# Identification of Endoglin as an epigenetically regulated tumour-suppressor gene in lung cancer

K O'Leary[1], A Shia[1,2], F Cavicchioli[1], V Haley[1], A Comino[3], M Merlano[4], F Mauri[5], K Walter[6], M Lackner[6], M B Wischnewsky[7], T Crook[8], C Lo Nigro[9] and P Schmid*,[1,2]

[1]Brighton and Sussex Medical School, University of Sussex, Brighton BN1 9RY, UK; [2]Barts Cancer Institute, Queen Mary University of London, Old Anatomy Building, Charterhouse Square, London EC1M 6BQ, UK; [3]Pathology Department, S. Croce General Hospital, via Coppino 26, 12100, Cuneo, Italy; [4]Medical Oncology, Oncology Department, S. Croce General Hospital, via Carle 25, 12100, Cuneo, Italy; [5]Department of Histopathology, Imperial College London, Hammersmith Campus, Du Cane Road, London, W12 0HS, UK; [6]Oncology Biomarker Development, Genentech, Inc., 550 Grandview Boulevard, South San Francisco, CA 94080, USA; [7]eScience Lab, Department of Biomathematics, University of Bremen, Bremen 28359, Germany; [8]Division of Cancer Research, Medical Research Institute, Jacqui Wood Cancer Centre, University of Dundee, Ninewells Hospital And Medical School, Dundee DD1 9SY, UK and [9]Laboratory of Cancer Genetics and Translational Oncology, Oncology Department, S. Croce Genreal Hospital, via Carle 25, Cuneo 12100, Italy

**Background:** The transforming growth factor-beta (TGF-$\beta$) pathway has been implicated in proliferation, migration and invasion of various cancers. Endoglin is a TGF-$\beta$ accessory receptor that modulates signalling. We identified Endoglin as an epigenetically silenced tumour-suppressor gene in lung cancer by means of a genome-wide screening approach, then sought to characterise its effect on lung cancer progression.

**Methods:** Methylation microarray and RNA sequencing were carried out on lung cancer cell lines. Epigenetic silencing of Endoglin was confirmed by methylation and expression analyses. An expression vector and a 20-gene expression panel were used to evaluate Endoglin function. Pyrosequencing was carried out on two independent cohorts comprising 112 and 202 NSCLC cases, respectively, and the impact of Endoglin methylation on overall survival (OS) was evaluated.

**Results:** Methylation in the promoter region resulted in silencing of Endoglin, which could be reactivated by demethylation. Increased invasion coupled with altered EMT marker expression was observed in cell lines with an epithelial-like, but not those with a mesenchymal-like, profile when Endoglin was absent. Methylation was associated with decreased OS in stage I but not in stages II–III disease.

**Conclusions:** We show that Endoglin is a common target of epigenetic silencing in lung cancer. We reveal a link between Endoglin silencing and EMT progression that might be associated with decreased survival in stage I disease.

The importance of epigenetic transcriptional silencing as a mechanism of tumour-suppressor gene inactivation is well recognised (Baylin and Herman, 2000; Teodoridis *et al.*, 2004; Esteller, 2007). In human cancers, aberrant methylation of genes is a far more frequent event than genetic mutation (Taby and Issa, 2010), therefore it represents a valuable source of new biomarkers and therapeutic targets. We identified Endoglin (CD105) by means of a combined screen to identify novel epigenetically regulated genes in lung cancer using methylation microarrays and RNA sequencing in a panel of lung cancer cell lines.

The Endoglin gene is located on chromosome 9q34.11 (Fernandez-Ruiz *et al.*, 1993) and the protein functions as an accessory receptor within the TGF-$\beta$ signalling pathway (Massague, 1998). Along with Betaglycan, it belongs to a class of TGF-$\beta$ receptors known as the accessory receptors. These influence TGF-$\beta$ signalling by regulating ligand access to type I and type II receptors (Massague, 1998), resulting in an altered phosphorylation state which facilitates ligand access and transduction of the intracellular signal (Guerrero-Esteo *et al.*, 2002).

Deregulation of TGF-$\beta$ signalling can result in chronic inflammation and fibrosis (Bartram and Speer, 2004) and has a vital role in cancer progression through enabling epithelial-to-mesenchymal transition (EMT) and metastatic colonisation (Meulmeester and Ten Dijke, 2011; Wendt *et al.*, 2011). However, it has also been shown to carry out tumour-suppressor functions (Siegel and Massague, 2003; Meulmeester and Ten Dijke, 2011), for example, via repression of the oncogenic Myc transcription factor (Meulmeester and Ten Dijke, 2011). The dual functions of this signalling pathway have confounded investigations into therapeutic targeting of the TGF-$\beta$ pathway (Gazdar *et al.*, 2004; Hawinkels and Ten Dijke, 2011; Smith *et al.*, 2012). Therefore, further research is warranted to illuminate the specific environmental cues which regulate these dual functions.

The role of Endoglin at a molecular level has, for the most part, only been examined in endothelial cells, where it has a crucial role in angiogenesis and therefore tumour progression (Li *et al.*, 1999; Raab *et al.*, 1999; Jonker and Arthur, 2002; Park *et al.*, 2013). However, there is evidence for expression of Endoglin in non-endothelial cells (Kassouf *et al.*, 2004; Postiglione *et al.*, 2005) and for an angiogenesis-independent role for Endoglin in tumour suppression. Endoglin-mediated suppression of invasion and migration has been shown in prostate cancer (Liu *et al.*, 2002; Craft *et al.*, 2007), oesophageal squamous cell carcinoma (Wong *et al.*, 2008), extravillous trophoblasts (Mano *et al.*, 2011), HER2-positive breast cancer cells (Henry *et al.*, 2011) and mouse keratinocyte cells (Perez-Gomez *et al.*, 2007). Given the role of Endoglin as a TGF-$\beta$ accessory receptor and the paradoxical effects of the TGF-$\beta$ pathway on tumour growth, we sought to investigate the epigenetic regulation of Endoglin in NSCLC and to define its function within these cancers.

## MATERIALS AND METHODS

**Cell culture and reagents.** Cell lines subjected to genome-wide screening were maintained and analysed at Genentech, Inc. (San Francisco, CA, USA) as previously described (Walter *et al.*, 2012). All other cell lines were provided by Michael Seckl (Imperial College London, London, UK). Authentication was carried out by STR profiling (Promega PowerPlex 18 System, Promega Corporation, Madison, WI, USA) on cell lines used in functional assays for which reference STR profiles were available (A549, HOP-62, H23, H322M). The cell lines A549, HCC193 and HCC95 were maintained in DMEM (PAA Laboratories, Teddington, UK) and all others in RPMI (PAA Laboratories) supplemented with 2 mM L-glutamine, 50 $\mu$g ml$^{-1}$ streptomycin, 0.5 Units ml$^{-1}$ penicillin and 10% heat inactivated foetal calf serum (Gibco, Paisley, UK) in a 37 °C humidified chamber with 5% $CO_2$ and 95% humidity. Transfected cell lines were maintained in medium containing G418 (Sigma, Dorset, UK) as follows: 300 $\mu$g ml$^{-1}$ for HCC95 and H23, 500 $\mu$g ml$^{-1}$ for A549, H322M and MOR.

**Genome-wide methylation and expression analysis.** Genome-wide screening was carried out at Genentech, Inc. Cell lines were profiled using the Infinium HumanMethylation450 Beadchip Kit (Illumina Inc., San Diego, CA, USA) as previously described (Walter *et al.*, 2012). Array data have been submitted to the Gene

Expression Omnibus database (accession number GSE36216). A single probe mapped to the short CpG island located within the Endoglin gene promoter region. Beta values for probes mapping to CpG island 'shore areas' (regions flanking CpG islands, up to 2 kb up and downstream) were also analysed, as mounting evidence suggests that shore areas may also directly affect gene expression (Irizarry *et al.*, 2009; Rao *et al.*, 2013).

For expression analysis, NSCLC cell line RNA was used to generate RNA-seq libraries using the TruSeq RNA Sample Preparation Kit (Illumina Inc.). Processed libraries were sequenced on the Illumina HiSeq 2000 Platform according to the manufacturer's recommendation (Illumina Inc.). In all, $75 \times 75$ base pair (bp) paired-end data were collected, and RNA-seq reads were aligned to the human genome version GRCh37 using GSNAP (Wu and Nacu, 2010). All short-read sequencing data were evaluated for quality control using the R Bioconductor Shortread package. To normalise for coverage depth, Reads Per Kilobase of transcript per Million mapped reads (RPKM) values were generated. RPKM expression values for Endoglin were then visualised in Excel (Microsoft, Redmond, WA, USA) and compared against methylation beta values at the Infinium probe site within the Endoglin gene.

**DNA extraction and bisuphfite conversion.** DNA was extracted from cell lines using the DNeasy Blood and Tissue Kit (Qiagen, Manchester, UK) as per manufacturer's instructions. DNA from formalin-fixed, paraffin-embedded (FFPE) tissues was extracted using an optimised protocol. Briefly, tissue slices were deparaffinised in xylene overnight and then washed twice in ethanol. Tissue was digested by incubating in 0.2 mg ml$^{-1}$ proteinase K in lysis buffer (300–500 $\mu$l) at 55 °C overnight. DNA was purified by rescuing the aqueous phase using phenol saturated with tris, followed by precipitation in isopropanol, then ethanol and finally resuspension in ddH$_2$O. Bisulphite modification of DNA was carried out using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA, USA) exactly as per the manufacturer's instructions, using either 500 ng of cell line DNA or 1 $\mu$g of FFPE tissue DNA.

**Methylation-specific PCR (MSP).** PCR was carried out using ThermoStart *Taq* PCR Master Mix (Thermo Scientific, Hemel Hempstead, UK) with addition of 0.5 $\mu$M primers and 10–20 ng bisulphite-converted DNA in a 20 $\mu$l reaction. Commercially available unmethylated and methylated DNA (Millipore, Watford, UK) was included in all bisulphite conversion and PCR steps as a control. See Supplementary Methods for all PCR cycling conditions.

**Pyrosequencing.** DNA was subjected to bisulphite conversion as described above, and the region to be sequenced was amplified by PCR (50 ng cell line/control DNA or 200–300 ng FFPE tissue DNA). PCR was carried out using the Pyromark PCR Kit (Qiagen) with the addition of 0.5 $\mu$M primers in a 30 $\mu$l reaction. In all, 3 $\mu$l of PCR product was resolved on 2% agarose gels to confirm the success of the reaction prior to sequencing. The biotinylated PCR products were captured by binding to streptavidin-coated sepharose beads and used as the template. Using the PyroMark Q96 Vacuum Prep Workstation (Qiagen), streptavidin-bound PCR products were immobilised, washed, denatured and allowed to anneal to the sequencing primer for 2 min at 80 °C. The non-biotinylated PCR primer was used as the sequencing primer and was extended by addition of nucleotides by the PyroMark Q96 ID pyrosequencer (Qiagen). Methylation at each CpG dinucleotide was calculated from the ratio of C to T in the bisulphite-modified sequence.

**Reverse transcription and semi-quantitative real-time PCR (qRT-PCR).** RNA was extracted using a phenol-based protocol. RNA (100 ng) was reverse transcribed into cDNA in a 20 $\mu$l reaction using the High Capacity cDNA Kit (Applied Biosystems, Foster City, CA, USA), as per the manufacturer's instructions. All

qRT-PCR reactions were conducted in triplicate. Commercially available Taqman gene expression assays (Applied Biosystems) together with Brilliant II QPCR Master Mix with low ROX (Agilent Technologies, Berkshire, UK) were used for each gene and analysed on the Stratagene Mx3000P QPCR System (Agilent Technologies). RPLP0 was used as a reference gene. Taqman expression assays were Hs00923996_m1 for Endoglin and Hs02992885_s1 for RPLPO.

**Western blotting.** Lysates were diluted in $2 \times$ Laemmli buffer with 10% $\beta$-mercaptoethanol and denatured for 10 min at 90 °C. Western blotting was performed using standard techniques. Primary antibodies used were mouse anti-CD105 (611314, BD Biosciences, Oxford, UK) at 1:600 dilution and rabbit anti-$\beta$-actin (4970L, Cell Signalling, Leiden, Netherlands) at 1:5000 dilution. Secondary antibodies were HRP-linked anti-mouse IgG (7076S, Cell Signalling) and HRP-linked anti-rabbit (7074S, Cell Signalling).

**Immunohistochemistry.** Endoglin expression was evaluated on TMA sections using the monoclonal antibody SN6h (Dako, Glostrup, Denmark) at 1:10 dilution, using a modified version of the antigen retrieval protocol described elsewhere (Shi et al., 1991). Briefly, the sections were rehydrated and heated in a microwave oven at 900 W for 20 min in Citrate buffer at pH 6. Primary antibody was used at 1:10 dilution for 1 h at room temperature and then processed with Polymer-HRP Kit (BioGenex, San Ramon, CA, USA) with diaminobenzidine development and Mayer haematoxylin counterstaining. Endothelial cells were used as a positive internal control. Negative controls were obtained by omitting the primary antibody. Staining was assessed using the semi quantitative immunohistochemical score method. Intensity of staining and the percentage of positive cells were assessed as previously described (Dhillon et al., 2010).

**Methylation reversal assays.** Cells were grown in culture medium until 60% confluent. Fresh medium containing 5 $\mu$M of 5-aza-2′-deoxycytidine (AZA) (Sigma) was added, and cells were incubated for 5 days. Trichostatin A (0.3 nM; TSA; Sigma) was added to media for the final 16 h prior to harvest. On day 5, cells were counted, washed in PBS and pelleted at $1 \times 10^6$ cells per pellet and stored at $-20$ °C until DNA, RNA and protein extraction.

**Transfections.** For siRNA transfection, 10 nM of an Endoglin-targeting siRNA (SI02663024, Qiagen) or a non-silencing siRNA control (Qiagen) were transfected into cells using Lipofectamine RNAiMax (Invitrogen). Culture medium was changed after 24 h. For ectopic expression of Endoglin, cells were transfected with purified pcDNA3.1 alone or pcDNA3.1 containing the long isoforms Endoglin cDNA (a gift from Professor Clare Isacke, Institute of Cancer Research, UK). Plasmid transfections were carried out using X-tremeGENE HP DNA Transfection Reagent (Roche, Welwyn Garden City, UK) as per the manufacturer's instructions.

**Cell-based assays.** Where indicated, TGF-$\beta$1 (Peprotech, Rocky Hill, NJ, USA) was added at a concentration of 10 ng ml$^{-1}$ and the TGF-$\beta$ receptor inhibitor SB 431542 (Sigma) at 20 $\mu$M. Cell viability and proliferation were measured using the MTT (3-[4,5-dimethylthiazol-2-yl]-2,5 diphenyl tetrazolium bromide) assay. Data are represented as background-corrected raw values and normalised to a control sample. For scratch wound-healing assays, cells were serum-starved overnight and scratched across the diameter of the well with a sterile pipette tip in the morning. Cells were washed twice with PBS and culture medium with or without TGF-$\beta$1 (10 ng ml$^{-1}$) was added. Scratch wounds were photographed immediately and after 16–30 h depending on speed of migration. Distance migrated was quantified using the ImageJ software, version 1.45 (US National Institutes of Health, Bethesda, MD, USA). For the tumour spheroid-based invasion assay, cell suspension were seeded into ultra-low attachment 96-well round-bottomed plates (Corning, Boston, MA, USA) and incubated for 4 days at 37 °C, 5% CO$_2$, 95% humidity. On day 4, 70 $\mu$l of medium was removed without disrupting the sphere and carefully replaced with 70 $\mu$l of Matrigel (BD Biosciences) with TGF-$\beta$ or inhibitor as indicated above. After 4 days, spheroids were photographed at $\times 40$ magnification using an inverted microscope (Olympus CKX-41, Center Valley, PA, USA) equipped with a digital camera (Micropix, Midhurst, UK). The BD BioCoat Tumor Invasion System (BD Biosciences) was used as per the manufacturer's instructions to measure invasion of the HCC95 cell line as these could not be induced to form spheroids. Data are represented as fluorescence at 24 h − fluorescence at time zero for each well. Technical replicates ($n=4$) are then averaged and normalised to the untreated empty vector control.

**EMT gene expression analysis.** To determine the expression levels of 20 EMT-related genes (Walter et al., 2012), cell line cDNA was analysed using a Taqman Low Density Array (TLDA, Applied Biosystems) at Genentech. Total cDNA (200 ng) was amplified using Taqman Universal PCR MasterMix (Applied Biosystems) on a TLDA card. All samples were assayed in triplicate. Two custom-designed reference genes, AL-1377271 and VPS-33B, and a commercial reference gene assay, 18S (Hs99999901_s1, Applied Biosystems), were also included in the panel. A mean of the Ct values for the three reference genes was calculated for each sample, and the expression levels of EMT target genes were determined using the $2^{-\Delta CT}$ method. dCt values were used to cluster cell lines according to EMT gene expression levels using the Cluster v.3.0 and Treeview v.1.60 softwares (Eisen Lab, University of California, Berkeley, CA, USA; http://rana.lbl.gov/EisenSoftware.htm).

**Primary tissues.** One hundred and twelve lung cancer tissues, 20 tumour-adjacent tissues and 10 biopsy tissues from cancer-free emphysema patients were obtained as FFPE specimens with informed patient consent and local ethics committee approval from the S. Croce General Hospital, Cuneo, Italy and are referred to in the text as 'cohort 1'. Tissues were from surgical resection samples of NSCLC patients. Two hundred and two lung cancer tissues and 20 tumour-adjacent tissues from biopsies or surgical resection were obtained as FFPE sections with informed patient consent and local ethics committee approval from Imperial College Healthcare NHS Trust, London. These are referred to in the text as 'cohort 2'. In all cases, the diagnosis and presence of adequate tumour representation in the specimen was confirmed by histopathological analysis. Patient characteristics are summarised in Supplementary Table S1.

**Statistical analysis.** Associations between Endoglin methylation status and clinicopathological characteristics were analysed using the Chi-squared test, Fishers exact test or Kendall's tau-$b$ statistic, where appropriate. Kaplan–Meier analysis was used to compare survival of methylated and unmethylated populations. Multivariate Cox regression analysis was carried out to estimate hazard ratios for the effect of Endoglin methylation together with potential confounding variables on overall survival. Differences in functional traits of cell lines were analysed by one or two-way analysis of variance, as appropriate. In all cases, $P$-values $<0.05$ were considered significant. Analyses were carried out using the IBM SPSS Statistics 21 software (IBM Corporation, Armonk, NY, USA) or GraphPad Prism 6 software (GraphPad Software, Inc., La Jolla, CA, USA).

## RESULTS

**Endoglin is epigenetically downregulated in lung cancer cell lines.** Endoglin was identified in a genome-wide screen for novel epigenetically regulated genes in lung cancer using methylation

microarrays (Illumina Human 450K Methylation BeadChip) coupled with RNA-seq (Illumina Inc.) in a panel of 65 lung cancer cell lines. A single methylation array probe mapped to the Endoglin CpG island (Figure 1A) and displayed high methylation values ($\beta$-value $> 0.5$) for more than half of the cell line panel. These highly methylated cell lines showed minimal basal expression of the gene, while cell lines with lower methylation values showed higher expression levels (Figure 1B). Based on this, Endoglin was selected for validation and further investigation.

Methylation of Endoglin was confirmed in a sub-panel of 18 lung cancer cell lines using pyrosequencing (Figure 2B) and MSP (Figure 2C), which correlated for all cell lines. Endoglin expression was confirmed by qPCR quantification of mRNA and western blotting quantification of the 98 kDa Endoglin protein. Protein levels were found to reflect mRNA levels, and both RNA and protein levels were negligible in cell lines with methylation $> 50\%$ (Figures 3A and B).

To validate epigenetic mechanisms as a direct cause of Endoglin transcriptional silencing, we performed methylation reversal assays using the demethylating agent AZA and/or the HDAC inhibitor Trichostatin A (TSA). In the highly methylated cell lines MOR and H322M, removal of epigenetic marks resulted in significant re-expression of Endoglin mRNA and protein. Re-expression was seen upon treatment with AZA and AZA combined with TSA, although no synergistic effect of the combined treatment was observed. Treatment with TSA alone had no effect on Endoglin expression (Figures 3C and D). In contrast, treatment with AZA and/or TSA had no effect in cell lines without Endoglin methylation (Figures 3C and D, HOP-62 cell line), confirming that the effects of AZA on Endoglin expression are a result of the direct effect on the methylation status of the promoter region. Together, these findings provide conclusive evidence that methylation-dependent epigenetic regulation is a common mechanism of Endoglin silencing in lung cancer cell lines.

**Loss of Endoglin drives EMT progression in lung cancer cell lines with an epithelial-like, but not a mesenchymal-like, expression signature.** Next, we sought to determine the functional

significance of Endoglin in lung cancer. We hypothesised that loss of Endoglin may affect TGF-$\beta$ signalling. As this pathway involves a delicate balance of pro- and anti-invasion, migration and proliferation signals, a range of assays were carried out to quantify these functional traits, including MTT assays for cell viability and proliferation, 3D spheroid models or transwell assays to study invasion and scratch wound-healing assays to analyse migration. Assays were performed with and without addition of TGF-$\beta$1 and/or a TGF-$\beta$ receptor inhibitor and Endoglin levels were modified using ectopic expression and RNA interference, respectively. The following cell lines were selected: the highly methylated and Endoglin-silenced NSCLC cell lines HCC95, H23, A549 and H322M with and without transfection with a pcDNA3.1 plasmid to ectopically express the gene, as well as the unmethylated and Endoglin expressing cell line HOP-62, with and without transient silencing with an Endoglin-targeting siRNA.

We identified two distinct phenotypes within the panel of transfected cell lines. The first phenotype, which includes the cell lines H322M and HCC95, refers to cells in which absence of Endoglin leads to increased invasion coupled with decreased migration and proliferation (Figures 4A and B). In contrast, in the second group which included the cell lines H23, A549 and HOP-62, Endoglin expression had no effect on migration, invasion or proliferation either with or without TGF-$\beta$. Representative data for the cell line H23 is shown in Figure 4C, data for additional cell lines are available in Supplementary Figure S1.

As Endoglin is involved in TGF-$\beta$ signalling, a pathway known to be capable of stimulating EMT (Zavadil and Bottinger, 2005; Kalluri and Weinberg, 2009; Wendt *et al.*, 2011), we hypothesised that loss of Endoglin may be driving EMT progression. If this is the case, cells that are epithelial-like could be induced by loss of Endoglin to undergo EMT while, conversely, cells that have already transitioned to a mesenchymal profile would not be affected. To address this, cell lines were analysed using an optimised 20-gene EMT marker expression panel (Walter *et al.*, 2012) to classify their EMT profile. EMT marker expression profiles were generated by cluster analysis following quantification of EMT marker expression. Expression data for the cell lines mentioned above
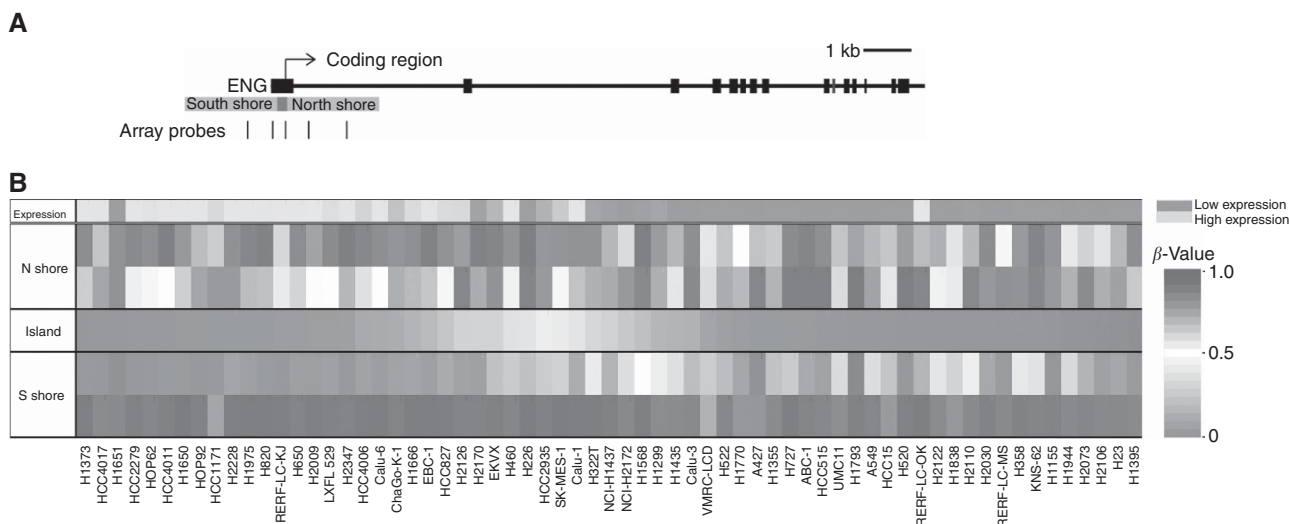


Figure 1. A genome-wide screen identifies Endoglin as a candidate epigenetically regulated gene in lung cancer. (**A**) Representation of the Endoglin gene with location of methylation array probes. Thick black horizontal lines represent exons. The location of the CpG island is shown in dark grey, with the shore areas in light grey. Location of methylation array probes that target a single CpG dinucleotide are shown underneath as vertical black lines. (**B**) Heat map representation of microarray data showing Endoglin expression and methylation levels in a panel of lung cancer cell lines. Heat map shows methylation as beta values, calculated from the ratio of methylated signal to total signal. Cell lines are arranged in order of increasing methylation of the array-targeted CpG site within the Endoglin CpG island. Expression is represented as reads per kilobase per million (RPKM) from high-throughput RNA-seq (Illumina Inc.). RPKM values are colour-coded from lowest (green) to highest (yellow). The full colour version of this figure is available at *British Journal of Cancer* online.
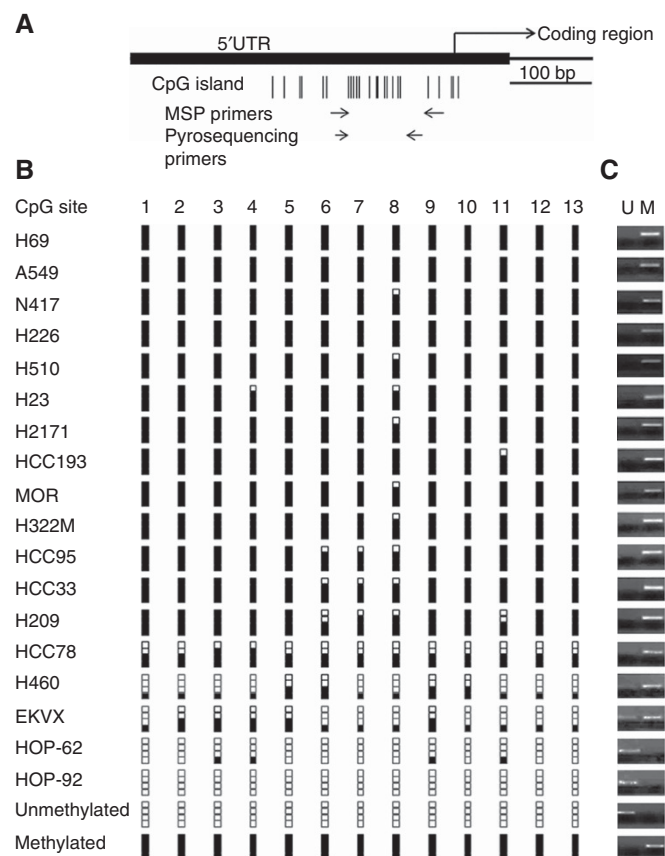
**Figure 2. Methylation analysis of the Endoglin promoter region.**
(**A**) Representation of the Endoglin gene with CpG island and primer locations. Individual CpG sites within the island are represented by vertical lines. The island overlaps the 5′UTR of the gene and the beginning of the coding region. Forward and reverse primers are represented by horizontal arrows. (**B**) Endoglin methylation as measured by pyrosequencing. Methylation levels for cell lines are represented as percentage of methylation at each of 13 CpG sites as follows: 0–5%: zero black squares; 6–25%: one black square; 26–50%: two black squares; 51–75%: three black squares; and 76–100%: four black squares. (**C**) Endoglin methylation measured by methylation-specific PCR. M and U refer to PCR reactions carried out with primers targeting methylated and unmethylated sequence, respectively. Results are shown as PCR products of both reactions for each cell line.

were combined with an extended panel of 80 cell lines published previously (Walter *et al.*, 2012; data not shown) for cluster analysis.

As expected, cell lines showing no Endoglin-dependent functional changes (H23, A549 and HOP-62, referred to as Group 2) displayed a mesenchymal-like profile while cell lines that showed an Endoglin-dependent invasive phenotype (HCC95 and H322M, referred to as Group 1) are epithelial-like (Figure 5A). Analysis of individual EMT marker genes in representative cell lines confirmed that absence of Endoglin in epithelial-like cells induces pro-EMT expression changes (Figure 5B), while having no effect on marker expression in cells that are mesenchymal (Figure 5C).

Together, these findings suggest a tumour-suppressing effect of Endoglin expression in tumours that have not undergone EMT. We therefore hypothesised that transcriptional silencing of Endoglin may be associated with increased invasion and consequently an increased risk of disease recurrence in patients with NSCLC. As EMT is linked with increased loco-regional and distant metastatic spread, we also hypothesised that the effects of Endoglin on patient outcome should be most relevant for patients with early-stage disease where EMT is less likely to have taken place.

**Endoglin is methylated in primary NSCLC and predicts poor clinical outcome in early-stage disease.** To test these hypotheses, Endoglin methylation was analysed in two independent, single-centre cohorts of NSCLC patients who underwent complete surgical resection. Cohort 1 comprised of only patients with stage I disease, whereas cohort 2 included patients with stages I–III NSCLC. The median follow-up is 7.24 years for cohort 1 and 1.38 years for cohort 2.

FFPE tissues from the resected primary tumours were analysed for methylation using the pyrosequencing assay validated in our cell line panel. As with cell lines, methylation of primary tissues was calculated as the average across 13 CpG sites within the amplicon generated by the pyrosequencing primers. A panel of non-tumour lung tissue samples were used as controls for the definition of a cutoff for methylation in tumour samples. For each cohort, the cutoff was defined as the average of control tissues plus 3 s.ds.

Applying this to our panel of tumour samples, cutoff values of 47% and 32% methylation were used for cohorts 1 and 2, respectively. This led to 42% of cohort 1 and 23% of cohort 2 being classified as methylated for Endoglin. Methylation-dependent silencing of Endoglin expression was confirmed by immunohistochemistry in cohort 1; representative samples are shown in Figures 6D–G. Analysis of the clinicopathological parameters suggested that Endoglin methylation might be correlated with gender (higher incidence in male patients) and squamous cell histology, but results were inconsistent across both cohorts (Supplementary Table S1).

Kaplan–Meier analysis of both cohorts by stage suggested a stronger effect of Endoglin on survival for patients with stage I disease (Figure 6A) compared with stages II–III disease (Figure 6B), although statistical significance was not reached. A multivariate Cox proportional hazards analysis was carried out to control for potential confounding factors, such as age, gender, grade, stage and histology. In this model, Endoglin was found to be a significant predictor of overall survival in stage I but not stages II–III cancers (Figure 6C). Stage 1 disease is characterised by the lack of metastatic lymph node involvement, whereas stage II or III disease commonly involves metastatic infiltration of loco-regional lymph nodes. Therefore, our findings are in keeping with the hypothesis that Endoglin methylation is of prognostic relevance in early-stage cancers that might not have undergone EMT.

## DISCUSSION

We have shown that methylation-dependent transcriptional silencing of Endoglin is a common event in NSCLC and is linked with an increased risk of disease recurrence in early disease stages. Our data also demonstrate that loss of Endoglin may induce a pro-invasive, antiproliferative phenotype together with EMT progression. These findings are supported by a number of published studies in other cancer types (Liu *et al.*, 2002; Craft *et al.*, 2007; Perez-Gomez *et al.*, 2007; Henry *et al.*, 2011; Mano *et al.*, 2011), which found increased invasion upon loss of Endoglin expression, while proliferation was either unaffected or not shown. However, while these studies also demonstrated increased cell migration, we observed a decrease that may be a reflection of a lower proliferation rate.

Indeed, reduced proliferation with acquired invasiveness has previously been reported although not in the context of Endoglin silencing. Studies on basal cell carcinoma and colorectal cancer primary tissues have suggested that infiltrative cells at the invasive front of the tumour show a lack of proliferation mediated by upregulation of the cell cycle inhibitory protein p16 (Palmqvist *et al.*, 2000; Svensson *et al.*, 2003). A similar phenotype was observed in a mouse *in vivo* model of glioblastoma, whereby
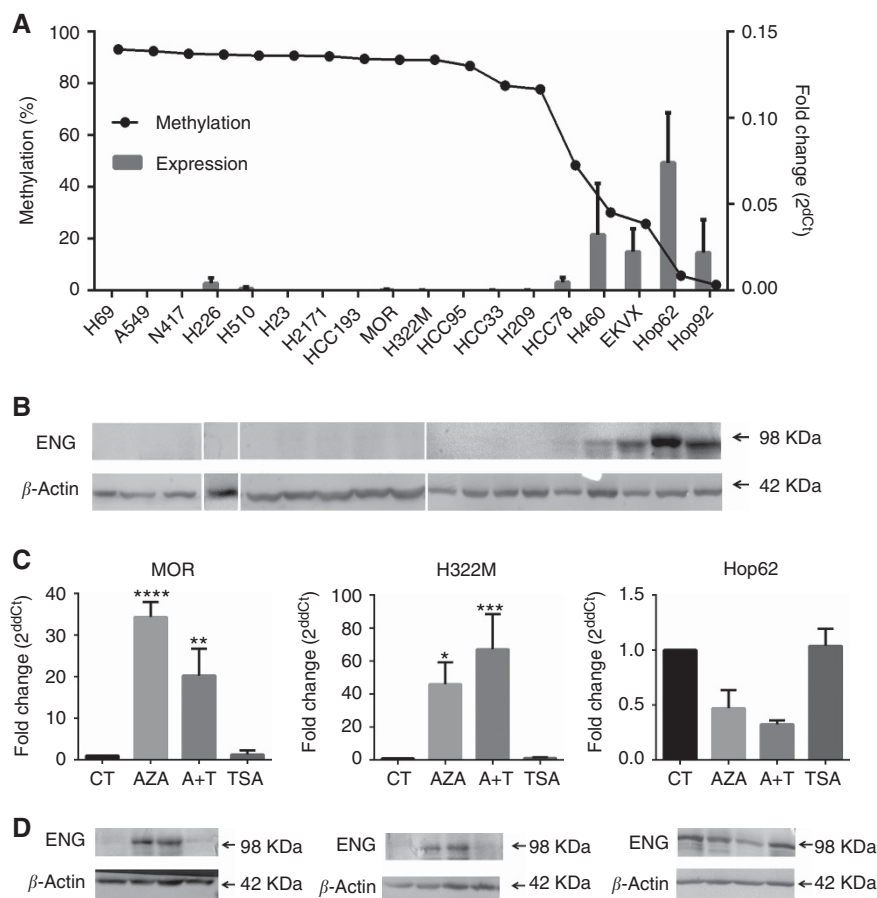
Figure 3. Methylation of Endoglin leads to gene silencing, which can be re-activated by demethylation. (**A**) Relative Endoglin mRNA expression was determined by qRT-PCR. Histogram bars show expression as the mean ± s.d. of three biological replicates, normalised to the reference gene (right y axis). Cell lines are shown in order of decreasing methylation by pyrosequencing (mean of three biological replicates) as indicated by the dotted line (left y axis). (**B**) Endoglin protein level was measured by western blotting. Blot shows cell lines in the same order as histogram above, with β-actin as a reference for each sample. (**C**) The cell lines MOR and H322M are methylated within the Endoglin promoter region and show significant upregulation of expression following treatment with the demethylating agent AZA. No upregulation was observed in the unmethylated cell line HOP-62. No effect of TSA was observed. (**D**) Protein levels are also upregulated upon demethylation treatment in the cell lines MOR and H322M. Blots show samples in the same order as histogram above, with β-actin as a reference protein. All experiments were carried out in triplicate.

invasive cells acquired stem-cell-like properties, this time mediated by a decrease in Erk and an increase in Akt activation (Molina et al., 2010). We assessed p16 levels in our cell line panel but found no evidence that p16 might be a mediator of growth inhibition in Endoglin-silenced cells (data not shown). Investigation of a wider panel of cell cycle regulators would be of interest but was beyond the scope of the current study.

Data presented here suggest that EMT status of the tumour may be a factor in determining whether or not Endoglin methylation has a functional and prognostic effect in NSCLC. The process of EMT occurs along a spectrum of changes, whereby there is a gradual reduction in expression of epithelial markers and an increase in mesenchymal markers (Kalluri and Weinberg, 2009). Therefore, it is reasonable to hypothesise that, as loss of Endoglin may drive EMT, cells already mesenchymal in their molecular profile would not be affected by Endoglin expression levels. Consequently, as EMT has a significant role in progression of early-stage disease one would expect transcriptional silencing of Endoglin to be most relevant in tumours that have not fully undergone EMT. Our translational data clearly support this hypothesis. We show that Endoglin methylation confers worse prognosis in stage I cancers, which by definition have not yet spread to lymph nodes, and therefore may be in the process of transitioning towards a more invasive phenotype if Endoglin is silenced. A prognostic panel that includes Endoglin methylation and EMT marker expression would therefore be worth evaluating in early-stage NSCLC and may help to characterise patients who could benefit from adjuvant therapy. More specifically, TGF-β-targeted therapies may hold potential for these patients in whom Endoglin silencing is conferring a pro-invasive function of TGF-β signalling.

In summary, we have confirmed epigenetic silencing of Endoglin and shown an association between Endoglin methylation and decreased survival in stage I NSCLC. Moreover, we have revealed a link between Endoglin silencing and EMT progression. Further research is warranted to further explore this mechanistic relationship and its potential as a prognostic marker and/or therapeutic target.
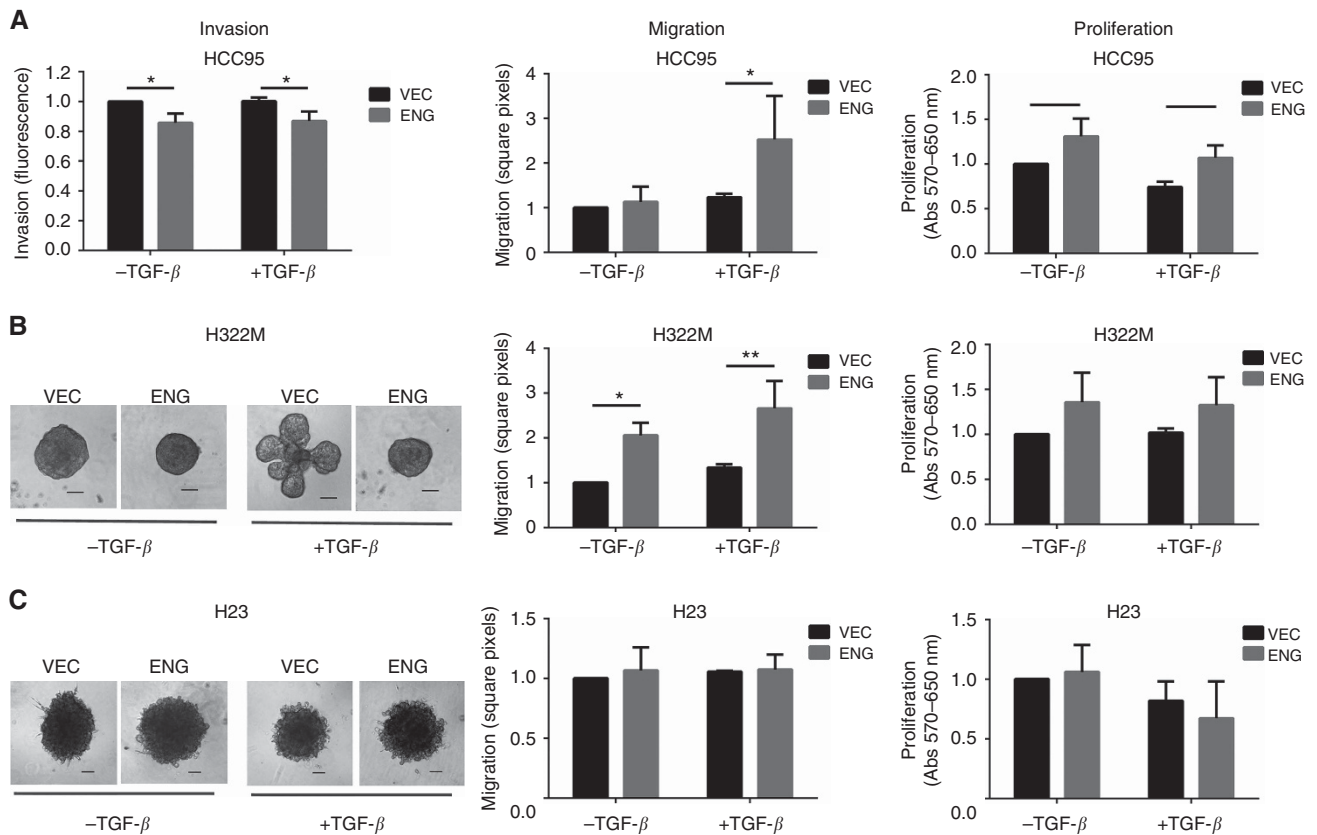
**Figure 4. Loss of Endoglin leads to increased invasion and decreased migration/proliferation.** Invasion was analysed using a 3D spheroid invasion model or transwell invasion assay (HCC95 cells only, as these could not be induced to form spheroids in 3D culture). Migration was analysed by scratch wound-healing assay over 24 h and proliferation by MTT over 5 days. (**A**, **B**) The cell lines HCC95 and H322M show increased invasion in VEC compared with ENG cells, together with decreased migration and proliferation, (**C**) while no effect was seen in the H23 cell line. Where indicated, TGF-$\beta$1 was added at 10 ng ml$^{-1}$. As Matrigel naturally contains TGF-$\beta$, a TGF-$\beta$ receptor inhibitor was added to $-$TGF-$\beta$ wells in 3D culture assays. Endoglin expression levels were checked by qPCR. Histograms show mean of three biological replicates with s.d. VEC and ENG refer to empty vector and Endoglin-transfected cells, respectively. Scale bar: 100 $\mu$m.
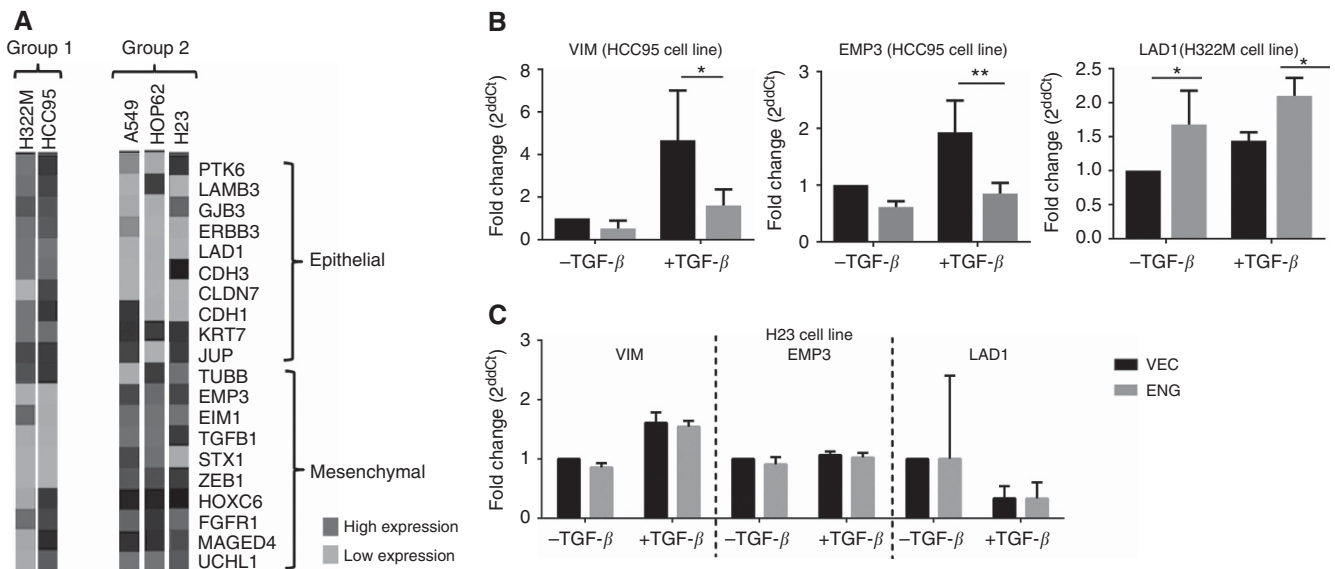


**Figure 5. Loss of Endoglin leads to an altered EMT profile in NSCLC cell lines.** (**A**) A 20-gene EMT expression panel was used to classify EMT status of cell lines in which Endoglin levels affected cell behaviour (Group 1 – H322M, HCC95) and those in which Endoglin levels had no effect (Group 2 – A549, HOP-62, H23). Expression levels of epithelial and mesenchymal markers are represented by a heat map. Changes in expression levels of individual marker genes were analysed for the cell lines H322M, HCC95 and H23. (**B**) The most significantly altered of these are shown for the cell lines H322M and HCC95 and (**C**) for the cell line H23, for which there is no change in expression of EMT markers.
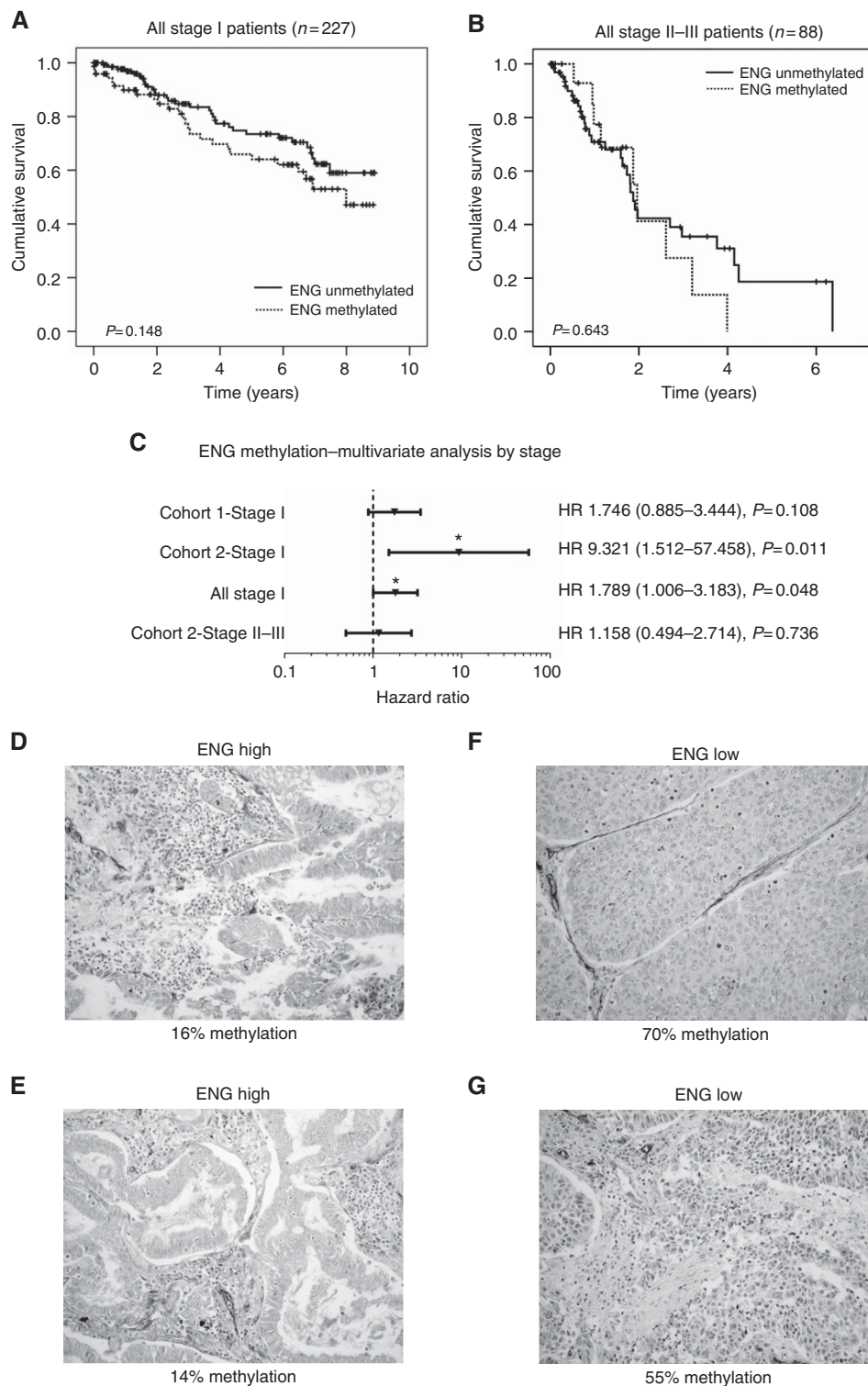
**Figure 6. Methylation of the Endoglin promoter region shows a trend towards shorter overall survival in early-stage NSCLC.** Kaplan–Meier plots show overall survival for patients with unmethylated Endoglin *vs* methylated Endoglin for (**A**) all stage I patients and (**B**) stages II–III patients and suggest a stage I-specific effect of Endoglin methylation. (**C**) This is supported by a multivariate Cox proportional hazards analysis. Forest plot shows hazard ratios and 95% confidence intervals adjusted for gender, age, grade, stage (for stages II–III group only) and histology. Endoglin methylation was significantly associated with survival in stage 1 patients from cohort 2 ($P = 0.011$) and from both cohorts combined ($P = 0.048$). NSCLC sections were stained with Endoglin mAb clone SN6h. Images of representative cores (**D**, **E**) with positive expression and (**F**, **G**) negative expression are shown, along with percentage of methylation of the Endoglin promoter region as measured by pyrosequencing.

## CONFLICT OF INTEREST

The aurthors declare no conflict of interest.

## REFERENCES

Bartram U, Speer CP (2004) The role of transforming growth factor beta in lung development and disease. *Chest* 125: 754–765.

Baylin SB, Herman JG (2000) DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet* 16: 168–174.

Craft CS, Romero D, Vary CP, Bergan RC (2007) Endoglin inhibits prostate cancer motility via activation of the ALK2-Smad1 pathway. *Oncogene* 26: 7240–7250.

Dhillon T, Mauri FA, Bellezza G, Cagini L, Barbareschi M, North BV, Seckl MJ (2010) Overexpression of the mammalian target of rapamycin: a novel biomarker for poor survival in resected early stage non-small cell lung cancer. *J Thorac Oncol* 5: 314–319.

Esteller M (2007) Epigenetic gene silencing in cancer: the DNA hyper-methylome. *Hum Mol Genet* 16(Spec No 1): R50–R59.

Fernandez-Ruiz E, St-Jacques S, Bellon T, Letarte M, Bernabeu C (1993) Assignment of the human endoglin gene (END) to 9q34–>qter. *Cytogenet Cell Genet* 64: 204–207.

Gazdar AF, Miyajima K, Reddy J, Sathyanarayana UG, Shigematsu H, Suzuki M, Takahashi T, Shivapurkar N (2004) Molecular targets for cancer therapy and prevention. *Chest* 125: 97S–101S.

Guerrero-Esteo M, Sanchez-Elsner T, Letamendia A, Bernabeu C (2002) Extracellular and cytoplasmic domains of endoglin interact with the transforming growth factor-beta receptors I and II. *J Biol Chem* 277: 29197–29209.

Hawinkels LJ, Ten Dijke P (2011) Exploring anti-TGF-beta therapies in cancer and fibrosis. *Growth Factors* 29(4): 140–152.

Henry LA, Johnson DA, Sarrio D, Lee S, Quinlan PR, Crook T, Thompson AM, Reis-Filho JS, Isacke CM (2011) Endoglin expression in breast tumor cells suppresses invasion and metastasis and correlates with improved clinical outcome. *Oncogene* 30(9): 1046–1058.

Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP (2009) The human colon cancer methylome shows similar hypo- and hyper-methylation at conserved tissue-specific CpG island shores. *Nat Genet* 41: 178–186.

Jonker L, Arthur HM (2002) Endoglin expression in early development is associated with vasculogenesis and angiogenesis. *Mech Dev* 110: 193–196.

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

Kassouf W, Ismail HR, Aprikian AG, Chevalier S (2004) Whole-mount prostate sections reveal differential endoglin expression in stromal, epithelial, and endothelial cells with the development of prostate cancer. *Prostate Cancer Prostatic Dis* 7: 105–110.

Li DY, Sorensen LK, Brooke BS, Urness LD, Davis EC, Taylor DG, Boak BB, Wendel DP (1999) Defective angiogenesis in mice lacking endoglin. *Science* 284: 1534–1537.

Liu Y, Jovanovic B, Pins M, Lee C, Bergan RC (2002) Over expression of endoglin in human prostate cancer suppresses cell detachment, migration and invasion. *Oncogene* 21: 8272–8281.

Mano Y, Kotani T, Shibata K, Matsumura H, Tsuda H, Sumigama S, Yamamoto E, Iwase A, Senga T, Kikkawa F (2011) The loss of endoglin promotes the invasion of extravillous trophoblasts. *Endocrinology* 152: 4386–4394.

Massague J (1998) TGF-beta signal transduction. *Annu Rev Biochem* 67: 753–791.

Meulmeester E, Ten Dijke P (2011) The dynamic roles of TGF-beta in cancer. *J Pathol* 223: 205–218.

Molina JR, Hayashi Y, Stephens C, Georgescu MM (2010) Invasive glio-blastoma cells acquire stemness and increased Akt activation. *Neoplasia* 12: 453–463.

Palmqvist R, Rutegard JN, Bozoky B, Landberg G, Stenling R (2000) Human colorectal cancers with an intact p16/cyclin D1/pRb pathway have up-regulated p16 expression and decreased proliferation in small invasive tumor clusters. *Am J Pathol* 157: 1947–1953.

Park S, Dimaio TA, Liu W, Wang S, Sorenson CM, Sheibani N (2013) Endoglin regulates the activation and quiescence of endothelium by participating in canonical and non-canonical TGF-beta signaling pathways. *J Cell Sci* 126: 1392–1405.

Perez-Gomez E, Villa-Morales M, Santos J, Fernandez-Piqueras J, Gamallo C, Dotor J, Bernabeu C, Quintanilla M (2007) A role for endoglin as a suppressor of malignancy during mouse skin carcinogenesis. *Cancer Res* 67: 10268–10277.

Postiglione L, Di Domenico G, Caraglia M, Marra M, Giuberti G, Del Vecchio L, Montagnani S, Macri M, Bruno EM, Abbruzzese A, Rossi G (2005) Differential expression and cytoplasm/membrane distribution of endoglin (CD105) in human tumour cell lines: implications in the modulation of cell proliferation. *Int J Oncol* 26: 1193–1201.

Raab U, Lastres P, Arevalo MA, Lopez-Novoa JM, Cabanas C, de la Rosa EJ, Bernabeu C (1999) Endoglin is expressed in the chicken vasculature and is involved in angiogenesis. *FEBS Lett* 459: 249–254.

Rao X, Evans J, Chae H, Pilrose J, Kim S, Yan P, Huang RL, Lai HC, Lin H, Liu Y, Miller D, Rhee JK, Huang YW, Gu F, Gray JW, Huang TM, Nephew KP (2013) CpG island shore methylation regulates caveolin-1 expression in breast cancer. *Oncogene* 32: 4519–4528.

Shi SR, Key ME, Kalra KL (1991) Antigen retrieval in formalin-fixed, paraffin-embedded tissues: an enhancement method for immunohistochemical staining based on microwave oven heating of tissue sections. *J Histochem Cytochem* 39: 741–748.

Siegel PM, Massague J (2003) Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat Rev Cancer* 3: 807–821.

Smith AL, Robin TP, Ford HL (2012) Molecular pathways: targeting the TGF-beta pathway for cancer therapy. *Clin Cancer Res* 18: 4514–4521.

Svensson S, Nilsson K, Ringberg A, Landberg G (2003) Invade or proliferate? Two contrasting events in malignant behavior governed by p16(INK4a) and an intact Rb pathway illustrated by a model system of basal cell carcinoma. *Cancer Res* 63: 1737–1742.

Taby R, Issa JP (2010) Cancer epigenetics. *CA Cancer J Clin* 60: 376–392.

Teodoridis JM, Strathdee G, Brown R (2004) Epigenetic silencing mediated by CpG island methylation: potential as a therapeutic target and as a biomarker. *Drug Resist Updat* 7: 267–278.

Walter K, Holcomb T, Januario T, Du P, Evangelista M, Kartha N, Iniguez L, Soriano R, Huw L, Stern H, Modrusan Z, Seshagiri S, Hampton GM, Amler LC, Bourgon R, Yauch RL, Shames DS (2012) DNA methylation profiling defines clinically relevant biological subsets of non-small cell lung cancer. *Clin Cancer Res* 18: 2360–2373.

Wendt MK, Tian M, Schiemann WP (2011) Deconstructing the mechanisms and consequences of TGF-beta-induced EMT during cancer progression. *Cell Tissue Res*.

Wong VC, Chan PL, Bernabeu C, Law S, Wang LD, Li JL, Tsao SW, Srivastava G, Lung ML (2008) Identification of an invasion and tumor-suppressing gene, Endoglin (ENG), silenced by both epigenetic inactivation and allelic loss in esophageal squamous cell carcinoma. *Int J Cancer* 123: 2816–2823.

Wu TD, Nacu S (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* 26: 873–881.

Zavadil J, Bottinger EP (2005) TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* 24: 5764–5774.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)