, we have discovered that ZEB1-driven EMT is a critical event in the malignant progression of HBECs to full malignancy and EMT and expression of the EMT-TF ZEB1 is observed in early stage NSCLC tumors. We studied the importance of specific oncogenetic context in the response of bronchial epithelial cells to known tumor progression stimuli where microenvironmental (e.g. exposure to TGF-β) and specific additional oncogenetic (e.g. MYC, but not BCL2 or PTEN) cues induce a ZEB1-dependent EMT in HBECs transformed with both p53 and KRAS oncogenic manipulations, but not in non- or single-oncogene manipulated HBECs. Functionally, we found ZEB1 has a causal role in HBEC progression and in promoting tumorigenicity of lung cancer cell lines, in part, by directly binding to the promoter of ESRP1 where repression of ESRP1 transcription leads to increased mesenchymal-type splicing of CD44. The ZEB1-driven increase in CD44s, the mesenchymal isoform, switches cells from a CD44^{lo} to a CD44^{hi} state. Importantly, we found that CD44^{hi} status could select for a small, highly invasive, highly transformed,

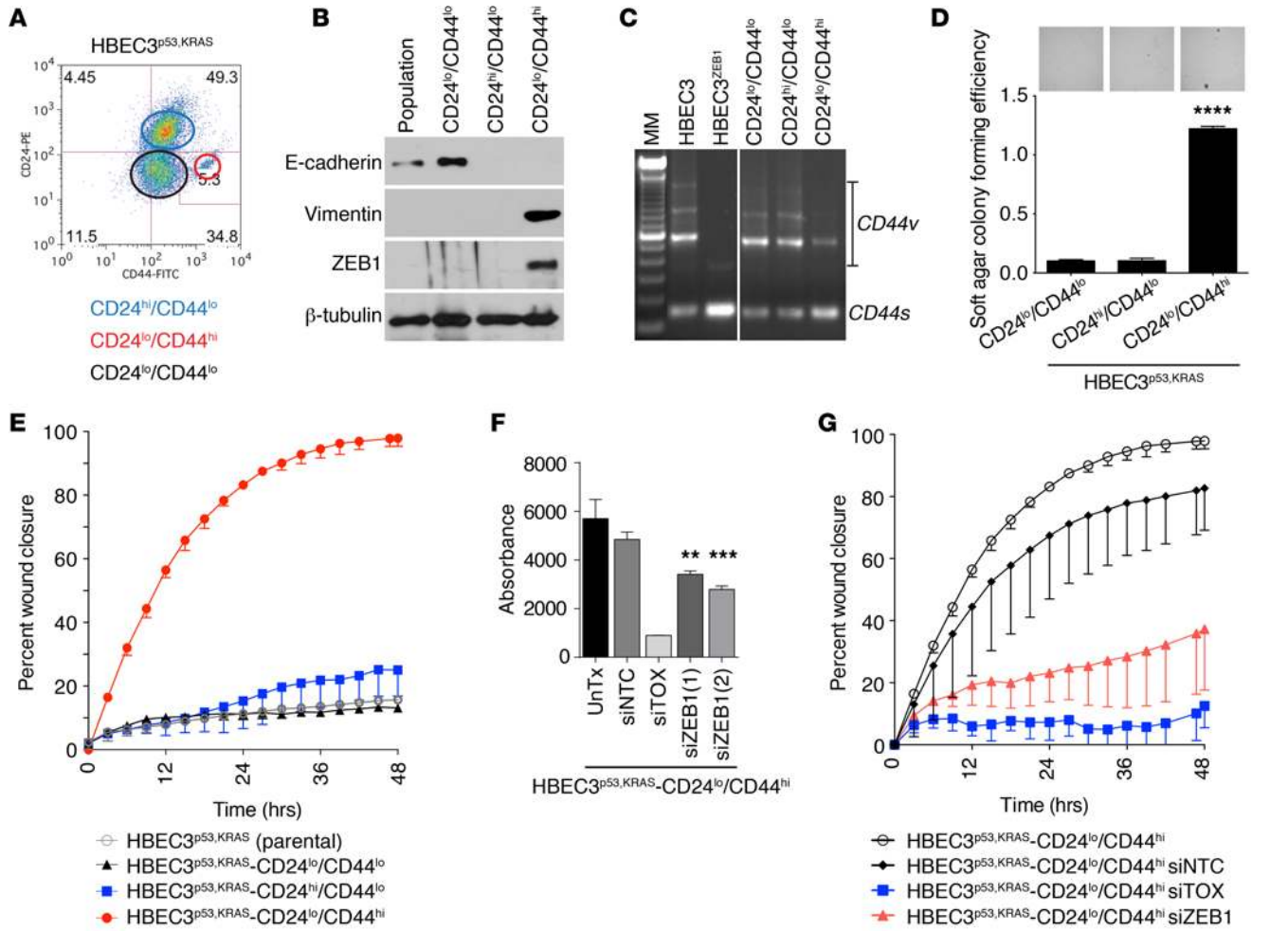


Figure 12. A ZEB1-driven switch from CD44v to CD44s selects for highly transformed cells in premalignant HBEC3^{p53,KRAS} cells. (A) Three populations of HBEC3^{p53,KRAS}, CD24^{lo}CD44^{lo} (black), CD24^{hi}/CD44^{lo} (blue), and CD24^{lo}CD44^{hi} (red) were sorted for subsequent analysis. Numbers in each corner represent the percentage of cells within that quarter. Gates were drawn in control cells (HBEC3) to represent CD24^{lo}CD44^{hi} cells where the number within the boxed region represents percentage of CD24^{lo}CD44^{hi} cells. (B) Immunoblot for ZEB1 and EMT markers in CD24/CD44 sorted populations of HBEC3^{p53,KRAS}. (C) CD44 isoform expression in CD24/CD44 sorted populations of HBEC3^{p53,KRAS}. Line indicates lanes were run on the same gel but were noncontiguous. (D) Anchorage-independent (soft agar) colony formation of CD24/CD44 sorted populations of HBEC3^{p53,KRAS} (mean ± SD). (E) Cellular invasion of CD24/CD44 sorted populations of HBEC3^{p53,KRAS} (mean ± SD). (F) Anchorage-independent (soft agar) colony formation of CD24/CD44 sorted populations of HBEC3^{p53,KRAS} following ZEB1 knockdown (mean ± SD). (G) Cellular invasion of HBEC3^{p53,KRAS}-CD24^{lo}CD44^{hi} sorted cells following ZEB1 knockdown (mean ± SD). β-Tubulin was used as a loading control. P values were obtained by 1-way ANOVA (D, F) and a nonlinear regression model (E, G). Data are presented as mean ± SD (D–G) and are representative of at least 3 independent experiments. n = 3. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

ZEB1-driven subpopulation in premalignant, partially transformed HBEC^{p53,KRAS}.

Acquiring mesenchymal properties is thought to confer a malignant epithelial cell with metastatic capability, and as such, EMT and metastasis are generally considered late events in tumorigenesis. A recent study, however, showed they can occur in early “preinvasive” stages of pancreatic adenocarcinoma (16). Here, we show 2 lines of evidence that EMT and expression of ZEB1 are early events in lung cancer pathogenesis: occurring in nonmalignant HBECs harboring oncogenic mutations, and early stage IB primary lung adenocarcinomas and squamous cell carcinomas. Stage IB tumors do not, by definition, exhibit local nodal metastasis, indicating that detection of ZEB1⁺ cells can occur before detectable metastatic disease. To extend these findings, studies of ZEB1 in histologically normal human bronchial epithelium and preneoplastic lesions are required.

We show genetic context is a critical factor in EMT-driven tumorigenesis in HBECs (Figure 13A). TGF-β-induced EMT exhibited a TGF-β paradox effect in HBECs requiring oncogenic mutations in both p53 and KRAS to switch TGF-β to a protumorigenic function. This switch may be specifically associated with these 2 very common mutations in lung cancer, with a spontaneously acquired mutation that may, for example, disable TGF-β signaling (35) or transformation of the cells to a more malignant state that is more responsive to oncogenic stimuli. MYC-induced EMT in HBECs was also dependent upon genetic context but sensitive to vitamin D pathway activation and independent of TGF-β signaling. Thus, multiple, independent factors can induce EMT in HBECs via induction of ZEB1, but they require either specific oncogenic mutations or a minimum oncogenic state. The ZEB1 promoter has been shown to exist in either a repressed or

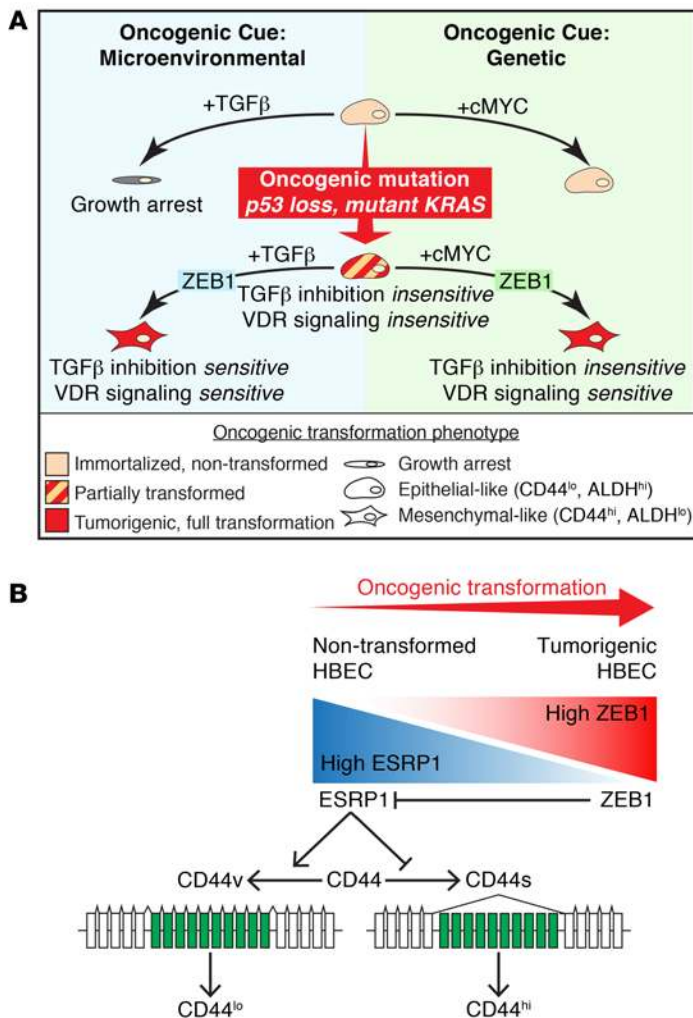


Figure 13. Tumorigenic progression of HBECs with EMT and ZEB1.

(A) The importance of genetic context in the response of bronchial epithelial cells to microenvironmental or genetic oncogenic cues. Nontransformed HBECs undergo either growth arrest or negligible transformation in response to TGF-β (microenvironmental) or exogenous expression of MYC (genetic), respectively. These same cues drive oncogenic progression in HBECs harboring p53 and KRAS oncogenic mutations where cells undergo an EMT, become fully tumorigenic, and acquire a CD44^{hi} profile. Disparate responses to TGF-β inhibition and VDR activation suggest these oncogenic cues signal through independent pathways. However, both require expression of ZEB1 to achieve EMT. (B) Working model for the role of ZEB1 in the malignant transformation of HBECs. ZEB1 expression endogenously increases with escalating oncogenic transformation. ZEB1 promotes tumorigenicity by directly repressing transcription of ESRP1 leading to decreased epithelial splicing (CD44v) and, as a result, increased mesenchymal splicing (CD44s) of CD44 that correlates with a CD44^{hi} profile.

ZEB2 were required for ESRP1 repression (42). ESRP1 is an epithelial cell-type-specific regulator of an extensive splicing network, including genes involved in cell polarity, cell-cell adhesion, and migration (43), and as such, is an important regulator of EMT (44). We extended the understanding of ZEB1 biology by showing its repression of ESRP1, and the resulting increased mesenchymal splicing of its target, the cell-surface antigen CD44, drives a switch in cells from an epithelial CD44^{lo} status to a mesenchymal CD44^{hi} status. Using our HBEC model for the oncogenic transformation of bronchial epithelial cells, we found oncogenic progression leads to a concordant increase in CD44^{hi} cells, which represent a highly transformed subpopulation, potentially tumor-initiating cells (Figure 13B). CD44s expression is increased in high-grade breast tumors (45) and predicts for poor prognosis in resected NSCLC (46). This correlates with our linkage of the CD44s isoform to a CD44^{hi} profile and demonstrates that these cells represent a highly transformed subpopulation.

CD44^{hi} has not been conclusively established as a marker for lung CSCs (47), but our data showed that CD44^{hi} cells but not CD44^{lo} cells were able to regenerate both CD44 populations (CD44^{lo} and CD44^{hi}). With increasing malignant transformation in HBECs, we observed an increase in CD24^{lo}CD44^{hi} cells with a concomitant decrease in ALDH⁺ cells. This could reflect interchangeable, oncogenically driven states in lung stem cells comparable to those of normal and malignant breast stem cells that have been proposed to exist in distinct, interchangeable states where EMT-like CSCs are CD24^{lo}CD44^{hi} and MET-like CSCs are ALDH⁺ (48).

This research has revealed several potential translational applications for ZEB1-driven changes relevant for lung cancer: early detection through the ZEB1/EMT/CD44 molecular biomarker panel to identify lung epithelial cells with oncogenic changes; determination of prognosis of resected NSCLCs based on tumor expression of ZEB1/EMT biomarkers; and ZEB1 or its downstream genes focused on chemoprevention and/or targeted tumor therapy directed at TGF-β and VDR.

In conclusion, we have established an important functional role for ZEB1 in lung cancer pathogenesis and demonstrate that ZEB1 and induction of EMT may occur early in lung tumorigenesis in premalignant cells. These findings provide information on the role of EMT in lung tumorigenesis, showing the importance of

bivalent/“poised” chromatin configuration in breast cancer and mammary epithelial cell lines, where the “poised” state is readily induced in response to stimuli (36). Oncogenic progression of HBECs may cause a switch in ZEB1 promoter chromatin configuration, allowing the cells to respond to genetic and microenvironmental stimuli that engage EMT and transform cells. Importantly, we showed TGF-β- and MYC-induced EMT in p53 and KRAS^{V12}-transformed HBEC3 cells required induction of ZEB1, indicating that EMT-TF is a crucial regulator of EMT in HBECs.

Induction of ZEB1 in oncogenically progressed HBECs had a causal effect on tumorigenesis, promoting anchorage-independent growth, cellular migration, and invasion. This finding was extended to a large panel of NSCLC and SCLC cell lines and showed that ZEB1 also promoted tumor growth and metastatic capacity in vivo. This builds upon previous in vitro (15) and in vivo (37–40) studies of ZEB1 to establish its important functional role in lung cancer development.

Our data support earlier studies showing ZEB1 expression is inversely correlated with E-cadherin expression in lung cancer (10, 14), a known target of ZEB1 repression. However, we also found ZEB1 directly binds to and represses the epithelial splicing protein ESRP1, which was recently confirmed in lung cancer cells (41) and differs from mouse mammary epithelial cells, where both ZEB1 and

specific oncogenetic context and preexisting oncogenic mutations to engage EMT machinery. They also establish ZEB1 as an important molecular target to limit metastatic disease in lung cancer, although further research into targeted inhibition of ZEB1, as recently shown in Snail inhibitor Co(III) conjugates (49), is required.

Methods

Extended materials and methods are provided in Supplemental Methods.

Cell lines. The CDK4/TERT-immortalized HBEC lines used in this study have been described previously (3, 4, 50). Lung cancer cell lines were established by our laboratory (51). Culturing conditions are described in Supplemental Table 7. All cell lines were DNA fingerprinted using a PowerPlex 1.2 Kit (Promega) and confirmed free of mycoplasma by an e-Myco Kit (Boca Scientific). ZEB1 knockdown and exogenous overexpression were introduced using pSuper-RetroshZEB1 (Supplemental Table 9) (a gift from Thomas Brabletz, University of Freiburg, Freiburg, Germany) and pMSCV-ZEB1 retroviral vectors, respectively. Cell lines were transduced as described previously (3) and selected with the appropriate antibiotic.

In vitro assays. Preparation of total cell lysates and Western blotting were performed as described previously (4), with primary antibodies listed in Supplemental Table 8. Levels of secreted TGF- β 1 were measured with the TGF- β 1 EMAX Immunoassay ELISA Kit (Promega). siRNA reverse transfections were performed as described previously (52), and cells were harvested 48 hours after transfection for in vitro tumorigenicity assays. siRNA, proliferation, anchorage-dependent, and -independent (soft agar) colony formation, migration (scratch), Matrigel invasion, and ChIP assays were performed as previously described (3, 4, 53), with additional details provided in Supplemental Methods. FACS was performed as previously described (33) using dual staining for CD24 and CD44 (Supplemental Table 8) with propidium iodide exclusion of nonviable cells. ALDH activity was measured using Aldefluor kits (Stem Cell Technologies Inc.) as described previously (33).

In vivo tumorigenicity assays and histologic analysis. Subcutaneous and intravenous xenograft growth were evaluated as previously described in female NOD/SCID mice (4, 53).

RNA isolation and qRT-PCR of mRNA expression. RNA isolation and qRT-PCR were performed as previously described (4). qRT-PCR was performed using validated TaqMan primers and probes (Supplemental Table 9) (Applied Biosystems), and relative expression was calculated using the $2^{-\Delta\Delta CT}$ method.

Microarray analysis. mRNA microarray analysis was performed as previously described (4) using Illumina HumanWG6 v3 and HumanHT-12 v4 Expression BeadChips (Illumina Inc.). All original microarray data were deposited in the NCBI's Gene Expression Omnibus (GEO GSE77925).

Statistics. Statistical tests were performed with GraphPad Prism v6.02 (GraphPad Software Inc.) and SPSS v19.0.0 (SPSS Inc.) software. Correlation coefficients were calculated using a Pearson or Spearman statistic, depending upon sample. Differences between means were calculated by 2-tailed *t* tests or 1-way ANOVA with Holm-Šidák multiple comparisons test. In vitro and in vivo growth curves were tested for significant difference in slopes using nonlinear regression model (least squares fit) with an exponential growth equation. Incucyte migration curves were compared using a nonlinear regression model (least squares fit) with a sigmoidal dose response (vari-

able slope). Kaplan-Meier survival plots and log-rank tests were used to assess differences in outcome. A *P* value of less than 0.05 was considered significant.

Study approval. All animal studies were approved by the Institutional Animal Care and Use Committee at the University of Texas Southwestern Medical Center, Dallas, Texas, USA.

Author contributions

All authors read the manuscript and provided feedback. JEL designed the research study, conducted experiments, acquired data, analyzed data, and wrote the manuscript. VN conducted experiments and acquired data. JKO conducted experiments, acquired data, and analyzed data. RKF conducted experiments and acquired data. DD conducted experiments and acquired data. PDD conducted experiments, acquired data, analyzed data, and edited the manuscript. JPS conducted experiments and acquired data. SKH conducted experiments, acquired data, and edited the manuscript. AA conducted experiments, acquired data, and edited the manuscript. MS provided reagents and edited the manuscript. LG analyzed data. CB provided reagents. IIW provided reagents. AFG provided reagents and designed the research study. NKH designed the research study and edited the manuscript. JDM designed the research study and edited the manuscript.

Acknowledgments

The authors wish to thank members of the Minna laboratory for their technical assistance, scientific discussion, and manuscript comments; Alex Pertsemliadis (Greehey Children's Cancer Research Institute at the University of Texas Health Science Center, San Antonio, Texas, USA) for scientific discussion; Thomas Brabletz for providing the pSuperRetroPuro-shZEB1 plasmid; Jennifer Richer (University of Colorado Health Sciences Center, Denver, Colorado, USA) for the ZEB1-pCIneo plasmid; and Rolf A. Brekken (University of Texas Southwestern Medical Center) for TGF- β monoclonal antibody (clone 1D11) and control antibody 13C4 monoclonal antibody. This work was supported by the NCI Lung Cancer SPORE (P50CA70907), CTD2N CA176284, NASA NSCOR (NNJ05HD36G), and Cancer Prevention and Research Institute of Texas (CPRIT) (RP120732). JEL was supported by a National Health and Medical Research Council of Australia Overseas-Based Biomedical Training Fellowship (494511), Thoracic Society of Australia and New Zealand/Allen & Hanburys Respiratory Research Fellowship, and Cancer Australia/Cure Cancer Australia Priority Driven Collaborative Cancer Research Scheme (1070515).

Address correspondence to: John D. Minna, Hamon Center for Therapeutic Oncology Research and the Simmons Comprehensive Cancer Center, The University of Texas Southwestern Medical Center, 6000 Harry Hines Blvd., Dallas, Texas 75390-8593, USA. Phone: 214.648.4900; E-mail: John.Minna@utsouthwestern.edu.

J.P. Sullivan's present address is: Verily Life Sciences, Mountain View, California, USA.

J.K. Osborne's present address is: Children's Hospital Boston, Boston, Massachusetts, USA.

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin*. 2012;62(1):10–29.
2. Larsen JE, Minna JD. Molecular biology of lung cancer: clinical implications. *Clin Chest Med*. 2011;32(4):703–740.
3. Sato M, et al. Multiple oncogenic changes (K-RAS(V12), p53 knockdown, mutant EGFRs, p16 bypass, telomerase) are not sufficient to confer a full malignant phenotype on human bronchial epithelial cells. *Cancer Res*. 2006;66(4):2116–2128.
4. Sato M, et al. Human lung epithelial cells progressed to malignancy through specific oncogenic manipulations. *Mol Cancer Res*. 2013;11(6):638–650.
5. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell*. 2009;139(5):871–890.
6. Sánchez-Tilló E, et al. EMT-activating transcription factors in cancer: beyond EMT and tumor invasiveness. *Cell Mol Life Sci*. 2012;69(20):3429–3456.
7. Massagué J. TGFbeta in Cancer. *Cell*. 2008;134(2):215–230.
8. Brabletz S, Brabletz T. The ZEB/miR-200 feedback loop--a motor of cellular plasticity in development and cancer? *EMBO Rep*. 2010;11(9):670–677.
9. Dohadwala M, et al. Cyclooxygenase-2-dependent regulation of E-cadherin: prostaglandin E(2) induces transcriptional repressors ZEB1 and snail in non-small cell lung cancer. *Cancer Res*. 2006;66(10):5338–5345.
10. Gemmill RM, et al. ZEB1-responsive genes in non-small cell lung cancer. *Cancer Lett*. 2011;300(1):66–78.
11. Fang Y, et al. Protein expression of ZEB2 in renal cell carcinoma and its prognostic significance in patient survival. *PLoS ONE*. 2013;8(5):e62558.
12. Zhang J, Lu C, Zhang J, Kang J, Cao C, Li M. Involvement of ZEB1 and E-cadherin in the invasion of lung squamous cell carcinoma. *Mol Biol Rep*. 2013;40(2):949–956.
13. Spaderna S, et al. The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer. *Cancer Res*. 2008;68(2):537–544.
14. Ohira T, et al. WNT7a induces E-cadherin in lung cancer cells. *Proc Natl Acad Sci U S A*. 2003;100(18):10429–10434.
15. Takeyama Y, et al. Knockdown of ZEB1, a master epithelial-to-mesenchymal transition (EMT) gene, suppresses anchorage-independent cell growth of lung cancer cells. *Cancer Lett*. 2010;296(2):216–224.
16. Rhim AD, et al. EMT and dissemination precede pancreatic tumor formation. *Cell*. 2012;148(1-2):349–361.
17. Danielpour D, Dart LL, Flanders KC, Roberts AB, Sporn MB. Immunodetection and quantitation of the two forms of transforming growth factor-beta (TGF-beta 1 and TGF-beta 2) secreted by cells in culture. *J Cell Physiol*. 1989;138(1):79–86.
18. Oida T, Weiner HL. Depletion of TGF-beta from fetal bovine serum. *J Immunol Methods*. 2010;362(1-2):195–198.
19. Palmer HG, et al. Vitamin D(3) promotes the differentiation of colon carcinoma cells by the induction of E-cadherin and the inhibition of beta-catenin signaling. *J Cell Biol*. 2001;154(2):369–387.
20. Larriba MJ, González-Sancho JM, Bonilla F, Muñoz A. Interaction of vitamin D with membrane-based signaling pathways. *Front Physiol*. 2014;5:60.
21. Soucek L, Helmer-Citterich M, Sacco A, Jucker R, Cesareni G, Nasi S. Design and properties of a Myc derivative that efficiently homodimerizes. *Oncogene*. 1998;17(19):2463–2472.
22. Zeller KI, et al. Global mapping of c-Myc binding sites and target gene networks in human B cells. *Proc Natl Acad Sci U S A*. 2006;103(47):17834–17839.
23. Agrawal P, Yu K, Salomon AR, Sedivy JM. Proteomic profiling of Myc-associated proteins. *Cell Cycle*. 2010;9(24):4908–4921.
24. Director's Challenge Consortium for the Molecular Classification of Lung Adenocarcinoma, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med*. 2008;14(8):822–827.
25. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511(7511):543–550.
26. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012;489(7417):519–525.
27. Yanagawa J, et al. Snail promotes CXCR2 ligand-dependent tumor progression in non-small cell lung carcinoma. *Clin Cancer Res*. 2009;15(22):6820–6829.
28. Aigner K, et al. The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity. *Oncogene*. 2007;26(49):6979–6988.
29. Ponta H, Sherman L, Herrlich PA. CD44: from adhesion molecules to signalling regulators. *Nat Rev Mol Cell Biol*. 2003;4(1):33–45.
30. Louderbough JM, Schroeder JA. Understanding the dual nature of CD44 in breast cancer progression. *Mol Cancer Res*. 2011;9(12):1573–1586.
31. Visvader JE, Lindeman GJ. Cancer stem cells: current status and evolving complexities. *Cell Stem Cell*. 2012;10(6):717–728.
32. Mani SA, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*. 2008;133(4):704–715.
33. Sullivan JP, et al. Aldehyde dehydrogenase activity selects for lung adenocarcinoma stem cells dependent on notch signaling. *Cancer Res*. 2010;70(23):9937–9948.
34. Jiang F, et al. Aldehyde dehydrogenase 1 is a tumor stem cell-associated marker in lung cancer. *Mol Cancer Res*. 2009;7(3):330–338.
35. Massagué J. TGFbeta signalling in context. *Nat Rev Mol Cell Biol*. 2012;13(10):616–630.
36. Chaffer CL, et al. Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell*. 2013;154(1):61–74.
37. Liu Y, et al. Zinc finger E-box binding homeobox 1 promotes invasion and bone metastasis of small cell lung cancer in vitro and in vivo. *Cancer Sci*. 2012;103(8):1420–1428.
38. Yang Y, et al. ZEB1 sensitizes lung adenocarcinoma to metastasis suppression by PI3K antagonism. *J Clin Invest*. 2014;124(6):2696–2708.
39. Gibbons DL, et al. Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. *Genes Dev*. 2009;23(18):2140–2151.
40. Ahn YH, et al. ZEB1 drives prometastatic actin cytoskeletal remodeling by down-regulating miR-34a expression. *J Clin Invest*. 2012;122(9):3170–3183.
41. Roche J, et al. Global Decrease of Histone H3K27 Acetylation in ZEB1-Induced Epithelial to Mesenchymal Transition in Lung Cancer Cells. *Cancers (Basel)*. 2013;5(2):334–356.
42. Horiguchi K, et al. TGF-beta drives epithelial-mesenchymal transition through deltaEF1-mediated downregulation of ESRP. *Oncogene*. 2012;31(26):3190–3201.
43. Warzecha CC, et al. An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. *EMBO J*. 2010;29(19):3286–3300.
44. De Craene B, Bex G. Regulatory networks defining EMT during cancer initiation and progression. *Nat Rev Cancer*. 2013;13(2):97–110.
45. Brown RL, et al. CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. *J Clin Invest*. 2011;121(3):1064–1074.
46. Ko YH, et al. Prognostic significance of CD44s expression in resected non-small cell lung cancer. *BMC Cancer*. 2011;11:340.
47. Alameer M, Peacock CD, Matsui W, Ganju V, Watkins DN. Cancer stem cells in lung cancer: Evidence and controversies. *Respirology*. 2013;18(5):757–764.
48. Liu S, Clouthier SG, Wicha MS. Role of microRNAs in the regulation of breast cancer stem cells. *J Mammary Gland Biol Neoplasia*. 2012;17(1):15–21.
49. Harney AS, Meade TJ, LaBonne C. Targeted inactivation of Snail family EMT regulatory factors by a Co(III)-Ebox conjugate. *PLoS ONE*. 2012;7(2):e32318.
50. Ramirez RD, et al. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. *Cancer Res*. 2004;64(24):9027–9034.
51. Phelps RM, et al. NCI-Navy Medical Oncology Branch cell line data base. *J Cell Biochem Suppl*. 1996;24:32–91.
52. Greer RM, et al. SMAC mimetic (JP1201) sensitizes non-small cell lung cancers to multiple chemotherapy agents in an IAP-dependent but TNF-alpha-independent manner. *Cancer Res*. 2011;71(24):7640–7648.
53. Osborne JK, et al. NeuroD1 regulates survival and migration of neuroendocrine lung carcinomas via signaling molecules TrkB and NCAM. *Proc Natl Acad Sci U S A*. 2013;110(16):6524–6529.