Slug/Snail2 is involved in the repression of proliferation genes by TGF-β in bronchial epithelial progenitor cells and is deregulated in abnormal epithelium

Chamseddine Ben Brahim[1,2], Charlotte Courageux[1,2], Ariane Jolly[3], Bérengère Ouine[4], Aurélie Cartier[4], Pierre de la Grange[3], Leanne de Koning[4], Pascale Leroy[1,2*]

[1] INSERM UMR1152, Physiopathology and Epidemiology of Respiratory Diseases, Paris, France

[2] Faculty of Medicine, Paris Diderot University, Bichat Campus, Paris, France.

[3] GenoSplice, Paris, France

[4] Institut Curie, PSL Research University, Department of Translational Research, RPPA platform, Paris, France

*Corresponding author

E-mail: pascale.leroy@inserm.fr (PL)

Abstract

Slug/Snail2 belongs to the Epithelial-Mesenchymal Transition (EMT)-inducing transcription factors that are involved in development and diseases. Slug is also highly expressed in normal adult stem/progenitor cells of several epitheliums, and in such is unique among these transcription factors. By comparing primary bronchial basal cells from normal subjects to those from subjects with Chronic Obstructive Pulmonary Disease (COPD), a respiratory disease in which subjects present many anomalies of their bronchial epithelium and higher levels of Transforming Growth Factor (TGF)-β in their lungs, in an air-liquid interface culture system that allows regenerating a bronchial epithelium similar to the one *in vivo*, we reveal that Slug has higher expression levels in basal/progenitor cells from COPD when in presence of TGF-β but that it does not repress the epithelial marker E-cadherin either in normal or in COPD cells, even when treated with TGF-β. To investigate Slug role in human primary bronchial basal/progenitor cells we performed loss of function experiments to determine Slug downstream genes and we characterized the impact of TGF-β on these genes. We show that Slug downstream genes are different in normal and COPD subjects. In particular, we identified a set of proliferation-related genes whose expression is decreased by TGF-β when cells are induced to differentiate, and that are among the genes repressed downstream of Slug in normal but not in COPD cells. In COPD the levels of expression of these genes are higher than in normal cells in presence of TGF-β □□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□, and they positively correlate with the effect of TGF-β on Slug. Our findings show that Slug is involved in the repression of proliferation genes by TGF-β□in normal basal/progenitor cells, but that in contrast, in subjects

that present many anomalies in their bronchial epithelium this function of Slug is lost and Slug and proliferation genes are simultaneously but independently regulated by TGF-β.

Introduction

Slug (Snail2) belongs with Snail (Snail1) and Smuc (Snail3) to the Snail family of zinc finger transcription factors and has been studied essentially for its function as an inducer of Epithelial-Mesenchymal Transition (EMT) (1). Like other EMT-inducing transcription factors it is expressed in a wide variety of tissues and cell types, during embryogenesis and organ formation and in adult normal and pathological, and it is involved in EMT-related processes and overexpressed in numerous carcinoma. However, in contrast to other EMT-inducing transcription factors, Slug is expressed in adult stem/progenitor cells in normal epitheliums. Recently, it has been shown to control stem/progenitor cell growth during mammary gland morphogenesis (2) and to repress differentiation in epidermal progenitor cells (3). In the lung, Slug is expressed in spheroids of stem/progenitor cells isolated from human adult airway epithelia (4) and, by microarray analysis, it was found enriched in both mouse and human airway basal cells (5, 6).

Basal cells are the adult stem/progenitor cells of the airway epithelium: they can self-renew and/or differentiate to repair the epithelium after injury (7). These cells can be isolated from bronchial tissue and, when grown in an air-liquid interface (ALI) culture system, reconstitute the pseudostratified bronchial epithelium (8). In this system, basal cells keep the memory of the characteristics of their source tissue up to passage 3 to 4; e.g. when isolated from patients with Chronic Obstructive Pulmonary Disease (COPD), they reconstitute an epithelium with anomalies characteristic of COPD (9).

COPD is a respiratory disease characterized by a progressive and irreversible loss of respiratory capacity, with subjects at stage GOLD (Global initiative for chronic Obstructive Lung Disease) 1/A having mild symptoms and at GOLD 4/D having severe symptoms. One of the characteristics of COPD subjects is a remodeling of the airways with, in particular, many anomalies of the epithelium lining the airways such as basal cells hyperplasia, goblet cell metaplasia or squamous metaplasia (10, 11). Cigarette smoke is the main cause of COPD with 80-90% of patient developing COPD being smokers. Several studies have shown that cells exposed to cigarette smoke keep the memory of the exposure and establish a "field of injury" all along the airway epithelium (12). The hypothesis is that the anomalies of the epithelium would result from an imbalance in the basal/progenitor cell fate, with this imbalance progressively worsening with the disease stage and leading to an increase of the extent of the anomalies. These deregulations are thought to be irreversible, since the anomalies are maintained even after subjects quit smoking (13). Noticeably, COPD subjects have an increased risk of lung cancer compared to non-COPD smokers, in particular of squamous carcinoma in the proximal bronchus (14, 15).

Our goal was to better understand Slug function in bronchial epithelium basal/progenitor cells. We hypothesized that Slug is involved in basal/progenitor cell fate i.e. self-renewal and differentiation, and, that its deregulation therefore leads to anomalies in the bronchial epithelium. Slug is a transcription factor and we focused on determining the downstream genes that it, directly or indirectly, regulates. The ALI cell culture system allows regenerating, from primary basal cells, a bronchial epithelium with characteristics close to the in vivo source tissue, providing a good model to characterize the genetic program that takes place during differentiation.

We used COPD subjects to compare with normal subjects as they present many anomalies in their bronchial epithelium that can be reproduced in the ALI cell culture system. Moreover, features of EMT have been reported in COPD airways (9, 16) and Transforming Growth Factor (TGF)-β is found at higher levels in COPD lung tissues (17). We also wanted to characterize the effects of TGF-β, since it has been shown to regulate Slug expression (18), to be, in certain conditions, an EMT-inducing factor (19), and to play a role in stem/progenitor cell fate (20), we also wanted to characterize its effects. We thus determines Slug expression and characterized its downstream genes in bronchial epithelial progenitor cells, comparing cells from normal and COPD subjects at the onset of differentiation in presence or absence of TGF-β.

Materials and methods

Study subjects and cell isolation

Human lung tissues, non-COPD (non-smokers (n=6) and smokers (n=6)) and COPD (smokers (n=6) and Ex-smokers (n=6)) (see S1 Appendix for characteristics), were obtained from patients undergoing lung lobectomy for peripheral lung carcinoma for removal of a primary lung tumor and who gave informed consent. The study was approved by the ethics committee of Paris Nord, IRB 00006477 Paris 7 University, France. COPD patients were diagnosed according to the GOLD guidelines. Lung tissues used in this study were dissected as far away as possible from the tumor. Primary human bronchial epithelial cells were isolated from a piece of large bronchus according to standard protocol.

Cell culture and ALI differentiation system

Primary human bronchial epithelial cells (Passage 1 or 2) were expanded on flasks coated with collagen I (BD Biocoat) in bronchial epithelial growth medium (BEGM), composed of bronchial epithelial basal medium (BEBM) supplemented with the SingleQuots kit (Lonza) and incubated at 37°C in 5% de CO2. Cells were expanded and differentiated in ALI (Air Liquid Interface condition) according to Fulcher et al (8) with the following modifications: Cells ≤90% confluent were plated in BEGM at a density of 1.5 x 10^5 / cm2 on Transwell cell culture inserts with a polyethylene (PET) membrane and 0.4 μm pores (Corning Costar) coated with Collagen IV (Sigma). The next day, medium was replaced with ALI medium. Undifferentiated cells correspond to cells submerged and at day 3 post-plating. Cells at the onset of differentiation correspond to cells put in ALI condition (removal of the medium at the top of the culture) at day 3 post-plating and further cultured for 6 days. TGF-β1 (Peprotech) was added at a concentration of 1 ng/ml 2 days before the switch to ALI condition.

shRNA lentivirus transduction

Protocol for transduction was obtained from D. Bryant (21) and adapted as following: Cells were plated at a density of 10^5/cm2 in BEGM on Transwell cell culture inserts with a PET membrane and 0.4μm pores (Corning Costar) coated with Collagen IV (Sigma), and allow to adhere 3-4 h at 37˚C, 5% CO2. shRNA lentiviral transduction particles from Mission RNAi TRC2 corresponding to SNAI2 specific or control non-targeting sequences (see S2 Appendix for sequence) were added on top of the cells at a MOI of 10. Inserts were swirled and placed O.N. at 37˚C, 5% CO2. Medium was diluted by half with BEGM and insert placed back at 37˚C, 5% CO2 for 1 day. Medium was replaced with ALI medium. At day 4 post-transduction, cells were

changed to ALI conditions. Cells were lysed for RNA or protein analysis at day 6 post-transduction.

Immunofluorescence staining of cell layers

All steps were done on a rocking table. Cells on inserts were rinsed with phosphate buffer saline (PBS)+ 1x, fixed with cold 4% paraformaldehyde in PBS+ for 10 min at 4°C, then permeabilized with 0.2% Triton-X100 in PBS for 10 min. For staining, blocking, incubations and washing were done with PFS (0.7% fish gelatin (Sigma) and 0.025% saponin in PBS 1x). Cells were first blocked with PFS for 30 min at room temperature, then incubated with primary antibodies at 4°C overnight. After extensive wash with PFS, cells were incubated at R.T. for 1 h with the corresponding Alexa Fluor conjugated secondary antibody at a dilution of 1:1000, then washed extensively. Cells were quickly rinsed twice with PBS 1x, adding Hoescht 33342 at 1 µg/ml in the first wash. Filters with the cells were cut out and mounted between a glass slide and a coverslip using ProLong gold (Invitrogen). Slides were left a minimum of 48 h at room temperature before observation under the microscope. Slides were frozen at - 20˚C for long term storage. (see S3 Appendix for antibody references).

RNA extraction

Cells were rinsed with phosphate buffer saline (PBS) before to be homogenized in RNA lysis buffer (NucleoSpin RNA Kit, Macherey-Nagel), supplemented with 1% □-Mercaptoethanol. Lysates were vortexed and proceeded immediately or snapped Frozen and stored at -80˚C. Total RNA was extracted on column using the NucleoSpin RNA Kit (Macherey-

Nagel). DNase I treatment was done directly on the column. RNA concentration was determined using a NanoDrop.

cDNA synthesis

For each sample, 0.3 μg of total RNA was reverse transcribed as previously described (22) with the following modifications: Briefly, total RNA was annealed with 0.1 mg/ml oligo(dT)15 primer (Promega) and cDNA synthesis was performed with M-MLV Reverse Transcriptase (Promega) for 1 h at 42˚C. A control without reverse transcriptase was done for each series of cDNAs.

Quantitative Real-Time Polymerase Chain Reaction (qPCR)

Quantitative PCR was done on the QuantStudio 6 Flex with the QuantStudio Real-Time PCR software using SYBR green PCR master mix (Applied BioSystems). Primers (see S2 Appendix for primer sequences) were designed to be used with an annealing temperature of 60°C and an elongation time of 1 min. For a given gene target, cDNA volume used was chosen to be in a Ct range from 20 to 30, using 0.125 μM each forward and reverse primer. Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) was used to normalize for cDNA amounts between samples.

Protein extraction

Cells were rinsed with cold Tris 20 mM pH 7.5 and lysed directly on inserts with RIPA buffer (Tris pH 7.5, 50 mM; NaCl, 150 mM, NP-40 (leg10) or Triton X-100, 1%: Sodium deoxycholate, 1%; SDS 0.1%). Cells were scrapped and lysates were placed on a wheel for 30 min at 4˚C. DNA was sheared with a needle and lysates spun down at 14K for 10 min at 4˚C. Protein concentration in the lysates was determined using BCA protein assay (Pierce). Lysates were proceed immediately or snapped frozen and stored at -80˚C.

Reverse Phase Protein Array (RPPA)

RPPA was done by the RPPA platform at Curie Institute on samples prepared in RIPA buffer (as described in protein extraction) and stored at -80°C. RPPA was done as following: Protein concentration was determined (Pierce BCA reducing agent compatible kit, ref 23252). Samples were printed onto nitrocellulose covered slides (Supernova, Grace Biolabs) using a dedicated arrayer (2470 arrayer, Aushon Biosystems). Five serial dilutions, ranging from 1000 to 62.5 µg/ml, and four technical replicates per dilution were printed for each sample. Arrays were labeled with 32 specific antibodies, of which 8 were described in this study (see S3 Appendix for antibody references) or without primary antibody (negative control), using an Autostainer Plus (Dako). Briefly, slides were incubated with avidin, biotin and peroxydase blocking reagents (Dako) before saturation with TBS containing 0.1% Tween-20 and 5% BSA (TBST-BSA). Slides were then probed overnight at 4°C with primary antibodies diluted in TBST-BSA. After washes with TBST, arrays were probed with horseradish peroxidase-coupled secondary antibodies (Jackson ImmunoResearch Laboratories, Newmarket, UK) diluted in TBST-BSA for 1 h at RT. To amplify the signal, slides were incubated with Bio-Rad Amplification Reagent for 15 min at RT. The arrays were washed with TBST, probed with IRDye 800CW Streptavidin (LiCOR)

diluted in TBST-BSA for 1 h at RT and washed again in TBST. For staining of total protein, arrays were incubated 30 min Super G blocking buffer (Grace Biolabs), rinsed in water, incubated 5 min in 0,000005% Fast green FCF (Sigma) and rinsed again in water. The processed slides were dried by centrifugation and scanned using a Innoscan 710-AL microarray scanner (Innopsys). Spot intensity was determined with MicroVigene software (VigeneTech Inc). All primary antibodies used in RPPA have been previously tested by Western Blotting to assess their specificity for the protein of interest. Raw data were normalized using Normacurve (23), which normalizes for fluorescent background per spot and a total protein stain.

Gene expression profiling

Microarray analysis was performed on biological triplicate samples. Total RNA were amplified and labeled before hybridization onto Affymetrix human Gene 2.1 ST GeneChip according the manufacturer, by the Genomics platform at Curie Institute, Paris (24). Array datasets were controlled using Expression console (Affymetrix) and further analyses and visualization were made using EASANA (GenoSplice, www.genosplice.com), which is based on the FAST DB annotations (25, 26) Gene Array data were normalized using quantile normalization. Background corrections were made with antigenomic probes and probes were selected as described previously (27). Only probes targeting exons annotated from FAST DB transcripts were selected to focus on well-annotated genes whose mRNA sequences are in public databases (25, 26). Bad-quality selected probes (e.g., probes labeled by Affymetrix as 'cross-hybridizing') and probes whose intensity signal was too low compared to antigenomic background probes with the same GC content were removed from the analysis. Only probes with a DABG P value ≤ 0.05 in at least half of the arrays were considered for statistical analysis (27).

Only genes expressed in at least one compared condition were analyzed. To be considered expressed, the DABG P value had to be ≤ 0.05 for at least half of the gene probes. We performed an unpaired Student's t-test to compare gene intensities in the different biological replicates. Genes were considered significantly regulated when fold-change was ≥ 1.2 and P value ≤ 0.05. Significant KEGG pathways (28), REACTOME pathways (29) and GO terms were retrieved using DAVID (30) from union of results of all, up- and down-regulated genes separately. Data set GEO ID numbers are GSE122957 and GSE123129.

Image capture and analysis

Images of cell layers stained by immunofluorescence were captured using a SP8 Leica confocal microscope equipped with a X 40 objective. The following excitation sources where used: a diode 405-nm, an argon laser 488-nm, a diode 561-nm, a diode 633-nm. Detectors were PMT and HyD. Digital images were analyzed using ImageJ software.

Statistical Analysis

Biological replicates were n ≥ 11 and data generated by at least 2 independent experiments. For RNA or protein levels, mean is ±SD and statistical analysis was carried out by parametric paired or unpaired two-sided t-test as appropriate. For fold-change, mean is ±SEM and statistical analysis was carried out by a one sample two-sided t-test. Correlations were computed as Pearson correlation coefficients and P value determined by two-sided test. Significance was accepted when P value < 0.05.

Results

Slug/Snail2 is the only EMT-inducing transcription factor highly expressed in basal cells and it co-expresses with epithelial and mesenchymal markers

Slug/Snai2 gene has been found by microarray analysis to be highly expressed in basal cells of the human airway epithelium (5, 6). To understand the role of Slug in these stem/progenitor cells, we used primary basal cells isolated from human bronchial epithelium and first characterized these cells for the expression of Slug.

We confirmed by immunocytochemistry that primary basal cells grown at confluence and maintained undifferentiated are all progenitor cells as shown by the expression of the marker p63, and that they all express Slug that co-localizes with p63 in their nuclei **(Fig 1A).** Similar expression profiles were obtained with cells from COPD subjects **(S1 Fig).** Slug is an EMT-inducing transcription factor and we also determined the expression of other transcription factors with this property. While Slug is highly expressed, the expression of Snail1 and Zeb1 is close to background and Twist1 does not exceed 10% of Slug levels, and this in both normal and COPD cells **(Fig 1B).** To determine the epithelial status of these progenitor cells, we studied the expression of epithelial and mesenchymal EMT-related markers. We found that among the genes coding for junction proteins, the epithelial marker E-cadherin (E-cad/CDH1) is expressed while N-cadherin (N-cad/CDH2), an EMT-related mesenchymal marker, is not expressed **(Fig 1C).** We also observed a high level of expression of the genes coding for the cytoskeletal proteins cytokeratin 5 (KRT5) and vimentin (Vim), respectively an epithelial and a mesenchymal marker,

as well as expression of ACTA2, the gene coding for the mesenchymal marker α-smooth muscle actin (a-sma). (**Fig 1D**). In conclusion, both COPD and normal epithelial progenitors co-express epithelial and mesenchymal markers.

Fig 1. Bronchial epithelial progenitors express Slug in their nuclei and co-express epithelial and mesenchymal markers.

Primary bronchial epithelial basal cells were grown on filters and analyzed undifferentiated at confluence, either by fluorescent immunocytochemistry (A) or by RT-qPCR (B-D). (A) Normal cells were fixed and labeled simultaneously with progenitor cell marker p63 (white) and Slug (green) antibodies and with Hoechst as a marker of nuclei (blue). Bars are 20 µm. (B-D) RNA were extracted from normal and COPD cells and analyzed by RT-qPCR. GAPDH was used to normalize cDNA amounts between samples and results were calculated as a ratio on GAPDH. Data are for n $\geq$ 11 and experiments were done at least in duplicate. (B) Expression of the EMT-inducing transcription factors Snai1, Twist1 and Zeb1 shown as a percentage of Slug mRNA, with mean ±SEM (C) Expression of the mesenchymal marker N-cad/CDH2 shown as a percentage of the epithelial marker E-cad/CDH1 mRNA, with mean ±SEM (D) mRNA levels of KRT5 gene, an epithelial cytoskeletal marker and ACTA 2 and Vim genes, mesenchymal cytoskeletal markers. Results shown are log2 (ratio on GAPDH) and are presented as a scatter plot with the mean ±SD. ns: non significant.

Slug/Snail2 expression is increased in COPD cells when compared to normal cells at the onset of differentiation in presence of TGF-β.

We then determined the expression of Slug and the EMT-related markers in cells induced to differentiate using the ALI cell culture system that allows basal/progenitor cells to differentiate into a characteristic bronchial pseudostratified epithelium. The complete differentiation takes 3 to 4 weeks in this system. However, we were interested in the onset of the genetic differentiation program and therefore studied the early timepoints. TGF-β can play a role in stem/progenitor cell fate (20); it also regulates Slug expression, is a potential EMT-inducing factor and is expressed at

higher levels in COPD airways compared to normal (17, 18). We thus compared the expression of Slug and the EMT-related markers at the onset of differentiation in absence or presence of TGF-β to determine if progenitor cells from normal and COPD responded differently to TGF-β. We used low concentration of TGF-β (1ng/ml) to be in conditions similar to the physiological range.

TGF-β significantly increases Slug mRNA expression while its effect on the expression of E-cad/CDH1 is not statistically significant. Slug remains the most highly expressed EMT-inducing transcription factor and among the cell-cell junction markers, E-cad/CDH1 is still the most expressed when compared to the mesenchymal marker N-cad/CDH2. Similar effects are found for both normal and COPD cells **(Fig 2 A-B and D-E)**. Immunocytochemistry shows that while Slug is expressed in the nuclei and E-cad is mainly expressed at the cell-cell junctions in cells at the onset of differentiation, TGF-β leads to delocalization of E-cad to the cytoplasm, with similar profiles of expression being observed in normal and COPD cells **(S2A-D Fig).** Among the cytoskeletal markers, Vim is highly upregulated by TGF-β, while KRT5 and ACTA2 levels do not change significantly. Similar effects are observed in normal and COPD cells **(Fig 2C).**

Fig 2. Effect of TGF-β on Slug and EMT-associated markers in normal and COPD bronchial epithelial progenitors.

Primary bronchial epithelial basal cells, normal and COPD, were grown on filters and at confluence changed to ALI culture to induce differentiation, without TGF-β or in presence of 1ng/ml of TGF-β. Cells were analyzed at day 6 of ALI culture for mRNA expression. RNA were extracted from normal and COPD cells and analyzed by RT-qPCR. GAPDH was used to normalize cDNA amounts between samples and results were calculated as a ratio on GAPDH. Data shown represent the mean for $n \geq 11$, and experiments were done at least in duplicate. (A-C) Results are presented as the fold-change induced by TGF-β on mRNA expression of Slug (A), E-cad/CDH1 (B) and genes coding for cytoskeletal proteins (C) with mean ±SEM, and compare normal and COPD cells. (D) Expression of the EMT-inducing transcription factors Snail1, Twist1 and Zeb1 shown as a percentage of Slug mRNA, with mean ±SEM. (E) Expression of the mesenchymal marker N-cad/CDH2 shown as a percentage of the epithelial marker E-

cad/CDH1 mRNA, with mean ±SEM. Statistical significance is at P value < 1.00E-03 *** as indicated. ns: non significant.

In response to TGF-β, the resulting levels of Slug mRNA are similar in normal and COPD cells, and this is also the case for E-cad/CDH1 and Vim mRNA (**Fig 3A)**. However, when we determined the expression at the protein level by Reverse Phase Protein Array (RPPA), which allows studying all the samples simultaneously, thus, reducing variability due to technical bias (31), we found that TGF-β leads to significantly higher levels of Slug in COPD cells. E-cad protein levels are also higher in COPD in presence of TGF-β, while Vim levels do not differ significantly between normal and COPD cells **(Fig 3B**). Both in normal and COPD cells, and in absence or presence of TGF-β, a strong positive correlation between Slug and E-cad protein levels is observed (**Fig 3C).** These results indicate that Slug is deregulated in COPD basal/progenitor cells at the onset of differentiation in presence of TGF-β.

Fig 3. Slug, E-cad/CDH1 and Vim expression in normal and COPD bronchial epithelial progenitors.

Primary bronchial epithelial basal cells, normal and COPD, were grown on filters and at confluence changed to ALI culture to induce differentiation, without TGF-β or in presence of 1ng/ml of TGF-β. Cells were analyzed at day 6 of ALI culture either for mRNA expression (A) or for protein expression (B-C). (A) Comparison of the mRNA levels in normal and COPD cells in presence of TGF-β of Slug, E-cad/CDH1 and Vim. RNA were extracted from normal and COPD cells and analyzed by RT-qPCR. GAPDH was used to normalize cDNA amounts between samples and results were calculated as a ratio on GAPDH. Data are for n $\geq$ 11 with experiments done at least in duplicate and results are shown as log2 (ratio on GAPDH) and presented as a scatter plot with the mean ±SD. (B-C) Comparison of the protein levels in normal and COPD cells in absence or presence of TGF-β, of Slug, E-cad and Vim. Proteins lysates were prepared from normal and COPD cells and analyzed by RPPA. Data shown are for n $\geq$ 11. (B) Protein levels are shown as log2 (values) and are presented as scatter plots with the mean ±SD. (C) Correlations between Slug and E-cad protein levels in absence or presence of TGF-β for normal and COPD cells. Results are Pearson correlations calculated with log2 (expression levels) and are presented as scatter plots with a regression line. Statistical significance is at P value < 1.00E-02 ** or < 1.00E-03 *** as indicated. ns: non significant.

Microarray analysis of Slug knockdown identifies genes involved in proliferation that are repressed downstream of Slug in normal but not COPD cells

We found that Slug responds differently to TGF-β in normal and COPD bronchial progenitor cells and that it does not repress E-cad in either cell type. To better understand Slug function in these progenitor cells, we wanted to identify the genes downstream of Slug and to determine if these genes were different between normal and COPD cells. It has been reported that the expression level of Slug determines the differentiation status of keratinocytes (3), and we chose a Slug knockdown approach rather than an overexpression that could also change Slug specificity, therefore its downstream genes, in our cell model. We knocked down Slug in normal and COPD cells using shRNA from Mission RNAi that were validated and selected the most efficient shRNA that was within the coding strand in order to limit non-specific downstream genes (32). To determine the genes whose expression was modified by Slug knockdown compared to a control siRNA, we performed a total RNA microarray analysis. We used an Affymetrix human Gene 2.1 ST GeneChip, that allows to test >30,000 coding transcripts and 11,000 long intergenic non-coding transcripts.

Slug Knockdown resulted in a statistically significant decrease of Slug mRNA with knockdown levels determined by RT-qPCR of 43% (fold-change 1.77, P value <1.00E-04) and 54% (fold-change 2.17, P value 4.20E-03), respectively for normal and COPD cells. Slug knockdown level was low, however within the range expected for all Mission RNAi validated Slug sequences. Microarray analysis gave lower fold-change for Slug: 1.48 (P value 1.36E-02)

and 1.39 (P value 8.96E-03), respectively for normal and COPD samples, and for this reason we chose to set the minimum fold-change threshold at 1.2, instead of 1.5 usually used, for selecting genes among those with a statistically significant change of expression (P value ≤5.00E-02). We focused on the genes in normal cells with an expression significantly changed according to these criteria. **A**mong the 805 genes identified downstream of Slug in normal cells, there is a majority (514/63.9%) of upregulated, i.e. genes that would be repressed directly or indirectly by Slug. This is coherent with Slug being a repressor (33, 34). Since we found that Slug expression is deregulated in COPD cells at the onset of differentiation in presence of TGF-β, we were particularly interested in genes responding to differentiation and TGF-β. To identify these genes, we also performed a microarray analysis on RNA from basal cells at the onset of differentiation, in presence or absence of TGF-β, using undifferentiated cells as control. **S1 Table** is the list of the 514 genes upregulated in normal cells with their fold change in normal and COPD cells knocked down for Slug as well as their respective response to differentiation or TGF-β. Setting the minimum fold-change at 1.2, we found that 398 of these genes (77.4%, a large majority) respond to differentiation and/or TGF-β and, among these 193 (48.5%) respond to both. We selected the genes responding to both differentiation and TGF-β and classified them in 4 groups according to their combined type of response. **Fig 4A** shows histograms representing the number of genes for each group and small bars indicating the mean fold-increase induced by Slug knockdown for these genes in normal (blue) and COPD (red) cells. The large majority of the genes are repressed by TGF-β and among these, genes that are upregulated during differentiation have a similar mean fold-increase induced by Slug knockdown in normal and COPD cells. In contrast, the 68 genes that are downregulated during differentiation have a much lower mean fold-increase in COPD than in normal cells. These genes are thus repressed by Slug upon

differentiation ot TGFbeta in normal cells, but this repression is lost in COPD cells. A search for enriched gene pathways using KEGG, REACTOME and Gene Ontology (GO) databases revealed that 45 out of these 68 genes are involved in processes related to proliferation or cell cycle. **Table 1** shows that all these 45 genes, except for 3, are much less or not repressed downstream of Slug in COPD compared to normal cells.

We further studied the expression of 5 of these proliferation-related genes and first confirmed by RT-qPCR the difference in the fold change between normal and COPD cells; we also studied E-cad/CDH1, cytokeratin 5/KRT5 and Vim/VIM genes that were not among the 514 genes significantly increased by Slug knockdown in normal cells. **Fig 4B** shows the correlation between microarray and RT-qPCR values for these 8 genes in both normal and COPD cells. The strong correlation (r^2 values of 0.9318 and 0.8671, respectively for normal and COPD), validates the microarray analysis and also confirms that the 5 proliferation-related genes are upregulated by Slug knockdown in normal but not COPD cells. No significant effect of Slug knockdown is seen on CDH1 and KRT5 expression, while VIM expression decreases significantly and similarly for normal and COPD. This confirms that CDH1/E-cad is not a target of Slug in bronchial progenitor cells from both normal and COPD. KRT5 is also not regulated downstream of Slug, while in contrast, VIM is induced downstream of Slug and similarly in cells from normal and COPD subjects.

Table 1. List of 45 proliferation genes upregulated by Slug knockdown in normal bronchial basal/progenitor cells.

	KD Normal (Mean)		KD COPD (Mean)	
Gene Symbol	**Fold-Change (log2)**	**P-value**	**Fold-Change (log2)**	**P-value**
DTL	0.98	1.06E-02	0.11	3.05E-01
NCAPH	0.93	4.21E-02	-0.06	8.26E-01
CCNA2	0.87	4.62E-03	0.12	4.06E-01
UHRF1	0.87	6.60E-03	-0.04	8.24E-01
ASPM	0.86	1.57E-02	0.12	1.80E-01
FOXM1 // NRIP2	0.83	3.41E-02	0.11	4.89E-01
MCM10	0.77	3.66E-02	0.65	4.10E-02
FAM83D	0.75	1.40E-04	0.21	5.20E-01
CKAP2L	0.75	6.06E-03	-0.11	6.37E-01
FBXO5	0.72	4.98E-02	0.01	8.72E-01
CASC5	0.71	3.04E-02	0.31	1.94E-01
TYMS	0.68	3.04E-02	0.00	9.90E-01
PCNA	0.67	3.02E-03	0.26	3.54E-02
PRC1	0.65	4.02E-03	0.07	6.53E-01
MKI67	0.64	4.47E-02	-0.08	5.01E-01
CEP55	0.61	1.02E-02	-0.16	3.03E-01
CENPE	0.60	3.02E-02	0.18	5.11E-01
MIS18A	0.59	2.23E-02	0.26	4.04E-01
FANCI	0.59	2.26E-02	0.08	6.87E-01
GPSM2	0.54	1.16E-03	0.20	1.45E-01
BORA	0.54	3.05E-02	0.15	5.95E-01
CENPF	0.54	9.86E-03	-0.10	4.00E-01
DLGAP5	0.54	2.38E-02	-0.19	2.71E-01
AURKB	0.51	5.20E-04	0.33	1.15E-01
KIF20A	0.51	2.55E-02	0.11	4.57E-01
KIF20B	0.51	2.56E-02	-0.01	9.41E-01
UBE2C	0.50	2.10E-02	0.32	3.89E-01
CDC20	0.49	3.58E-03	0.07	7.63E-01
CCNB1	0.48	4.94E-02	-0.03	6.89E-01
GINS1	0.48	1.35E-02	-0.08	7.93E-01
FANCD2	0.46	6.88E-03	0.25	2.54E-02
SKA2	0.44	2.55E-02	0.11	5.64E-01
HELLS	0.43	3.01E-02	-0.11	1.53E-01
CKS2	0.43	3.29E-02	-0.15	4.78E-01
NASP	0.41	3.34E-02	0.12	3.31E-02
ALDH3A1	0.40	1.53E-02	1.12	1.46E-03
FEN1	0.38	2.94E-02	0.37	1.06E-01
RFC4	0.37	4.18E-02	0.11	6.10E-01
HIST2H4A // HIST2H4B	0.34	1.02E-03	0.49	2.03E-01
HJURP	0.32	1.72E-02	0.10	7.84E-01
CDKN3	0.32	3.37E-02	-0.04	5.49E-01
PIR // VEGFD	0.30	4.13E-02	0.10	5.06E-01
PRIM1	0.30	3.66E-03	-0.10	4.67E-01
RNASEH2A	0.29	1.09E-02	0.04	5.42E-01
CDCA7L	0.28	1.14E-02	-0.06	6.18E-01

Genes statistically significant upregulated by Slug knockdown in normal cells and downregulated by both differentiation and TGF-β.
Genes were selected for a log2 (fold-Change) ≥ 0.26. Statistical significance is at P value < 5.00E-02
Comparison with fold-change for Slug knockdown in COPD cells.

Fig 4. Microarray analysis of Slug knockdown identifies proliferation genes repressed in normal bronchial epithelial progenitors.

Knockdown with shRNA specific of Slug were performed on primary bronchial epithelial basal cells, normal and COPD, Knockdown cells were grown on filters and, at confluence, changed to ALI culture to induce differentiation. Cells were analyzed at day 2 of ALI culture for mRNA expression. RNA were extracted from normal and COPD cells and analyzed by microarray analysis on an Affymetrix chip. Data are for n ≥ 3 each, normal and COPD. (A) Genes significantly upregulated by Slug knockdown in normal cells and responding to both differentiation and TGF-β were classified in 4 groups according to their response to differentiation and TGF-β. Histograms represent the number of genes in each group with genes upregulated by TGF-β in red and genes downregulated by TGF-β in green, and genes upregulated during differentiation as checkboard and genes downregulated during differentiation as oblique lines. Horizontal bars represent the mean fold-increase for each group of genes, in blue for normal cells and in red for COPD cells. (B) Validation of microarray by RT-qPCR and comparison of Slug knockdown effect in normal (blue) and COPD (red) cells for CDH1, KRT5 and VIM genes as well as proliferation-related genes. Results are Pearson correlations calculated with log2 (fold-change) and are presented as scatter plots with regression line and R-square values.

TGF-β represses the expression of proliferation-related genes in progenitor cells at the onset of differentiation and Slug is involved in this repression in normal but not in COPD cells.

We found a set of proliferation-related genes that have an increased expression following Slug knockdown in normal but not in COPD cells; this suggests that these genes are directly or indirectly repressed by Slug in normal cells but that this regulation does not take place in COPD. This is supported by the strong and statistically significant inverse correlation between Slug and the mRNA levels of these genes that is observed in undifferentiated normal but not COPD cells (**Table 2**). Moreover, in normal cells, in addition to Ki67 mRNA, a strong inverse correlation also

exists between Slug and Ki67 protein, a widely used marker strictly related to proliferation (35) **(Fig 5A).** These correlations are found with Slug mRNA and not with the protein.

Table 2. Correlation between proliferation genes and Slug in normal and COPD undifferentiated basal cells

	Normal		COPD	
	Slug mRNA	**Slug Protein**	**Slug mRNA**	**Slug Protein**
PRC1 mRNA	**- 0.7331** (*6.70E-03*)	0.3264 (*3.27E-01*)	- 0.2567 (*4.46E-01*)	0.3178 (*3.41E-01*)
FOXM1 mRNA	**- 0.7680** (*3.50E-03*)	0.3295 (*3.23E-01*)	- 0.3640 (*2.71E-01*)	0.2530 (*4.53E-01*)
CCNA2 mRNA	**- 0.7615** (*4.00E-03*)	0.3514 (*2.89E-01*)	- 0.1772 (*6.02E-01*)	0.2739 (*4.15E-01*)
PCNA mRNA	**- 0.7365** (*6.30E-03*)	0.3437 (*3.01E-01*)	0.0286 (*9.33E-01*)	0.0519 (*8.80E-01*)
KI67 mRNA	**- 0.7322** (*6.80E-03*)	0.3329 (*3.17E-01*)	- 0.1092 (*7.49E-01*)	0.0833 (*8.08E-01*)

Pearson correlation coefficients and associated P value between expression levels.
Computation was performed on log2 of values.
Statistical significance was determined with a two-sided t-test. Correlations are significant at P value < 5.00E-02.

At the expression level in undifferentiated cells, only Ki67 mRNA is statistically significant higher in COPD when compared to normal cells, and other proliferation-related genes show only a tendency in the same direction (**Fig 5B)**. However, in cells at the onset of differentiation in presence of TGF-β, significant lower levels are found in normal cells including for Ki67 protein, with the exception of PCNA that shows only a weak not statistically significant tendency (**Fig 5C)**. TGF-β decreases the expression of these proliferation-related genes and its mean effect is in accordance with the mean difference of their expression levels between normal and COPD cells (**Fig 5D)**.

To determine the implication of Slug in the repression of the proliferation-related gene expression by TGF-β in normal cells, we analyzed the correlation between Slug and these genes. **Table 3 and Fig 5E** show that in normal cells, there are significant inverse correlations between the combined effects of differentiation and TGF-β on Slug and the proliferation-related gene mRNA levels, except for PCNA. A good tendency for an inverse correlation is also found for Ki67 protein. No such inverse correlations are found in COPD cells, showing that Slug is

involved in the repression of the proliferation-related gene expression by TGF-β in normal cells but that this function of Slug is lost in COPD cells.

Table 3. Correlation between the effect of differentiation and TGF-β on proliferation genes and on Slug in normal and COPD basal/progenitor cells

	Normal		COPD	
	Slug mRNA	**Slug Protein**	**Slug mRNA**	**Slug Protein**
PRC1 mRNA	**- 0.6983** (*1.15E-02*)	- 0.4189 (*2.00E-01*)	0.4013 (*2.21E-01*)	**0.6755** (*2.25E-02*)
FOXM1 mRNA	**- 0.6003** (*3.91E-02*)	- 0.3884 (2.38E-01)	0.5383 (*8.76E-02*)	**0.8093** (*2.50E-03*)
CCNA2 mRNA	**- 0.7056** (*1.04E-02*)	- 0.3021 (3.67E-01)	0.3527 (*2.87E-01*)	0.5576 (*7.47E-02*)
PCNA mRNA	- 0.2238 (*4.84E-01*)	- 0.1194 (7.27E-01)	0.2448 (*4.68E-01*)	0.1919 (*5.72E-01*)
KI67 mRNA	**- 0.5840** (*4.62E-02*)	- 0.4359 (1.80E-01)	0.2768 (*4.10E-01*)	0.5275 (*9.54E-02*)

Pearson correlation coefficients and associated P value between the effect of differentiation and TGF-β. Statistical significance was determined with a two-sided t-test. Correlations are significant at P value < 5.00E-02

Fig 5. Proliferation genes are more repressed by TGF-β in normal bronchial epithelial progenitors than in COPD.

Primary bronchial epithelial basal cells, normal and COPD, were grown on filters and analyzed undifferentiated at confluence or grown on filters and at confluence changed to ALI culture to induce differentiation, without TGF-β or in presence of 1 ng/ml of TGF-β. Cells were analyzed undifferentiated or at day 6 of ALI culture either for mRNA or for protein expression. RNA or proteins lysates were prepared from normal and COPD cells and analyzed respectively by RT-qPCR and RPPA. For RT-qPCR, GAPDH was used to normalize cDNA amounts between samples and results were calculated as a ratio on GAPDH. Data shown are for n ≥ 11. Correlations between Slug mRNA and Ki67 mRNA and protein levels in undifferentiated normal cells. Results are Pearson correlations calculated with log2 (expression levels) and are presented as scatter plots with a regression line. (B, C) Comparison of the mRNA levels of proliferation-related genes and protein levels of Ki67 between normal and COPD cells, undifferentiated (B), at the onset of differentiation without TGF-β or in presence of 1 ng/ml of TGF-β (C). Results shown are log2 (ratio on GAPDH) for mRNA or log2 (values) for protein and are presented as a scatter plot with the mean ±SD. (D) Comparison of TGF-β effect on the expression of proliferation-related genes and proliferation marker Ki67 protein between normal and COPD cells. Results are presented as the fold-change induced by TGF-β on mRNA expression with mean ±SEM. (E, F) Correlations between Ki67, mRNA or protein, and Slug mRNA in normal cells

(E) or Slug protein in COPD cells (F) for the effect of differentiation and TGF-β. Results are Pearson correlations calculated with the level of effect and are presented as scatter plots with a regression line. Statistical significance is at P value < 5.00E-02 *, < 1.00E-02 ** or < 1.00E-03 *** as indicated. ns: non significant.

The higher levels of Slug protein and of proliferation-related genes observed in COPD progenitor cells are differently induced by TGF-β

Slug protein and proliferation-related gene levels are higher in COPD cells in presence of TGF-β (**Fig 3B and Fig 5C**) and, in contrast to normal cells, positive correlations, significant or with a good tendency, are observed between the combined effects of differentiation and TGF-β on Slug protein levels and that on the proliferation-related gene levels, with the exception of PCNA. A strong positive correlation is also found between Slug and Ki67 protein (**Table 3 and Fig 5F**). To understand the link between the higher expression of Slug protein and of the proliferation-related gene in COPD progenitor cells in presence of TGF-β, we determined the expression of signaling pathways downstream of TGF-β, using RPPA. In addition to the canonical TGF-β signaling pathway Smad3-dependent, we also considered β-catenin (β-cat) signaling pathway known to crosstalk with Smad signaling and to be involved in proliferation (36). While TGF-β increases the level of phosphorylation of Smad3, i.e. activates Smad3-dependent pathway, both in normal and COPD cells and at similar levels, it turns on β-cat signaling (i.e. decreases phosphorylation of β-cat) only in COPD. In addition, TGF-β increases the levels of total Smad3 and β-cat also only in COPD (**Fig 6 A and B**).

Slug protein levels correlate positively with both total Smad3 and β-cat levels, but do not correlate with the phosphorylation levels of Smad3 or β-cat (**Fig 6C and D**). In contrast, a positive correlation, significant or as a strong tendency, is found between the activation of β-cat

(i.e. decrease of β-cat phosphorylation) and the levels of Ki67 protein and of the proliferation-related genes, with the exception of PCNA (**Table 4, Fig 6E).** These results show that, in COPD cells, the higher levels of Slug protein and of the proliferation-related gene expression found in presence of TGF-β, are induced by TGF-β through different downstream pathways.

Table 4. Correlation between proliferation genes and Smad3 or β-cat in COPD basal/progenitor cells in presence of TGF-β.

	Smad3		β-catenin	
	Phospho/Total	**Total**	**Phospho/Total**	**Total**
PRC1 mRNA	- 0.1911 *(5.74E-01)*	0.3377 (*3.10E-01*)	**- 0.6551 (***2.87E-02* **)**	**0.7683 (***5.70E-03* **)**
FOXM1 mRNA	- 0.2244 (*5.07E-01*)	0.291 (*3.85E-01*)	**- 0.6511 (***3.00E-02* **)**	**0.6877 (***1.94E-02* **)**
CCNA2 mRNA	- 0.2726 (4.17E-01)	0.3709 (*2.62E-01*)	**- 0.5391 (***8.70E-02* **)**	**0.7667 (***5.90E-03* **)**
PCNA mRNA	- 0.1091 (7.49E-01)	0.314 (*3.47E-01*)	- 0.3975 (*2.26E-01*)	**0.6769 (***2.22E-02* **)**
KI67 mRNA	- 0.2256 (5.05E-01)	0.2769 (*4.10E-01*)	**- 0.7155 (***1.33E-02* **)**	**0.7872 (***4.00E-03* **)**

Pearson correlation coefficients and associated P value between protein levels (phosphorylated/Total or Total). Computation was performed on log2 of values.
Statistical significance was determined with a two-sided t-test. Correlations are significant at P value < 5.00E-02

Fig 6. In COPD, TGF-β induced higher levels of Slug protein and proliferation genes through different pathways.

Primary bronchial epithelial basal cells, normal and COPD, were grown on filters and at confluence changed to ALI culture to induce differentiation, without TGF-β or in presence of 1 ng/ml of TGF-β. Cells were analyzed at day 6 of ALI culture for protein expression. Proteins lysates were prepared from normal and COPD cells and analyzed by RPPA. Data shown are for n ≥ 11. (A-B) Comparison of TGF-β effect on the ratio Phospho/Total or total protein expression levels of Smad3 (A) and β-cat □B□ between normal and COPD cells. Results are presented as the fold-change induced by TGF-β with mean ±SEM. □C-E□□ Correlations between Smad3 or β-cat and Slug or Ki67 protein levels in COPD cells in presence of TGF-β. Smad3 (Total protein level or ratio Phospho/Total) and Slug protein levels (C), β-cat (Total protein level or ratio Phospho/Total) and Slug protein levels (D) or Ki67 protein levels (E). Results are Pearson correlations calculated with log2 (expression levels) or log2 (Ratio Phospho/Total) and are presented as scatter plots with a regression line. Statistical significance is at P value < 5.00E-02 *, < 1.00E-02 ** or < 1.00E-03 *** as indicated. ns: non significant.

Discussion

In this work, we show that in bronchial basal/progenitor cells, Slug is co-expressed with epithelial and mesenchymal markers and that, despite being highly expressed in the cell nuclei, it does not repress E-cad, even in presence of TGF-β, and this in normal cells as well as COPD cells that have been described to reconstitute a bronchial epithelium with anomalies and features of EMT (9). Using a Slug knockdown approach and comparing between normal and COPD cells the genes downstream of Slug that respond to both differentiation and TGF-β, we identified a large set of genes involved in cell cycle and proliferation that are directly or indirectly repressed targets of Slug in normal but not in COPD basal/progenitor cells. By comparing the expression of several of these genes between normal and COPD cells, we found that their mRNA levels as well as Ki67 protein levels are lower in normal cells at the onset of differentiation when in presence of TGF-β. We also found a negative correlation between the effect of TGF-β on their expression and that on Slug at the onset of differentiation in normal cells, while in contrast, a mild positive correlation was observed in COPD. This shows that normal and COPD basal/progenitor cells respond differently to TGF-β, with Slug being involved in the repression of proliferation genes by TGF-β in normal cells, while this function of Slug is lost in COPD.

Bronchial basal/progenitor cells express both epithelial and mesenchymal markers simultaneously, both in normal and COPD cells and similarly in presence of TGF-β. In mammary stem cells a co-expression of epithelial and mesenchymal markers has also been reported leading to the concept of hybrid phenotype (37, 38). In our studies, apart from Slug, only the mesenchymal marker VIM is highly induced by TGF-β in normal and COPD cells, resulting in similar levels in both cell type. This induction of VIM expression is coherent with our finding that it belongs to the genes that are induced downstream of Slug and that it is induced at similar

554 levels in normal and COPD cells. These results are in agreement with a report showing that Slug
555 is important for the up-regulation of Vim by TGF-β in epithelial cells (18). On the contrary, E-
556 cad/ CDH1 gene is not among the genes whose expression is modified by Slug knockdown,
557 confirming that it is not a target of Slug in bronchial basal/progenitor cells. Co-expression of Slug
558 and E-cad have also been reported in mammary stem cells (39). In our study, we observe a
559 positive correlation between Slug and E-cad proteins in all conditions tested, indicating that Slug
560 does not repress E-cad in these cells even in COPD and in presence of TGF-β. Several studies
561 have reported features of EMT in COPD bronchial epithelium. However, this is based on studies
562 only revealing cells that double stain for epithelial and mesenchymal markers or on studies
563 showing only a slight lower expression of E-cad in COPD. Also theses studies concern cells from
564 *ex-vivo* epithelium or fully differentiated epithelium *in vitro* ALI culture, and suggest that COPD
565 cells may be rather imprinted for EMT, and in such, have a higher potential to enter EMT in
566 permissive conditions (9, 17, 40).

567 In search of Slug function in basal/progenitor cells, we identified genes involved in
568 proliferation that are repressed downstream of Slug in normal but not COPD cells. Among the
569 genes studied, PCNA is apart: It is the gene with the least difference between normal and COPD
570 of Slug KD effect, and this is coherent with the fact that there is no significant difference of
571 expression between normal and COPD and no correlation with Slug for the effect of TGF-β. In
572 addition to its role in DNA replication, PCNA has also a role in DNA repair, making it a poor
573 proliferation marker. In normal cells, Slug mRNA, but not protein, correlates with these
574 proliferation-related genes. We cannot exclude that, by an unknown mechanism, Slug mRNA is
575 involved in the repression of these genes; alternatively, Slug being regulated at the post-
576 translational level (41) they could be repressed by a form of Slug protein not recognized by the

antibody used, and this form would be absent in COPD. In contrast to normal cells, a positive correlation is found between the effect of TGF-β on these proliferation genes and on Slug. Our results confirm the involvement of Slug in proliferation previously reported and can provide an explanation to the fact that that some studies report Slug to be a repressor of proliferation (42-44), while others report a positive link with proliferation (2, 45, 46). Normal and COPD basal cells may represent two cell states with a different regulation of proliferation genes by Slug and TGF-β and this could be similar to what is observed in the different studies that use cell lines from different origin.

Our work reveals that genes downstream of Slug are different between normal and COPD cells and that they respond differently to TGF-β. In normal cells, Slug is involved in the response of proliferation genes to TGF-β while in COPD TGF-β regulates Slug and the proliferation genes through different mechanisms, with proliferation gene expression being linked to β-cat activation by TGF-β. In fact, Slug, β-cat and Smad3 protein have all been reported to be stabilized following GSK-3β inactivation (41, 47, 48) and we can speculate that in COPD cells, the increased levels of these proteins results from such a mechanism induced by TGF-β.

TGF-β function is complex and has a wide spectrum of effects depending on cell state: it is a cell cycle inhibitor in normal cells and a tumor promoter in malignant cells (49, 50). The difference between normal and COPD cells could reflect these antagonistic effects of TGF-β, COPD being in a premalignant state. Moreover, the higher expression of Slug protein seen in COPD basal cells in presence of TGF-β could increase during repeated injury. Slug levels of expression define its function (3) and overexpression of Slug induces EMT in epithelial cells (33). We can also speculate that such deregulations of the progenitor cells could ultimately lead to

599 a shift in Slug function, becoming an EMT-inducing factor. That could explain the higher EMT
600 features found in COPD epithelium and the increase risks for COPD to develop lung carcinomas.
601

602 # Acknowledgements

603 We thank the Pulmonary Department, the Pathology Department and the Thoracic Surgery
604 Department at Bichat-Claude Bernard University Hospital (Paris, France) and INSERM UMR
605 1152 for providing lung tissues and isolating the cells. We thank Audrey Rapinat and David
606 Gentien at the Genomics platform, Curie Institute (Paris, France) for Affymetrix GeneChip
607 hybridization. We thank Dusko Ilic and Pierre Savagner for critical reading of the manuscript and
608 helpful discussions. P.L. is supported by the French National Center for Scientific Research
609 (CNRS). This work was supported by a donation from Association Science et Technologie
610 (Groupe Servier) to P.L. and by funding from French National Institute for Medical Research
611 (INSERM).
612

613 # References

614 1. Nieto MA, Huang RY, Jackson RA, Thiery JP. Emt: 2016. Cell 2016;166(1):21-
615 45.
616 2. Nassour M, Idoux-Gillet Y, Selmi A, Come C, Faraldo ML, Deugnier MA, et al.
617 Slug controls stem/progenitor cell growth dynamics during mammary gland
618 morphogenesis. PLoS One 2012;7(12):e53498.

619 3. Mistry DS, Chen Y, Wang Y, Zhang K, Sen GL. SNAI2 controls the
620 undifferentiated state of human epidermal progenitor cells. Stem Cells
621 2014;32(12):3209-18.
622 4. Tesei A, Zoli W, Arienti C, Storci G, Granato AM, Pasquinelli G, et al. Isolation
623 of stem/progenitor cells from normal lung tissue of adult humans. Cell Prolif
624 2009;42(3):298-308.
625 5. Hackett NR, Shaykhiev R, Walters MS, Wang R, Zwick RK, Ferris B, et al. The
626 human airway epithelial basal cell transcriptome. PLoS One
627 2011;6(5):e18378.
628 6. Rock JR, Onaitis MW, Rawlins EL, Lu Y, Clark CP, Xue Y, et al. Basal cells as
629 stem cells of the mouse trachea and human airway epithelium. Proc Natl
630 Acad Sci U S A 2009;106(31):12771-5.
631 7. Rock JR, Randell SH, Hogan BL. Airway basal stem cells: a perspective on
632 their roles in epithelial homeostasis and remodeling. Dis Model Mech
633 2010;3(9-10):545-56.
634 8. Fulcher ML, Gabriel S, Burns KA, Yankaskas JR, Randell SH. Well-
635 differentiated human airway epithelial cell cultures. Methods Mol Med
636 2005;107:183-206.
637 9. Gohy ST, Hupin C, Fregimilicka C, Detry BR, Bouzin C, Gaide Chevronay H, et
638 al. Imprinting of the COPD airway epithelium for dedifferentiation and
639 mesenchymal transition. Eur Respir J 2015;45(5):1258-72.
640 10. Rigden HM, Alias A, Havelock T, O'Donnell R, Djukanovic R, Davies DE, et al.
641 Squamous Metaplasia Is Increased in the Bronchial Epithelium of Smokers

with Chronic Obstructive Pulmonary Disease. PLoS One 2016;11(5):e0156009.

11. Sohal SS, Walters EH. Role of epithelial mesenchymal transition (EMT) in chronic obstructive pulmonary disease (COPD). Respir Res 2013;14:120.

12. Steiling K, van den Berge M, Hijazi K, Florido R, Campbell J, Liu G, et al. A dynamic bronchial airway gene expression signature of chronic obstructive pulmonary disease and lung function impairment. Am J Respir Crit Care Med 2013;187(9):933-42.

13. Randell SH. Airway epithelial stem cells and the pathophysiology of chronic obstructive pulmonary disease. Proc Am Thorac Soc 2006;3(8):718-25.

14. Powell HA, Iyen-Omofoman B, Baldwin DR, Hubbard RB, Tata LJ. Chronic obstructive pulmonary disease and risk of lung cancer: the importance of smoking and timing of diagnosis. J Thorac Oncol 2013;8(1):6-11.

15. Young RP, Hopkins RJ, Christmas T, Black PN, Metcalf P, Gamble GD. COPD prevalence is increased in lung cancer, independent of age, sex and smoking history. Eur Respir J 2009;34(2):380-6.

16. Mahmood MQ, Sohal SS, Shukla SD, Ward C, Hardikar A, Noor WD, et al. Epithelial mesenchymal transition in smokers: large versus small airways and relation to airflow obstruction. Int J Chron Obstruct Pulmon Dis 2015;10:1515-24.

17. Mahmood MQ, Reid D, Ward C, Muller HK, Knight DA, Sohal SS, et al. Transforming growth factor (TGF) beta1 and Smad signalling pathways: A

likely key to EMT-associated COPD pathogenesis. Respirology 2017;22(1):133-140.

18. Slabakova E, Pernicova Z, Slavickova E, Starsichova A, Kozubik A, Soucek K. TGF-beta1-induced EMT of non-transformed prostate hyperplasia cells is characterized by early induction of SNAI2/Slug. Prostate 2011;71(12):1332-43.

19. Xu J, Lamouille S, Derynck R. TGF-beta-induced epithelial to mesenchymal transition. Cell Res 2009;19(2):156-72.

20. Yumoto K, Thomas PS, Lane J, Matsuzaki K, Inagaki M, Ninomiya-Tsuji J, et al. TGF-beta-activated kinase 1 (Tak1) mediates agonist-induced Smad activation and linker region phosphorylation in embryonic craniofacial neural crest-derived cells. J Biol Chem 2013;288(19):13467-80.

21. Bryant DM, Datta A, Rodriguez-Fraticelli AE, Peranen J, Martin-Belmonte F, Mostov KE. A molecular network for de novo generation of the apical surface and lumen. Nat Cell Biol 2010;12(11):1035-45.

22. Leroy P, Mostov KE. Slug is required for cell survival during partial epithelial-mesenchymal transition of HGF-induced tubulogenesis. Mol Biol Cell 2007;18(5):1943-52.

23. Troncale S, Barbet A, Coulibaly L, Henry E, He B, Barillot E, et al. NormaCurve: a SuperCurve-based method that simultaneously quantifies and normalizes reverse phase protein array data. PLoS One 2012;7(6):e38686.

685 24. Maubant S, Tahtouh T, Brisson A, Maire V, Nemati F, Tesson B, et al. LRP5
686 regulates the expression of STK40, a new potential target in triple-negative
687 breast cancers. Oncotarget 2018;9(32):22586-22604.
688 25. de la Grange P, Dutertre M, Correa M, Auboeuf D. A new advance in
689 alternative splicing databases: from catalogue to detailed analysis of
690 regulation of expression and function of human alternative splicing variants.
691 BMC Bioinformatics 2007;8:180.
692 26. de la Grange P, Dutertre M, Martin N, Auboeuf D. FAST DB: a website
693 resource for the study of the expression regulation of human gene products.
694 Nucleic Acids Res 2005;33(13):4276-84.
695 27. de la Grange P, Gratadou L, Delord M, Dutertre M, Auboeuf D. Splicing factor
696 and exon profiling across human tissues. Nucleic Acids Res 2010;38(9):2825-
697 38.
698 28. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and
699 interpretation of large-scale molecular data sets. Nucleic Acids Res
700 2012;40(Database issue):D109-14.
701 29. Haw R, Stein L. Using the reactome database. Curr Protoc Bioinformatics
702 2012;Chapter 8:Unit8 7.
703 30. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis
704 of large gene lists using DAVID bioinformatics resources. Nat Protoc
705 2009;4(1):44-57.
706 31. Akbani R, Becker KF, Carragher N, Goldstein T, de Koning L, Korf U, et al.
707 Realizing the promise of reverse phase protein arrays for clinical,

translational, and basic research: a workshop report: the RPPA (Reverse Phase Protein Array) society. Mol Cell Proteomics 2014;13(7):1625-43.

32. Coussens MJ, Corman C, Fischer AL, Sago J, Swarthout J. MISSION LentiPlex pooled shRNA library screening in mammalian cells. J Vis Exp 2011(58).

33. Bolos V, Peinado H, Perez-Moreno MA, Fraga MF, Esteller M, Cano A. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. J Cell Sci 2003;116(Pt 3):499-511.

34. Savagner P. Leaving the neighborhood: molecular mechanisms involved during epithelial-mesenchymal transition. Bioessays 2001;23(10):912-23.

35. Jurikova M, Danihel L, Polak S, Varga I. Ki67, PCNA, and MCM proteins: Markers of proliferation in the diagnosis of breast cancer. Acta Histochem 2016;118(5):544-52.

36. Logan CY, Nusse R. The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol 2004;20:781-810.

37. Jolly MK, Tripathi SC, Jia D, Mooney SM, Celiktas M, Hanash SM, et al. Stability of the hybrid epithelial/mesenchymal phenotype. Oncotarget 2016;7(19):27067-84.

38. Lu M, Jolly MK, Levine H, Onuchic JN, Ben-Jacob E. MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. Proc Natl Acad Sci U S A 2013;110(45):18144-9.

39. Ye GD, Sun GB, Jiao P, Chen C, Liu QF, Huang XL, et al. OVOL2, an Inhibitor of WNT Signaling, Reduces Invasive Activities of Human and Mouse Cancer

731 Cells and Is Down-regulated in Human Colorectal Tumors. Gastroenterology
732 2016;150(3):659-671 e16.
733 40. Sohal SS, Reid D, Soltani A, Ward C, Weston S, Muller HK, et al. Evaluation of
734 epithelial mesenchymal transition in patients with chronic obstructive
735 pulmonary disease. Respir Res 2011;12:130.
736 41. Kim JY, Kim YM, Yang CH, Cho SK, Lee JW, Cho M. Functional Regulation of
737 Slug/Snail2 is dependant on GSK-3b-mediated phosphorylation. FEBS J
738 2012;279(16):2929-39.
739 42. Sun Y, Shao L, Bai H, Wang ZZ, Wu WS. Slug deficiency enhances self-renewal
740 of hematopoietic stem cells during hematopoietic regeneration. Blood
741 2010;115(9):1709-17.
742 43. Turner FE, Broad S, Khanim FL, Jeanes A, Talma S, Hughes S, et al. Slug
743 regulates integrin expression and cell proliferation in human epidermal
744 keratinocytes. J Biol Chem 2006;281(30):21321-31.
745 44. Wang WL, Huang HC, Kao SH, Hsu YC, Wang YT, Li KC, et al. Slug is temporally
746 regulated by cyclin E in cell cycle and controls genome stability. Oncogene
747 2015;34(9):1116-25.
748 45. Bhat-Nakshatri P, Appaiah H, Ballas C, Pick-Franke P, Goulet R, Jr., Badve S, et
749 al. SLUG/SNAI2 and tumor necrosis factor generate breast cells with
750 CD44+/CD24- phenotype. BMC Cancer 2010;10:411.
751 46. Phillips S, Prat A, Sedic M, Proia T, Wronski A, Mazumdar S, et al. Cell-state
752 transitions regulated by SLUG are critical for tissue regeneration and tumor
753 initiation. Stem Cell Reports 2014;2(5):633-47.

47. Guo X, Ramirez A, Waddell DS, Li Z, Liu X, Wang XF. Axin and GSK3- control Smad3 protein stability and modulate TGF - signaling. Genes Dev 2008;22(1):106-20.

48. Liu C, Li Y, Semenov M, Han C, Baeq GH, Tan Y, et al. Control of beta-catenin phosphorylation/degradation by a dual mechanism. Cell 2002;108(6):837-47.

49. Derynck R, Akhurst RJ, Balmain A. TGF-beta signaling in tumor suppression and cancer progression. Nat Genet 2001;29(2):117-29.

50. Fynan TM, Reiss M. Resistance to inhibition of cell growth by transforming growth factor-beta and its role in oncogenesis. Crit Rev Oncog 1993;4(5):493-540.

Supporting information

S1 Fig. Bronchial epithelial progenitors from COPD express Slug in their nuclei.
Primary bronchial epithelial basal cells from COPD were grown on filters and analyzed undifferentiated at confluence by fluorescent immunocytochemistry. Cells were fixed and labeled simultaneously with progenitor cell marker p63 (white) and Slug (green) antibodies and with Hoechst as a marker of nuclei (blue). Bars are 20 μm.

S2A-D Fig. Bronchial epithelial progenitors express Slug in their nuclei and co-express E-cadherin with or without TGF-β. Normal (A-B) or COPD (C-D) primary bronchial epithelial basal cells were grown on filters and at confluence changed to ALI culture to induce differentiation, without TGF-β (A, C) or in presence of 1 ng/ml of TGF-β (B, D). Cells were analyzed at day 6 of ALI culture by fluorescent immunocytochemistry. Cells were fixed and labeled simultaneously with Slug (green) and E-cad (red) antibodies and with Hoechst as a marker of nuclei (blue). Bars are 20 μm.

779 **S1 Table. List of the 514 genes upregulated by Slug knockdown in normal bronchial basal/progenitor cells.**

780 Genes statistically significant upregulated by Slug knockdown in normal bronchial basal/progenitor cells:

781 Table shows the response of the genes to differentiation and TGFβ and fold-change for Slug knockdown in COPD

782 cells. Genes were selected for a log2 (fold-Change) ≥0.26. Statistical significance is at p-val < 5.00E-02

783

784 **S1 Appendix: Characteristics of the study subjects.**

785 Data are presented as Mean ±SD. Pack-year=1 year smoking 20 cigarettes per day. COPD, chronic obstructive

786 pulmonary disease; FEV1, forced expiratory volume in 1 s;

787

788 **S2 Appendix: Sequences of primers for PCR and shRNA**

789

790 **S3 Appendix: References for antibodies used in immunochemistry and RPPA**

791 CST: Cell Signaling Technology; DBS: Diagnostic BioSystems: BD: Becton Dickinson Biosciences

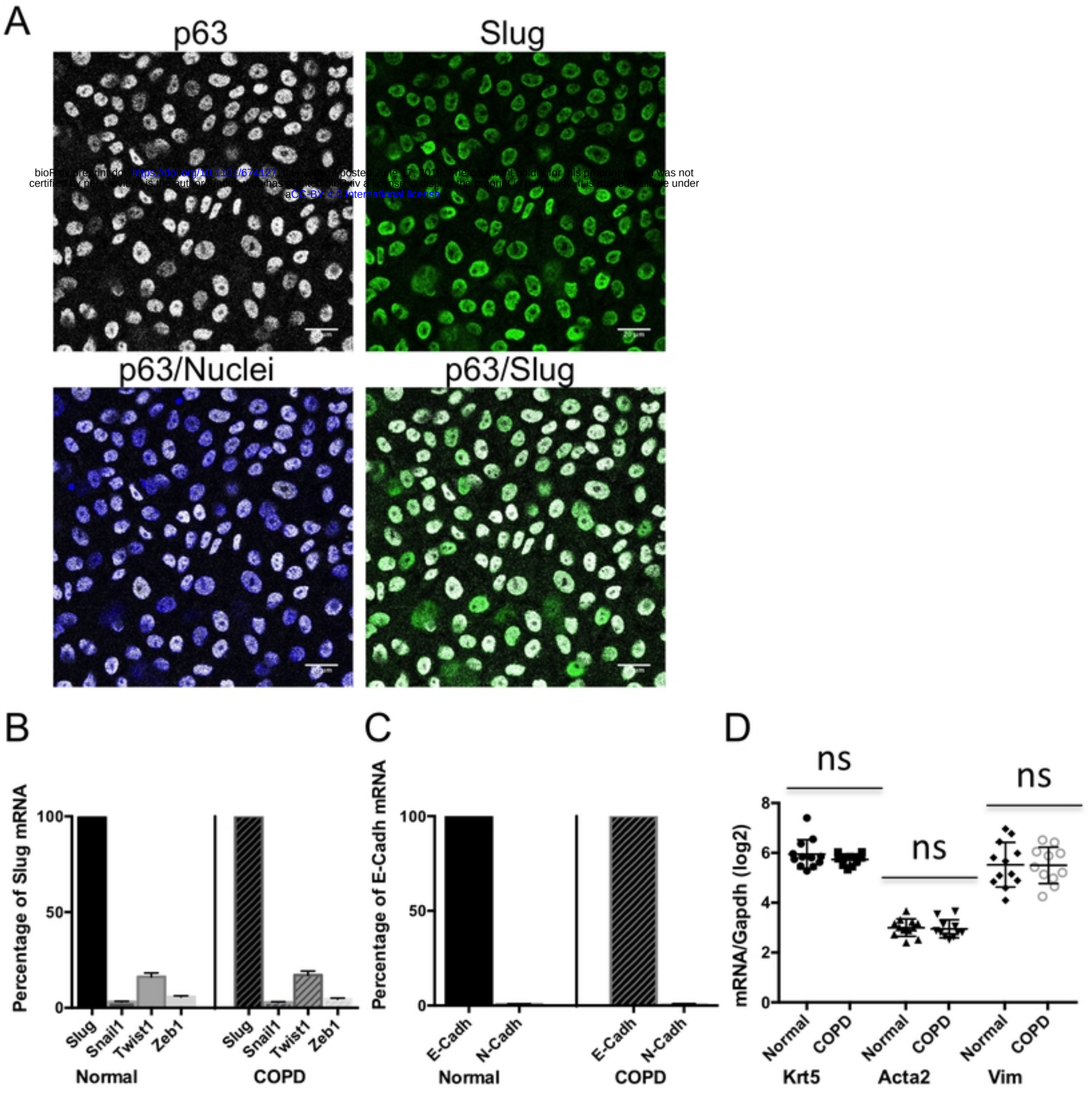


Figure 1

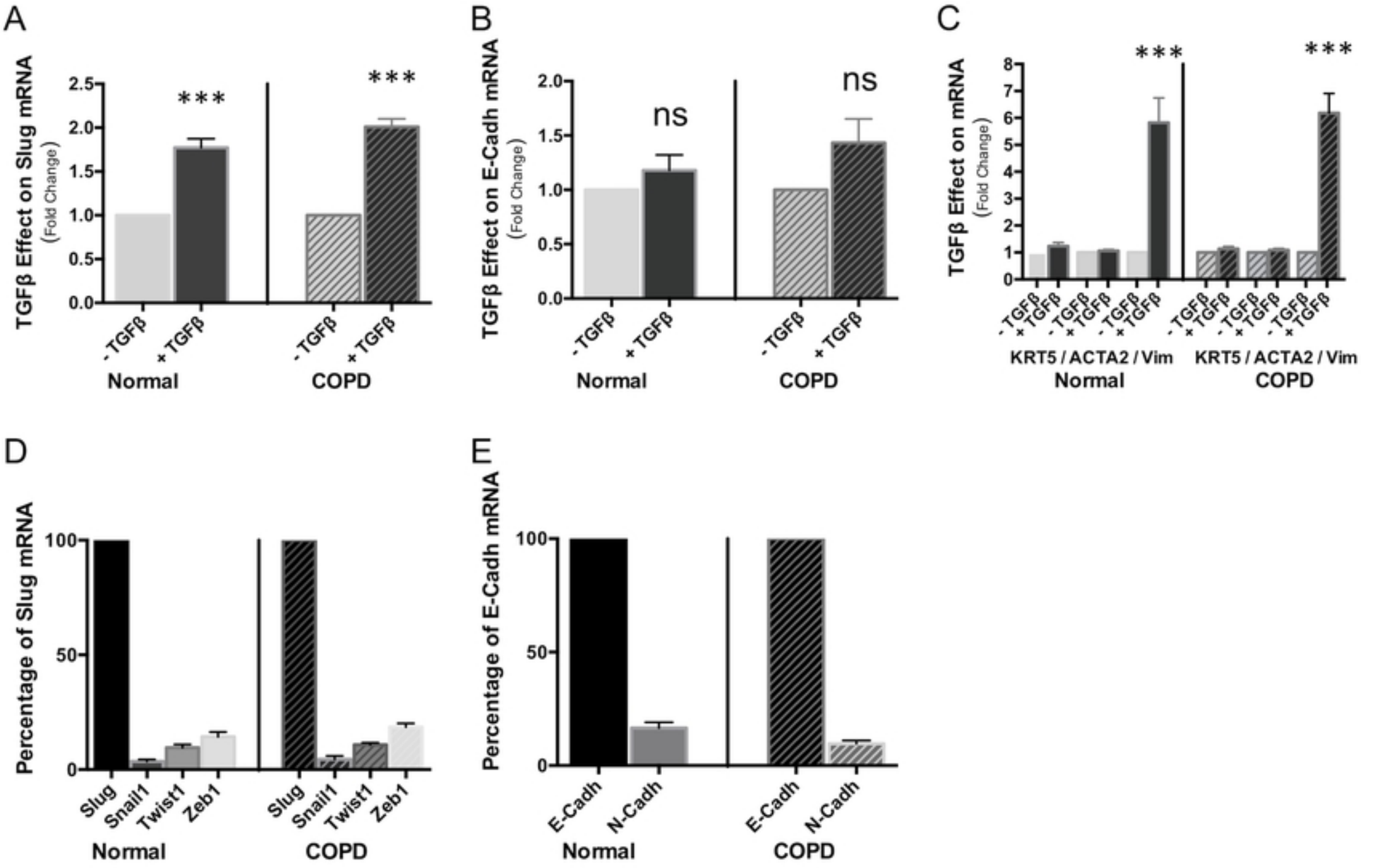
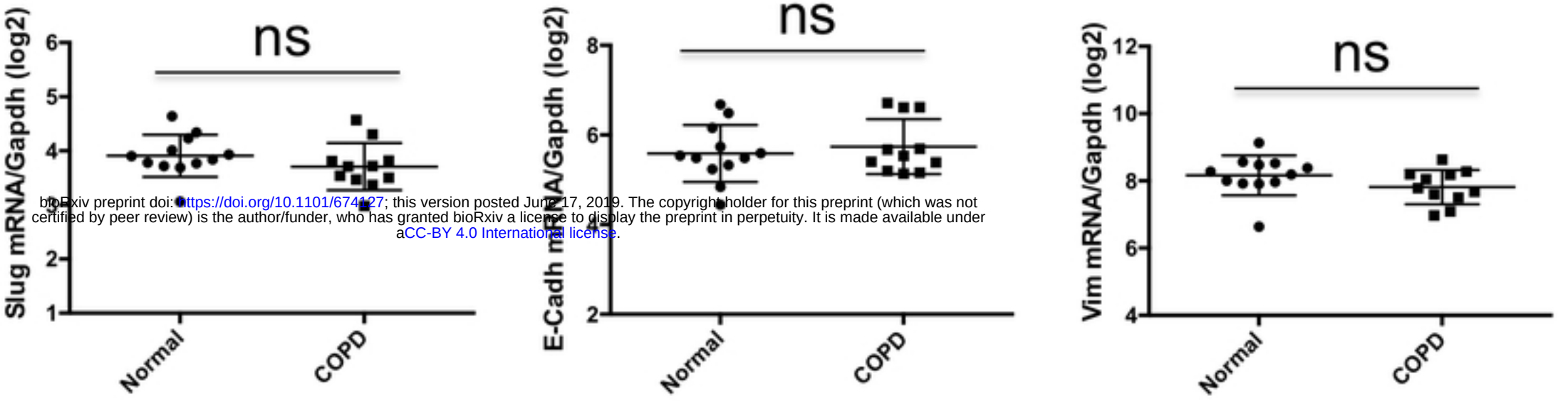
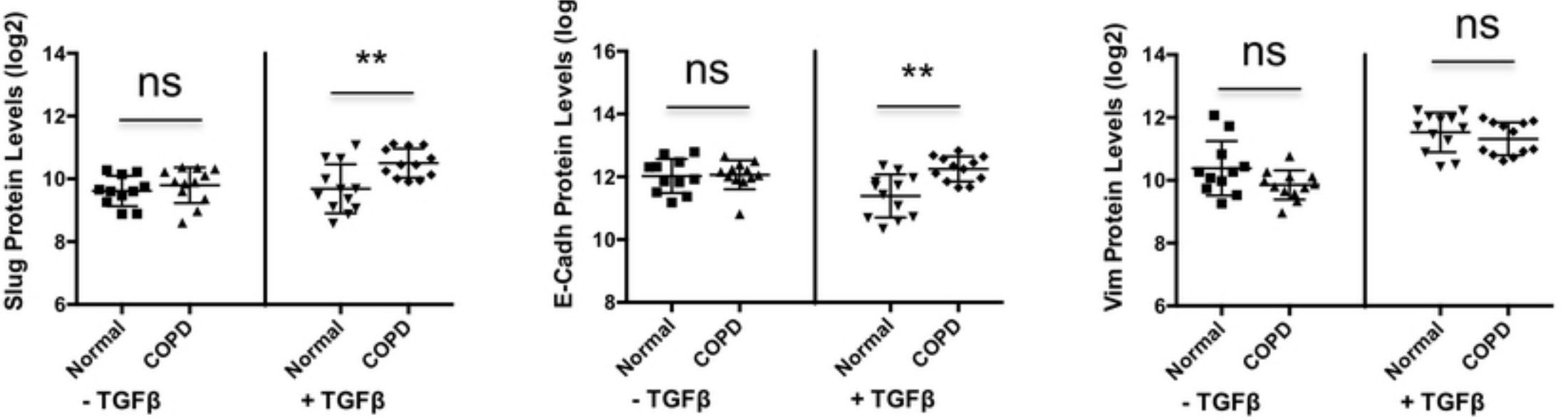

A
TGFβ Effect on Slug mRNA
(Fold Change)
2.5
2.0
1.5
1.0
0.5
0.0
***
***
- TGFβ
+ TGFβ
- TGFβ
+ TGFβ
Normal
COPD
B
TGFβ Effect on E-Cadh mRNA
(Fold Change)
2.0
1.5
1.0
0.5
0.0
ns
ns
- TGFβ
+ TGFβ
- TGFβ
+ TGFβ
Normal
COPD
C
TGFβ Effect on mRNA
(Fold Change)
8
7
6
5
4
3
2
1
0
***
***
- TGFβ
+ TGFβ
- TGFβ
+ TGFβ
- TGFβ
+ TGFβ
- TGFβ
+ TGFβ
- TGFβ
+ TGFβ
- TGFβ
+ TGFβ
KRT5 / ACTA2 / Vim
KRT5 / ACTA2 / Vim
Normal
COPD
D
Percentage of Slug mRNA
100
50
0
Slug
Snail1
Twist1
Zeb1
Slug
Snail1
Twist1
Zeb1
Normal
COPD
E
Percentage of E-Cadh mRNA
100
50
0
E-Cadh
N-Cadh
E-Cadh
N-Cadh
Normal
COPD

Figure 2

A



B



C

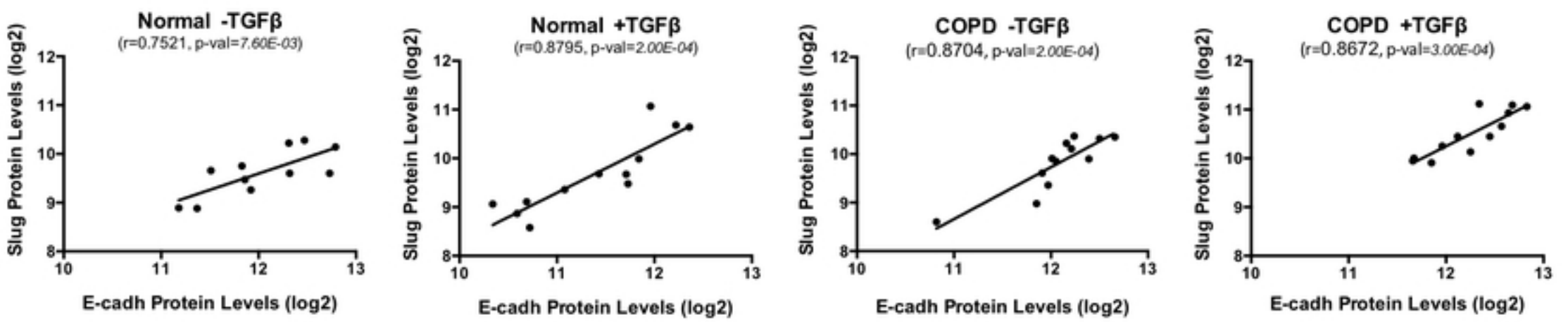
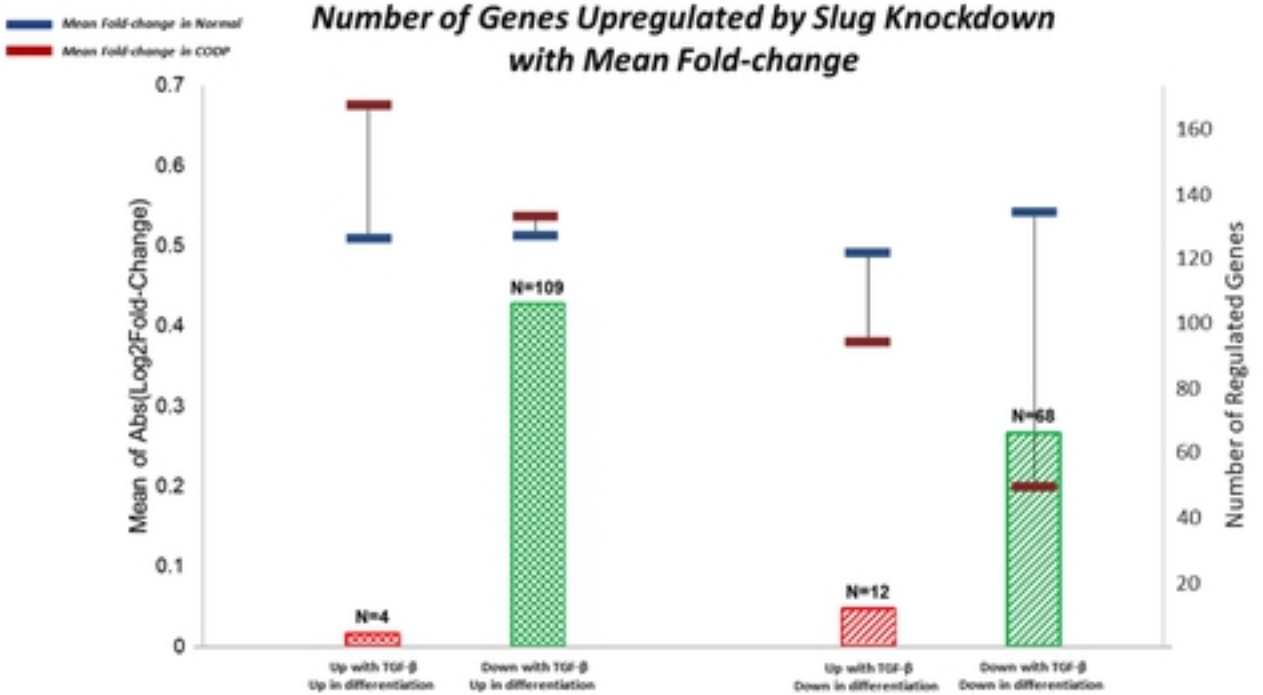


Figure 3

A


Mean Fold-change in Normal
Mean Fold-change in COOP
Number of Genes Upregulated by Slug Knockdown with Mean Fold-change
Mean of Abs(Log2Fold-Change)
Number of Regulated Genes
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
160
140
120
100
80
60
40
20
N=4
N=109
N=12
N=68
Up with TGF-β Up in differentiation
Down with TGF-β Up in differentiation
Up with TGF-β Down in differentiation
Down with TGF-β Down in differentiation

B

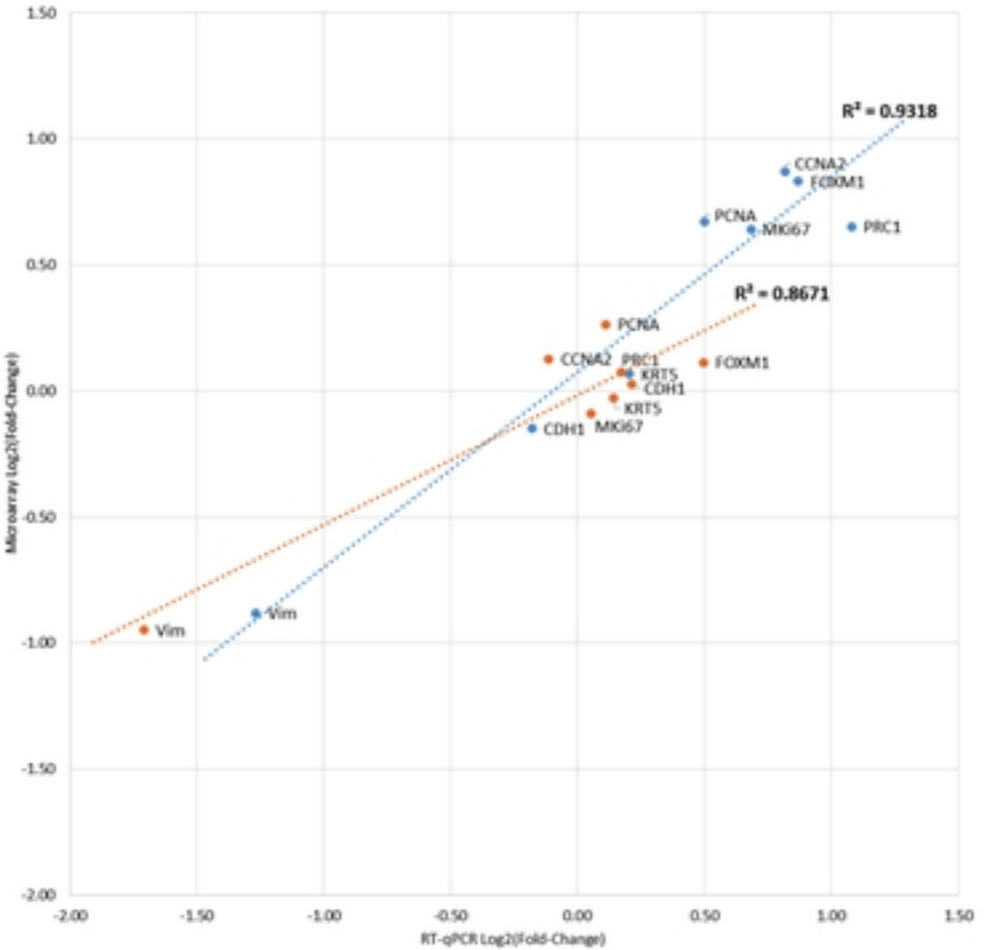
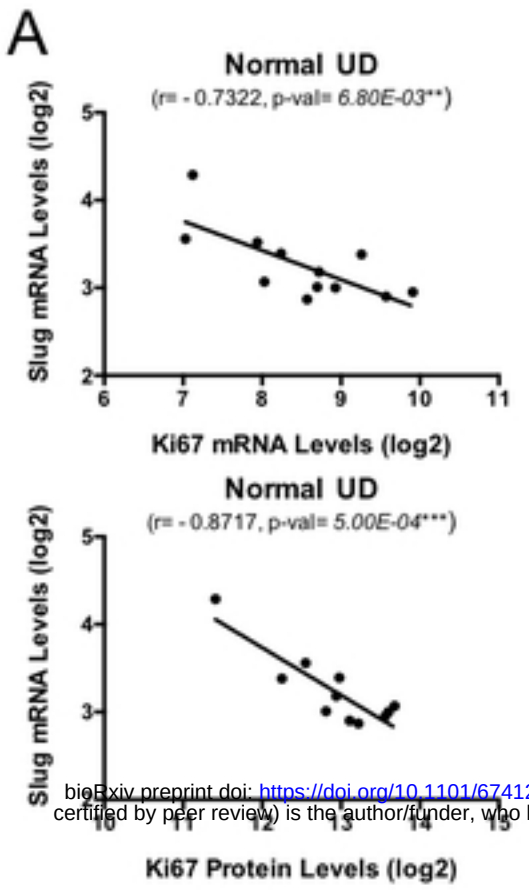
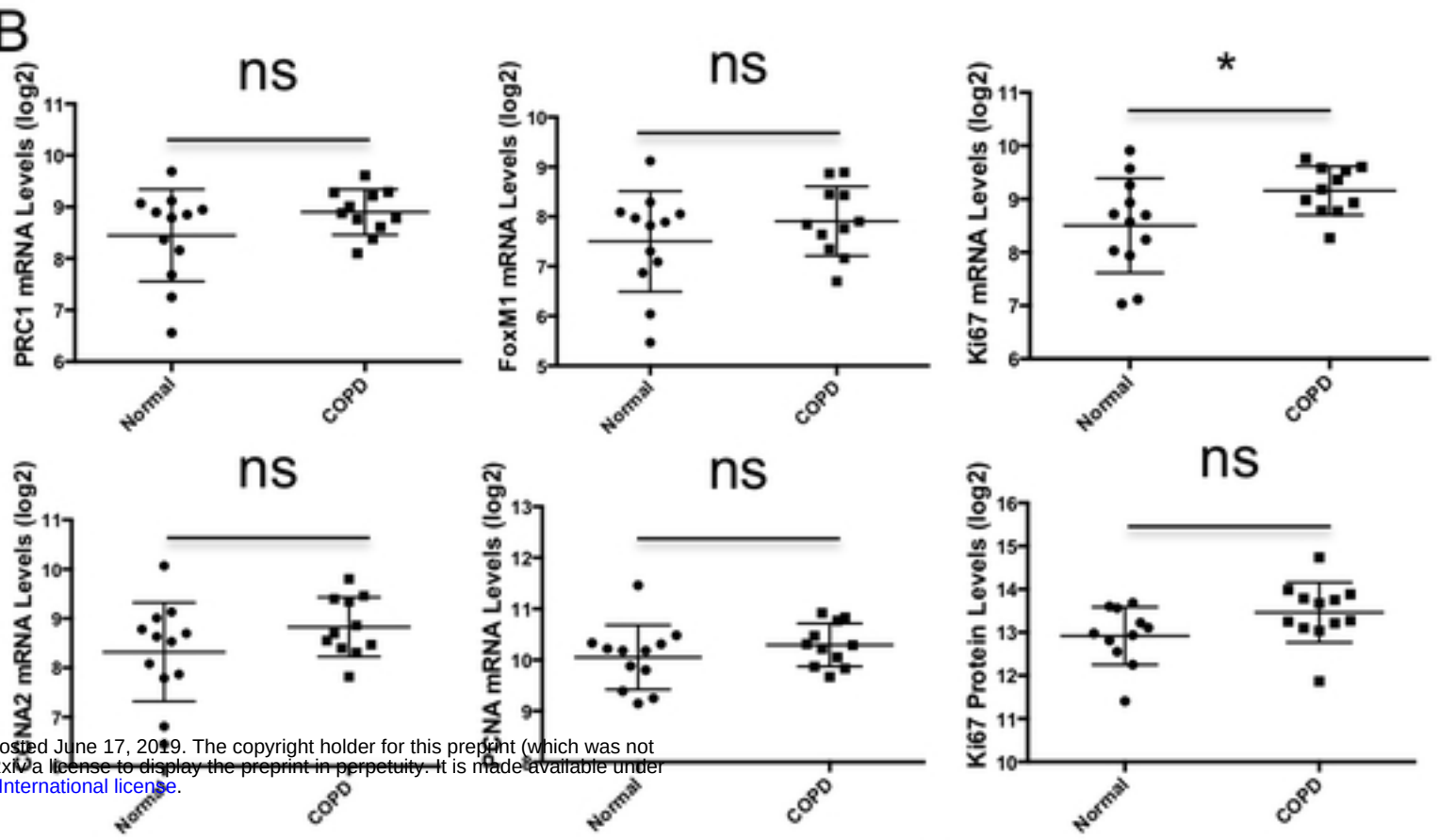
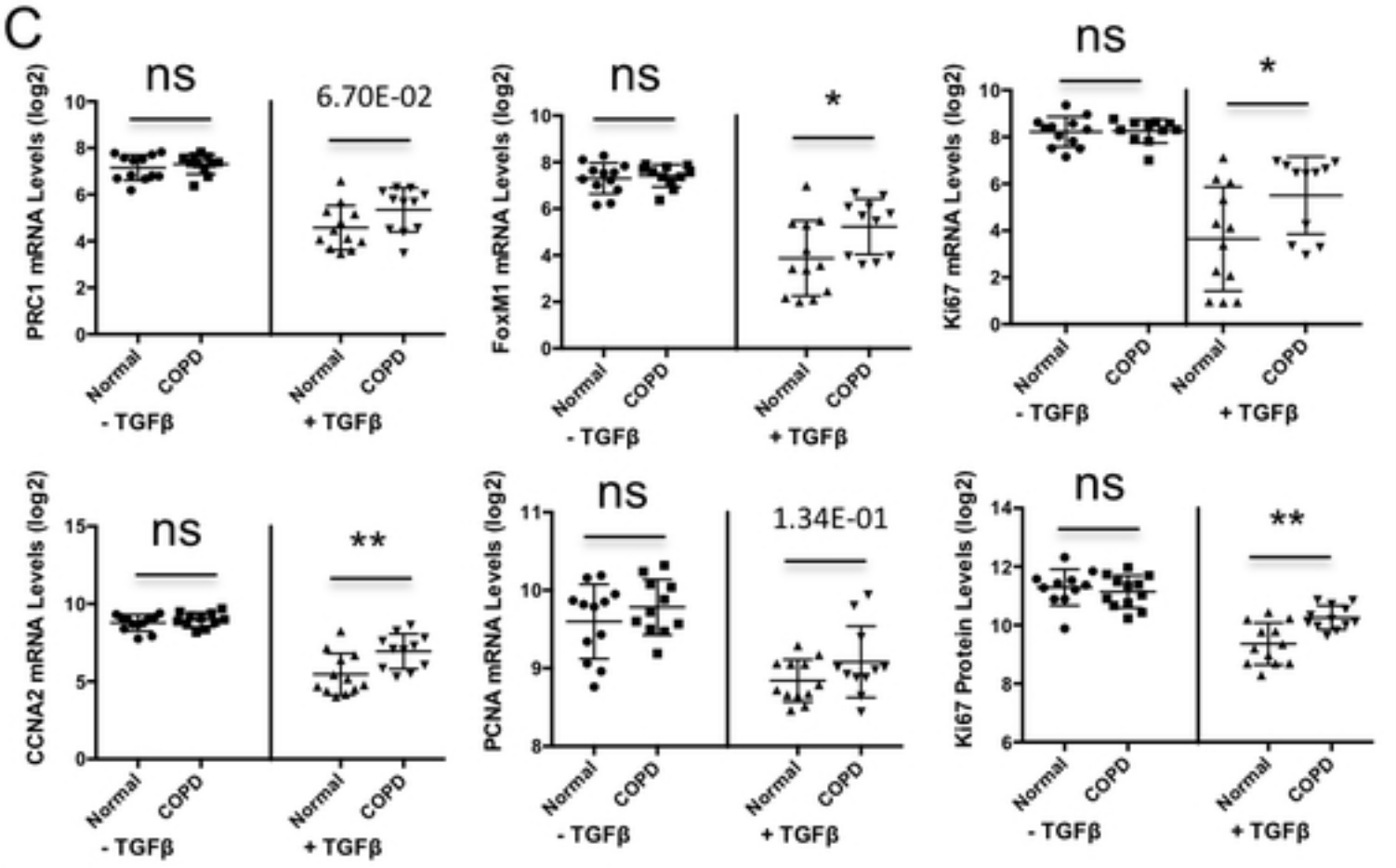
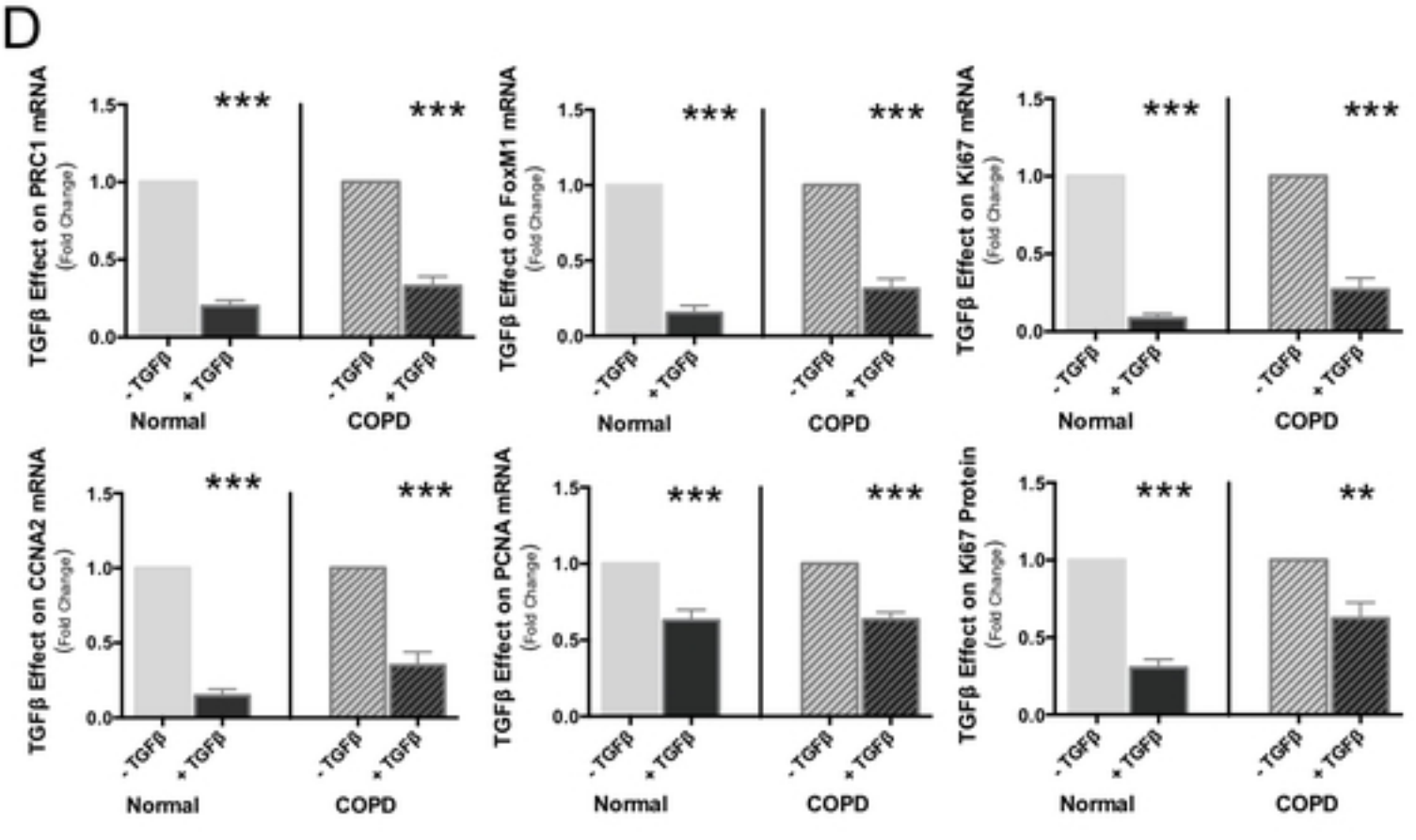

R² = 0.9318
R² = 0.8671
CCNA2
FOXM1
PCNA
MKI67
PRC1
KRT5
CDH1
Vim
Microarray Log2(Fold-Change)
RT-qPCR Log2(Fold-Change)
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00

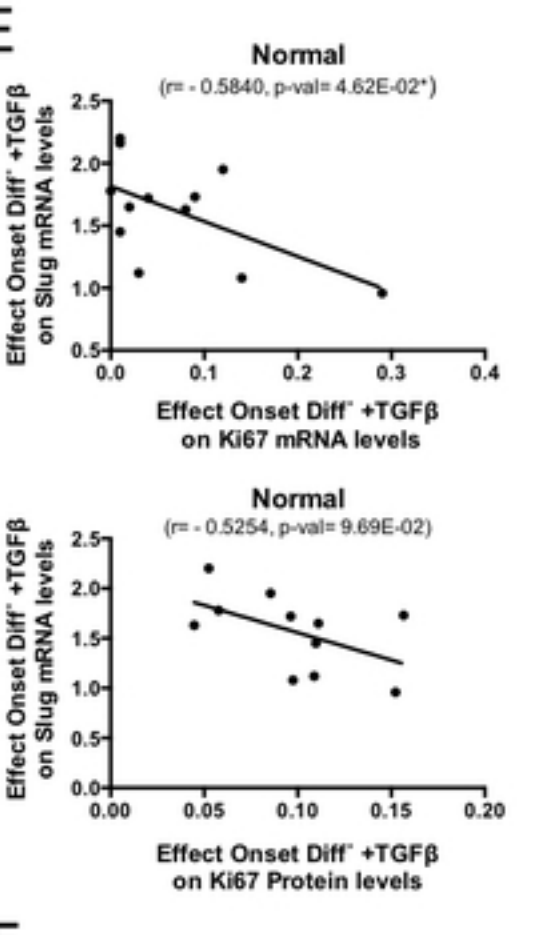
Figure 4











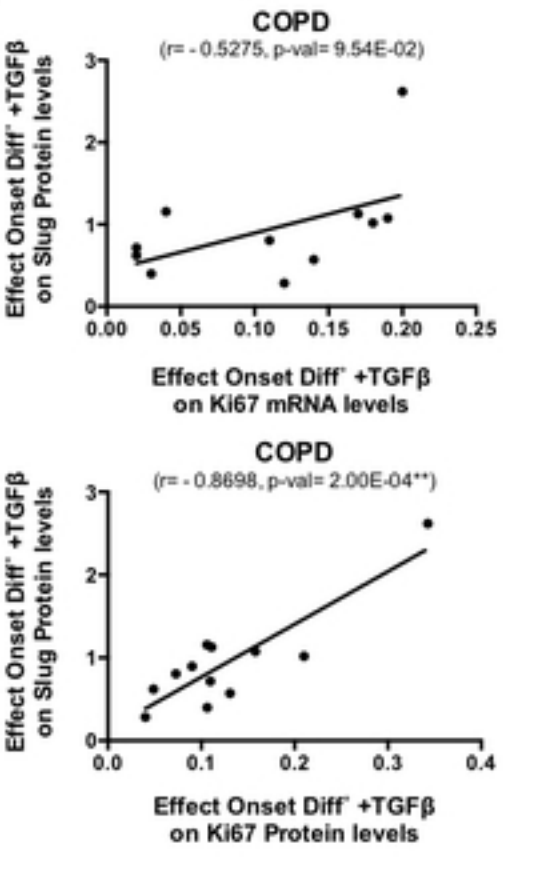


Figure 5

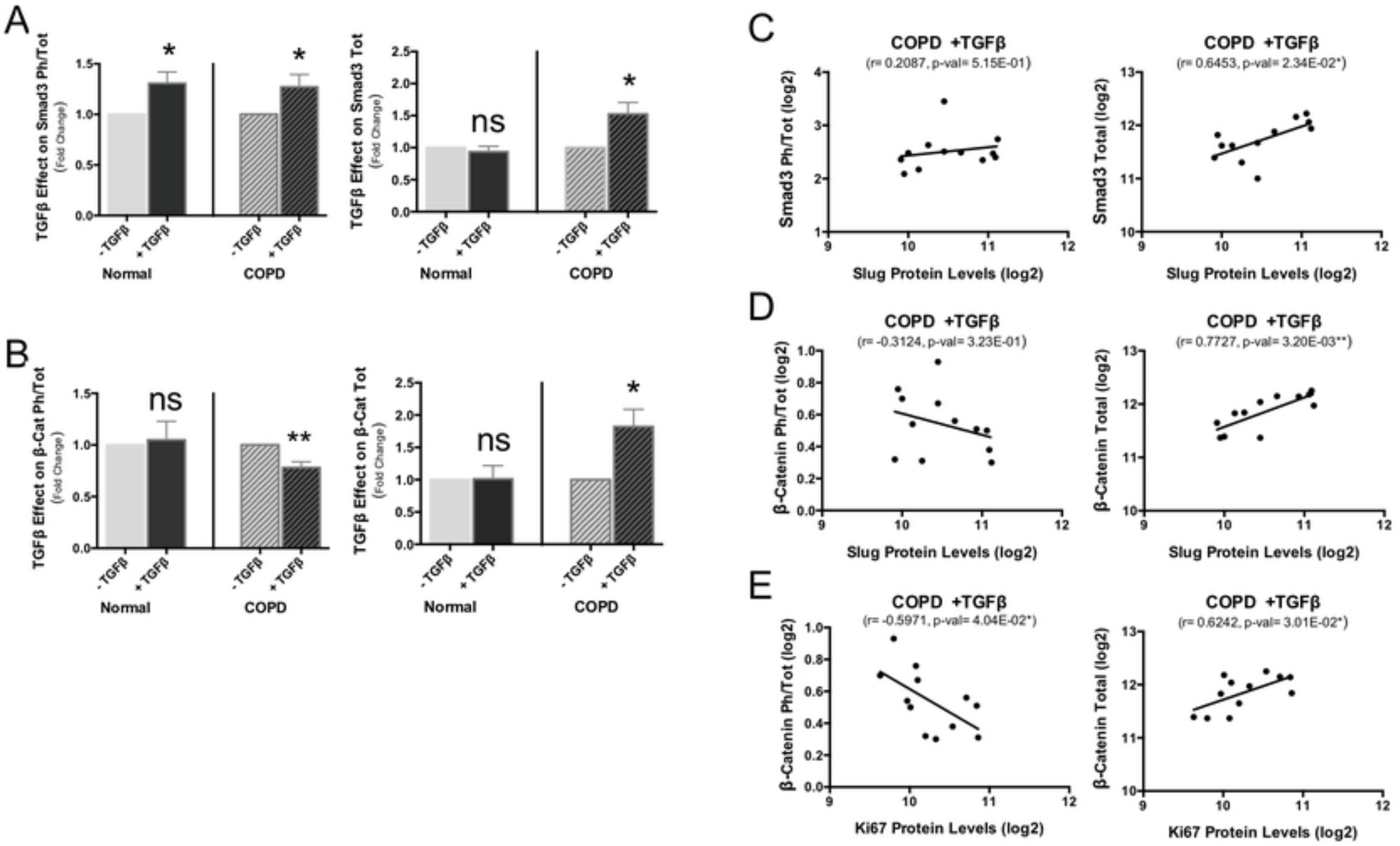

A
TGFβ Effect on Smad3 Ph/Tot
(Fold Change)
1.5
1.0
0.5
0.0
*
*
-TGFβ
+TGFβ
-TGFβ
+TGFβ
Normal
COPD
TGFβ Effect on Smad3 Tot
(Fold Change)
2.5
2.0
1.5
1.0
0.5
0.0
ns
*
-TGFβ
+TGFβ
-TGFβ
+TGFβ
Normal
COPD
B
TGFβ Effect on β-Cat Ph/Tot
(Fold Change)
1.5
1.0
0.5
0.0
ns
**
-TGFβ
+TGFβ
-TGFβ
+TGFβ
Normal
COPD
TGFβ Effect on β-Cat Tot
(Fold Change)
2.5
2.0
1.5
1.0
0.5
0.0
ns
*
-TGFβ
+TGFβ
-TGFβ
+TGFβ
Normal
COPD
C
COPD +TGFβ
(r= 0.2087, p-val= 5.15E-01)
Smad3 Ph/Tot (log2)
4
3
2
1
9
10
11
12
Slug Protein Levels (log2)
COPD +TGFβ
(r= 0.6453, p-val= 2.34E-02*)
Smad3 Total (log2)
13
12
11
10
9
10
11
12
Slug Protein Levels (log2)
D
COPD +TGFβ
(r= -0.3124, p-val= 3.23E-01)
β-Catenin Ph/Tot (log2)
1.0
0.8
0.6
0.4
0.2
0.0
9
10
11
12
Slug Protein Levels (log2)
COPD +TGFβ
(r= 0.7727, p-val= 3.20E-03**)
β-Catenin Total (log2)
13
12
11
10
9
10
11
12
Slug Protein Levels (log2)
E
COPD +TGFβ
(r= -0.5971, p-val= 4.04E-02*)
β-Catenin Ph/Tot (log2)
1.0
0.8
0.6
0.4
0.2
0.0
9
10
11
12
Ki67 Protein Levels (log2)
COPD +TGFβ
(r= 0.6242, p-val= 3.01E-02*)
β-Catenin Total (log2)
13
12
11
10
9
10
11
12
Ki67 Protein Levels (log2)

Figure 6