HHS Public Access

Author manuscript

J Thorac Oncol. Author manuscript; available in PMC 2016 August 01.

Published in final edited form as:

J Thorac Oncol. 2015 August ; 10(8): 1181–1188. doi:10.1097/JTO.0000000000000592.

Epigenetic Repression of *CCDC37* and *MAP1B* Links Chronic Obstructive Pulmonary Disease to Lung Cancer

Mathewos Tessema, PhD[*], **Christin M. Yingling, BS**[*], **Maria A. Picchi, MPH**[*], **Guodong Wu, PhD**[*], **Yushi Liu, PhD**[*,†], **Joel L. Weissfeld, PhD**[‡], **Jill M. Siegfried, PhD**[§], **Yohannes Tesfaigzi, PhD**[ϕ], and **Steven A. Belinsky, PhD**[*]

[*]Lung Cancer Program, Lovelace Respiratory Research Institute, Albuquerque, New Mexico

[ϕ]COPD Program, Lovelace Respiratory Research Institute, Albuquerque, New Mexico

[‡]Department of Epidemiology, University of Pittsburgh, Pittsburgh, Pennsylvania

[§]Department of Pharmacology, University of Minnesota, Minneapolis, Minnesota

Abstract

Introduction—Lung cancer and chronic obstructive pulmonary disease (COPD) share environmental risk factors. COPD also increases the risk of lung cancer; however, the molecular mechanisms are unclear.

Methods—An epigenome-wide association study of lung tumors and cancer-free lung tissue (CFLT) pairs from non-small cell lung cancer (NSCLC) cases with (n=18) or without (n=17) COPD was conducted using the HumanMethylation450 beadchip (HM450K). COPD-associated methylation of top-ranked genes was confirmed in a larger sample set, independently validated, and their potential as sputum-based biomarkers was investigated.

Results—Methylation of *CCDC37* and *MAP1B* was more prevalent in lung tumors from COPD than non-COPD cases [54/71 (76%) vs. 20/46 (43%), p=0.0013] and [48/71 (68%) vs. 17/46 (37%), p=0.0035], respectively, after adjustment for age, sex, smoking status, and tumor histology. HM450K probes across *CCDC37* and *MAP1B* promoters showed higher methylation in tumors than CFLT with the highest methylation seen in tumors from COPD cases ($p<0.05$). These results were independently validated using The Cancer Genome Atlas data. *CCDC37* methylation was more prevalent in sputum from COPD than non-COPD smokers ($p<0.005$) from two cohorts. *CCDC37* and *MAP1B* expression was dramatically repressed in tumors and CFLT from COPD than non-COPD cases, $p<0.02$.

Address for Corresponding and Reprints: Steven A. Belinsky (SAB) and Mathewos Tessema (MT), Lovelace Respiratory Research Institute, 2425 Ridgecrest Dr. SE, Albuquerque, NM. 87108, U.S.A. Phone: 505-348-9465 (SAB), 505-348-9123 (MT); Fax: 505-348-4990 (SAB), 505-348-8567 (MT); sbelinsk@LRRI.org (SAB), mtessema@LRRI.org (MT).

[†]YL has now moved to Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana.

Conflict of interest: The authors have no conflict of interest.

This article has supplemental data available online.

Disclosure: This study was supported by R01 CA164782 (Drs. Belinsky and Tesfaigzi) and in part by R21 CA161561 (Dr. Tessema), P30 CA118100 (Dr. Belinsky), R01 CA097356 (Dr. Belinsky), and P50 CA090440 (Dr. Siegfried).

Conclusions—The reduced expression of *CCDC37* and *MAP1B* associated with COPD likely predisposes these genes to methylation that in turn, may contribute to lung cancer.

Keywords

EWAS; Methylation; COPD; NSCLC; Sputum

INTRODUCTION

Lung cancer and COPD are the leading causes of morbidity and mortality worldwide.[1] In the US, COPD is the third major cause of death next to cardiovascular and malignant diseases, while lung cancer is the largest cause of cancer mortality, being responsible for approximately a third of all cancer-related deaths.[2] In addition to sharing similar etiology, as cigarette smoking is the major risk factor for both diseases, multiple epidemiological studies have consistently demonstrated a strong link between COPD and lung cancer (recently reviewed in[3]). The presence of COPD increases the risk of lung cancer by 2–4-fold suggesting that the pathogenesis of the two diseases may also follow similar pathways.[4,5] The enormous human life and economic loss associated with these diseases and the lack of effective treatment for both have increasingly attracted intensive research. However, the molecular mechanisms by which COPD contributes to lung cancer development remain largely uncharacterized.[3]

Epigenetic silencing of genes through histone modifications and methylation of cytosines in promoter regions affects the expression of hundreds of genes and is a major causal event during lung cancer development.[6,7] The commonality and involvement of gene methylation in carcinogenesis together with the development of sensitive methylation assays have led to studies focused on establishing the utility of methylation as a biomarker in screening for cancer risk, prevention, treatment, and prognosis. Our group was the first to establish that cancer associated methylation of genes could be detected in sputum samples from high risk smokers prior to clinical diagnosis of lung cancer risk.[8,9] We have also evaluated the commonality of aberrant methylation of candidate genes in COPD and lung cancer. Methylation analysis of a panel of eight lung cancer-related genes in sputum DNA from 1,267 cancer-free smokers showed that methylation of *p16* (*CDKN2A*) and *GATA4* were significantly associated with lower lung function.[10] Similarly, Suzuki *et al*. used a candidate gene approach to identify that the genetic and epigenetic profile of lung tumors from COPD patients was distinct from that of non-COPD cases.[11] These findings suggest that the development of COPD through cigarette smoke induced DNA damage and remodeling drives the epigenetic silencing of cell growth-related genes, which could in turn contribute to the development of lung cancer. Identification of these abnormalities in sputum samples from smokers with COPD may provide new biomarkers for COPD development and progression, and early cancer detection.

Thus, the objective of this study was to conduct an epigenome wide association study (EWAS) to compare the methylation changes across the whole genome of lung tumor-normal pairs from NSCLC patients with or without COPD. The EWAS was performed using the HM450K that generates methylation data for over 480,000 CpGs genome-wide. For this

study, we focused on approximately 160,000 probes that are located within CpG islands in gene promoter regions (TSS 1500 as defined in[12]), excluding regions normally regulated by methylation such as imprinted and X-chromosome genes. Methylation of the identified genes was confirmed in larger sample sets using different methylation assays and the results were independently validated using HM450K data from The Cancer Genome Atlas (TCGA) database. Expression of selected genes was quantified in lung tumor-normal pairs and the relationship between methylation and expression was evaluated. Finally, the utility of detecting methylated genes in sputum for potential development of non-invasive biomarkers for COPD was investigated in two geographically distinct cohorts.

MATERIALS AND METHODS

Study subjects, COPD status, and samples

Frozen lung tumor samples from 117 NSCLC patients (71 with and 46 without COPD) were obtained from tumor banks at the University of New Mexico (UNM) and the Mayo Clinic. Only tumors with ≥ 75% tumor purity were considered suitable for study. For a subset of these cases, cancer-free lung tissues (CFLT) collected from sites most distant to the tumor in the resected lobe were evaluated along with the tumor pair. COPD status and severity was defined according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification.[13] Sputum samples from two longitudinal smoker cohorts, the Lovelace Smokers cohort (LSC, participants from the Albuquerque, NM metropolitan area since 2001[14–17]) and the Pittsburgh Lung Screening Study (PLuSS, established in 2002[18]) were studied. Characteristics of subjects whose tissue/sputum samples were studied are shown in Tables 1 and 2. Normal human bronchial epithelial cells (NHBEC) collected from 10 cancer-free smokers at UNM through diagnostic bronchoscopy[9] and peripheral blood mononuclear cells (PBMC) obtained from 10 healthy donors were used as normal control. All samples were obtained with written informed consent from patients, and the study was approved by each institute's Ethics Committee. In addition, five human bronchial epithelial cell lines (HBEC) and five human small airway epithelial cell lines (HSAEC) immortalized as described,[19] and 23 NSCLC cell lines were also studied. The cell lines used in this study and their sources, authentications, and handling methods are described in the online supplement.

DNA methylation and gene expression analysis

DNA extraction, modification, and methylation analysis using Combined Bisulfite Modification and Restriction Analysis (CoBRA), Methylation-Specific PCR (MSP), and (for sputum samples) Nested MSP assays were conducted as described.[20–22] The primer sequences and amplification conditions are shown in supplementary Tables S1 and S2. Genome-wide methylation of lung tumor-normal pairs was quantified using HM450K (Illumina, San Diego, CA).[12] Methylation data for lung tumor-normal pairs evaluated by TCGA was obtained from https://tcga-data.nci.nih.gov/tcga/.

Gene expression analysis was conducted using RNA isolated from lung tumor-CFLT pairs and cell lines as described.[23] Briefly, 1 μg total RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit from Applied Biosystems (Foster City, CA)

according to the manufacturer's protocol. Expression of *CCDC37*, *MAP1B*, and *BETA-ACTIN* was quantified using Hs00403623_m1, Hs00195485_m1, and 4310881E TaqMan assays (Applied Biosystems), respectively, as described.[24,25] All samples were analyzed at least twice in duplicate and expression levels were calculated using ΔΔCT method.[26]

Statistical analysis

Gene methylation and patient characteristics including age, sex, smoking, COPD status, and tumor histology were summarized with mean and standard deviation for continuous variables and proportions for categorical variables. HM450K data for lung tumors from COPD and non-COPD cases was compared using β-regression analysis.[27] The association between methylation and patient characteristics was assessed by Fisher's exact test. Gene expression levels were compared using two-tailed T-test for unequal variance (COPD vs. non-COPD cases) and pairwise T-test (tumor vs. normal pairs). All analyses were conducted in SAS 9.2 and R 2.14.

RESULTS

Genome-wide Screening of Cytosine Methylation in Lung Tumor-Normal Pairs

Tumors from 18 COPD and 17 non-COPD lung adenocarcinoma cases along with 6 CFLT pairs (3 COPD and 3 non-COPD) and PBMC from 3 cancer-free donors were screened for methylation using HM450K. A total of 138 probes within 43 genes showed significantly higher methylation in lung tumors from COPD compared to non-COPD cases. Exclusion of probes with features (shown in Figure 1) suggesting limited biomarker and/or gene-silencing potential narrowed the candidates to 108 probes within 30 genes. Methylation of these 30 genes was evaluated by CoBRA and MSP assays using PBMC from cancer-free smokers (n=5) and NSCLC cell lines (n=23). Based on the CoBRA/MSP results, 6 genes that were methylated in PBMC and 8 genes that were rarely (< 5%) methylated in NSCLC cell lines were excluded. The final 16 genes, which were verified by CoBRA/MSP for showing methylation in ≥2 NSCLC cell lines but not in PBMC, were evaluated by MSP using the same lung tumor-CFLT pairs interrogated by HM450K. The prevalence for methylation of 10 of the 16 genes was increased by ≥15% in tumors from COPD compared to non-COPD cases (Table S3). Methylation of these 10 genes was then evaluated in the remaining tumor-CFLT pairs and the aggregate results are summarized in Table 3. After adjustment for age, sex, smoking status, and tumor histology, lung tumors from COPD cases showed significantly higher prevalence for methylation of *CCDC37* and *MAP1B* ($p<0.005$). *CCDC37* was methylated in 54/71 (76%) lung tumors from COPD compared to 20/46 (43%) non-COPD cases, while *MAP1B* was methylated in 48/71 (68%) tumors with COPD versus 17/46 (37%) non-COPD subjects. Methylation of the remaining 8 genes was not significantly different between lung tumors from COPD and non-COPD cases.

Independent Quantitative Validation of *CCDC37* and *MAP1B* methylation

The level of methylation across *CCDC37* and *MAP1B* promoters was quantified using HM450K. The locations of 8 *CCDC37* and 6 *MAP1B* probes with respect to the promoter CpG island and first exon of each gene are depicted in the X-axes of Figure 2 and 3. Comparison of these probes using our HM450K data for normal lung and tumors from

COPD and non-COPD cases revealed that methylation of all 8 *CCDC37* probes was significantly increased in lung tumors compared to normal lung ($p<0.05$). Each probe showed the highest methylation in tumors from COPD cases including 4 probes (*φ) with significantly higher methylation in tumors from COPD than non-COPD subjects (Figure 2A). Using $\beta \geq 0.2$ as a cut-off to define hypermethylation, the prevalence for aberrant methylation of each probe was also increased significantly from normal lung to tumors in non-COPD cases and further increased in tumors from COPD cases (Table S4). These results were independently validated using the TCGA data that was similarly generated by HM450K from a separate lung cancer cohort. TCGA data for 10 normal lung tissue and 61 lung tumors from NSCLC cases whose spirometry data was available and met with our classification criteria of COPD (n=17) or non-COPD (n=44) were analyzed. Characteristics of patients whose samples were evaluated by TCGA and the data used in this study are shown Tables S5. With the exception of probe cg20952286, which is missing for all TCGA samples, the level and prevalence for methylation of the remaining 7 *CCDC37* probes was similar to our data (Figure 2B, Table S6). The identical results supported a combined analysis of the TCGA and our data sets, which revealed that the prevalence and level of methylation of all 7 *CCDC37* probes were significantly higher in lung tumors from COPD than non-COPD cases (Figures 2C, Table S7). Similarly, quantitative analysis of all 6 *MAP1B* probes using our samples (Figures 3A, Table S4), TCGA data (Figures 3B, Table S6), and combined analysis of the two data sets (Figures 3C, Table S7) also showed the lowest methylation in normal lung that was significantly increased in lung tumors from COPD cases and to a lesser extent in tumors from non-COPD cases. However, unlike *CCDC37,* the higher methylation of each *MAP1B* probe in COPD than non-COPD tumors was not statistically significant.

Methylation of *CCDC37* and *MAP1B* in Sputum

The potential use of *CCDC37* and *MAP1B* methylation as non-invasive biomarkers for COPD was evaluated using sputum samples from cancer-free smokers with or without COPD. The prevalence for methylation of *CCDC37* in sputum samples from LSC cohort was 60% (60/100) in COPD cases compared to 36% (40/111) in control (Table 4). After adjustment for age, sex, and smoking status, methylation of *CCDC37* in the sputum of COPD cases was significantly higher than control, odds ratio (OR) 2.74, 95% confidence interval (95% CI) 1.55—4.90 (p=0.0006). This result was independently validated using sputum samples from the geographically distinct PLuSS cohort, where *CCDC37* methylation was highly prevalent in COPD cases than controls, OR=2.79, 95% CI 1.37—5.66 (p=0.0046). A combined analysis of the two cohorts also confirmed that *CCDC37* was more frequently methylated in the sputum of COPD cases than controls, OR=2.55, 95% CI 1.68—3.87 ($p<0.0001$) (Table 4). In contrast, the prevalence for *MAP1B* methylation in sputum samples from the LSC cohort was similar between COPD cases (24%, 24/100) and controls (23%, 25/111), OR=1.10, 95% CI 0.58—2.08, p=0.78) and thus, was not analyzed in PLuSS samples.

CCDC37 and *MAP1B* expression in COPD and Lung Cancer

Quantitative analysis of *CCDC37* and *MAP1B* transcripts in lung tumor-CFLT pairs from adenocarcinoma patients revealed that CFLT from COPD cases expressed significantly

lower levels of both genes compared to CFLT from non-COPD cases, p=0.020 and 0.010, respectively (Figure 4A—B). Similarly, lung tumors from COPD cases expressed significantly lower *CCDC37* and *MAP1B* than tumors from non-COPD cases, p=0.010 and 0.003, respectively (Figure 4C—D). Comparison of *CCDC37* expression between lung tumors versus CFLT pairs within COPD or non-COPD cases separately also revealed significant repression in tumors of either group (Figure 4E). In contrast, *MAP1B* expression in lung tumors from COPD or non-COPD cases was similar to the CFLT pairs in the corresponding COPD group, likely reflecting the heterogeneity for expression between cell types comprising the normal tissue (Figure 4F).

DISCUSSION

This EWAS identified that epigenetic mediated repression of *CCDC37* and *MAP1B* is significantly associated with COPD and lung cancer. The dramatic reduction in expression of these genes is initiated in the cancer-free COPD lung. Expression is further reduced or maintained at lower levels in lung tumors from COPD cases with accompanying promoter hypermethylation. Together, these findings fit with Clark and Melki's model [28] that states methylation in cancer is a consequence rather than the cause of gene silencing. Hence, repression of these genes in COPD prior to lung cancer is likely predisposing for the increased methylation during lung carcinogenesis and may contribute to the increased risk of lung cancer observed in COPD patients. Moreover, *CCDC37* methylation was more prevalent in the sputum of cancer-free smokers with than without COPD. This finding, which was independently validated in a geographically distinct cohort, supports *CCDC37* silencing by methylation as an early event linked with COPD prior to cancer development. Although aberrant methylation of *CCDC37* in lung cancer and *MAP1B* in colorectal cancer has been previously reported,[29,30] this is the first study to link these epigenetic changes to COPD and lung cancer.

Epigenetic modifications are key molecular mechanisms that allow cells to regulate gene expression networks in response to environmental stimuli.[31] Cigarette smoke, the major risk factor for lung cancer and COPD, is an example of such environmental stimuli.[3] Our previous studies have demonstrated that *in vitro* exposure of HBEC to tobacco carcinogens initially repress expression of genes and microRNAs through histone modifications (such as deacetylation, methylation, and/or demethylation), which create a closed chromatin structure repressive to transcription.[7] Long-term epigenetic silencing of such repressed genes occurs during malignant transformation through hypermethylation of the promoter CpG islands. Studies from our group and others have shown that genes with constitutively lower expression and promoter activity are more prone to becoming methylated during cancer development within that organ.[28,32] These prior findings and concepts appear operative for regulating expression of *CCDC37* and *MAP1B* during the development of COPD and lung cancer.

CCDC37 is one of a large family of genes encoding for coiled-coil domain containing (CCDC) proteins. Currently there is no publication regarding the function of *CCDC37*. However, Yamamoto et al.[33] recently demonstrated that MAI1, a homologue of CCDC37 in Chlamydomonas reinhardtii, a single celled chlorophyte (green algae), is an essential

regulator of ciliary motility and speculated that CCDC37 might have similar function in humans and other higher organisms. Cilia motility plays a key role in clearing mucous from the lungs and is severely reduced in the COPD patient. Accumulating mucous will result in increased pulmonary inflammation and likely enhanced oxidative damage to DNA that in turn could play a role in further gene silencing through promoter hypermethylation.[17,34]

MAP1B encodes for microtubule-associated protein 1B (MAP1B), one of the major cytoskeletal proteins that is involved in various cellular activities including actin-based cell motility, molecular trafficking, autophagy, and cancer.[35–37] This large (2464 amino acids) protein is first synthesized as a polyprotein precursor which is rapidly cleaved to give rise to the respective heavy and light chains, termed MAP1B-HC and MAP1B-LC1.[38] Previous studies have demonstrated that MAP1B-LC1 interacts with Pes1 and p53 to modulate cell proliferation and apoptosis.[39,40] Thus, loss of function of this gene may contribute to the cancer cell's ability evade death signals and proliferate. Together, silencing of *CCDC37* and *MAP1B* during development of COPD may contribute to the development of malignant NSCLC.

Our group has been instrumental in demonstrating and validating aberrant methylation of genes in sputum samples from high risk smokers as a molecular marker for predicting lung cancer risk.[8,9] We and others have also reported that detection of methylation in sputum samples could link lung cancer with other risk factors such as smoking and COPD.[10,20,41,42] The dramatic repression of CCDC37 in cancer-free lung tissue from COPD cases, its further repression by methylation in lung cancer and the significant association of its methylation in sputum samples from cancer-free smokers with COPD support its potential as a biomarker for lung cancer risk in COPD patients. Thus, future studies should investigate if adding this novel biomarker to previously identified biomarker panels[10,43] will improve the sensitivity and specificity of predicting lung cancer risk among COPD patients. Functional assays evaluating the mechanism by which epigenetic silencing of these genes contribute to COPD and lung cancer could provide additional insight into the molecular link between these two major public health hazards.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank The Cancer Genome Atlas (TCGA) project. Data generated by TCGA pilot project, which was established by the NCI and NHGRI, was used to validate part of our findings. The dbGaP accession number for TCGA data is phs000178.v8.p7. Information about TCGA and the investigators and institutions that constitute the TCGA research network can be found at http://cancergenome.nih.gov/. We also acknowledge Dr. Eric Edell, pulmonologist at Mayo Clinic, Rochester, MN for providing lung tumor-normal pair samples for this study.

Financial support: This study was supported by R01 CA164782 (Drs. Belinsky and Tesfaigzi) and in part by R21 CA161561 (Dr. Tessema), P30 CA118100 (Dr. Belinsky), R01 CA097356 (Dr. Belinsky), and P50 CA090440 (Dr. Siegfried). The sponsors had no role in the design of the study, collection and analysis of data, or preparation of the manuscript.

References

1. Punturieri A, Szabo E, Croxton TL, et al. Lung cancer and chronic obstructive pulmonary disease: needs and opportunities for integrated research. J Natl Cancer Inst. 2009; 101:554–559. [PubMed: 19351920]
2. Murphy SL, Xu J, Kochanek KD. Deaths: final data for 2010. Natl Vital Stat Rep. 2013; 61:1–117. [PubMed: 24979972]
3. Houghton AM. Mechanistic links between COPD and lung cancer. Nat Rev Cancer. 2013; 13:233–245. [PubMed: 23467302]
4. Nakayama M, Satoh H, Sekizawa K. Risk of cancers in COPD patients. Chest. 2003; 123:1775–1776. [PubMed: 12740307]
5. Papi A, Casoni G, Caramori G, et al. COPD increases the risk of squamous histological subtype in smokers who develop non-small cell lung carcinoma. Thorax. 2004; 59:679–681. [PubMed: 15282388]
6. Belinsky SA. Gene-promoter hypermethylation as a biomarker in lung cancer. Nat Rev Cancer. 2004; 4:707–717. [PubMed: 15343277]
7. Tellez CS, Juri DE, Do K, et al. EMT and stem cell-like properties associated with miR-205 and miR-200 epigenetic silencing are early manifestations during carcinogen-induced transformation of human lung epithelial cells. Cancer Res. 2011; 71:3087–3097. [PubMed: 21363915]
8. Belinsky SA, Nikula KJ, Palmisano WA, et al. Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. Proc Natl Acad Sci U S A. 1998; 95:11891–11896. [PubMed: 9751761]
9. Belinsky SA, Palmisano WA, Gilliland FD, et al. Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. Cancer Res. 2002; 62:2370–2377. [PubMed: 11956099]
10. Sood A, Petersen H, Blanchette CM, et al. Wood smoke exposure and gene promoter methylation are associated with increased risk for COPD in smokers. Am J Respir Crit Care Med. 2010; 182:1098–1104. [PubMed: 20595226]
11. Suzuki M, Wada H, Yoshino M, et al. Molecular characterization of chronic obstructive pulmonary disease-related non-small cell lung cancer through aberrant methylation and alterations of EGFR signaling. Ann Surg Oncol. 2010; 17:878–888. [PubMed: 19841986]
12. Bibikova M, Barnes B, Tsan C, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011; 98:288–295. [PubMed: 21839163]
13. Pauwels RA, Buist AS, Calverley PM, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. Am J Respir Crit Care Med. 2001; 163:1256–1276. [PubMed: 11316667]
14. Hunninghake GM, Cho MH, Tesfaigzi Y, et al. MMP12, lung function, and COPD in high-risk populations. N Engl J Med. 2009; 361:2599–2608. [PubMed: 20018959]
15. Leng S, Liu Y, Thomas CL, et al. Native American ancestry affects the risk for gene methylation in the lungs of Hispanic smokers from New Mexico. Am J Respir Crit Care Med. 2013; 188:1110–1116. [PubMed: 24032348]
16. Leng S, Stidley CA, Liu Y, et al. Genetic determinants for promoter hypermethylation in the lungs of smokers: a candidate gene-based study. Cancer Res. 2012; 72:707–715. [PubMed: 22139380]
17. Leng S, Stidley CA, Willink R, et al. Double-strand break damage and associated DNA repair genes predispose smokers to gene methylation. Cancer Res. 2008; 68:3049–3056. [PubMed: 18413776]
18. Wilson DO, Weissfeld JL, Fuhrman CR, et al. The Pittsburgh Lung Screening Study (PLuSS): outcomes within 3 years of a first computed tomography scan. Am J Respir Crit Care Med. 2008; 178:956–961. [PubMed: 18635890]
19. Ramirez RD, Sheridan S, Girard L, et al. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. Cancer Res. 2004; 64:9027–9034. [PubMed: 15604268]

20. Belinsky SA, Liechty KC, Gentry FD, et al. Promoter hypermethylation of multiple genes in sputum precedes lung cancer incidence in a high-risk cohort. Cancer Res. 2006; 66:3338–3344. [PubMed: 16540689]
21. Tessema M, Willink R, Do K, et al. Promoter methylation of genes in and around the candidate lung cancer susceptibility locus 6q23-25. Cancer Res. 2008; 68:1707–1714. [PubMed: 18339850]
22. Tessema M, Yu YY, Stidley CA, et al. Concomitant promoter methylation of multiple genes in lung adenocarcinomas from current, former and never smokers. Carcinogenesis. 2009; 30:1132–1138. [PubMed: 19435948]
23. Tessema M, Klinge DM, Yingling CM, et al. Re-expression of CXCL14, a common target for epigenetic silencing in lung cancer, induces tumor necrosis. Oncogene. 2010; 29:5159–5170. [PubMed: 20562917]
24. Tessema M, Yingling CM, Liu Y, et al. Genome-wide unmasking of epigenetically silenced genes in lung adenocarcinoma from smokers and never smokers. Carcinogenesis. 2014; 35:1248–1257. [PubMed: 24398667]
25. Tessema M, Yingling CM, Thomas CL, et al. SULF2 methylation is prognostic for lung cancer survival and increases sensitivity to topoisomerase-I inhibitors via induction of ISG15. Oncogene. 2012; 31:4107–4106. [PubMed: 22158045]
26. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001; 25:402–408. [PubMed: 11846609]
27. Siegmund KD. Statistical approaches for the analysis of DNA methylation microarray data. Hum Genet. 2011; 129:585–595. [PubMed: 21519831]
28. Clark SJ, Melki J. DNA methylation and gene silencing in cancer: which is the guilty party? Oncogene. 2002; 21:5380–5387. [PubMed: 12154400]
29. Ferracin M, Gafa R, Miotto E, et al. The methylator phenotype in microsatellite stable colorectal cancers is characterized by a distinct gene expression profile. J Pathol. 2008; 214:594–602. [PubMed: 18241079]
30. Kwon YJ, Lee SJ, Koh JS, et al. Genome-wide analysis of DNA methylation and the gene expression change in lung cancer. J Thorac Oncol. 2012; 7:20–33. [PubMed: 22011669]
31. Petronis A. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. Nature. 2010; 465:721–727. [PubMed: 20535201]
32. Tessema M, Yingling CM, Grimes MJ, et al. Differential epigenetic regulation of TOX subfamily high mobility group box genes in lung and breast cancers. PLoS One. 2012; 7:e34850. [PubMed: 22496870]
33. Yamamoto R, Song K, Yanagisawa HA, et al. The MIA complex is a conserved and novel dynein regulator essential for normal ciliary motility. J Cell Biol. 2013; 201:263–278. [PubMed: 23569216]
34. Cuozzo C, Porcellini A, Angrisano T, et al. DNA damage, homology-directed repair, and DNA methylation. PLoS Genet. 2007; 3:e110. [PubMed: 17616978]
35. Benoist M, Palenzuela R, Rozas C, et al. MAP1B-dependent Rac activation is required for AMPA receptor endocytosis during long-term depression. Embo J. 2013; 32:2287–2299. [PubMed: 23881099]
36. Harrison B, Kraus M, Burch L, et al. DAPK-1 binding to a linear peptide motif in MAP1B stimulates autophagy and membrane blebbing. J Biol Chem. 2008; 283:9999–10014. [PubMed: 18195017]
37. Montenegro-Venegas C, Tortosa E, Rosso S, et al. MAP1B regulates axonal development by modulating Rho-GTPase Rac1 activity. Mol Biol Cell. 2010; 21:3518–3528. [PubMed: 20719958]
38. Langkopf A, Hammarback JA, Muller R, et al. Microtubule-associated proteins 1A and LC2. Two proteins encoded in one messenger. RNA J Biol Chem. 1992; 267:16561–16566. [PubMed: 1379599]
39. Lee SY, Kim JW, Jeong MH, et al. Microtubule-associated protein 1B light chain (MAP1B-LC1) negatively regulates the activity of tumor suppressor p53 in neuroblastoma cells. FEBS Lett. 2008; 582:2826–2832. [PubMed: 18656471]

40. Lerch-Gaggl AF, Sun K, Duncan SA. Light chain 1 of microtubule-associated protein 1B can negatively regulate the action of Pes1. J Biol Chem. 2007; 282:11308–11316. [PubMed: 17308336]

41. Leng S, Do K, Yingling CM, et al. Defining a gene promoter methylation signature in sputum for lung cancer risk assessment. Clin Cancer Res. 2012; 18:3387–3395. [PubMed: 22510351]

42. Leng S, Liu Y, Weissfeld JL, et al. 15q12 Variants, Sputum Gene Promoter Hypermethylation, and Lung Cancer Risk: A GWAS in Smokers. J Natl Cancer Inst. 2015; 107

43. Bruse S, Petersen H, Weissfeld J, et al. Increased methylation of lung cancer-associated genes in sputum DNA of former smokers with chronic mucous hypersecretion. Respir Res. 2014; 15:2. [PubMed: 24405663]

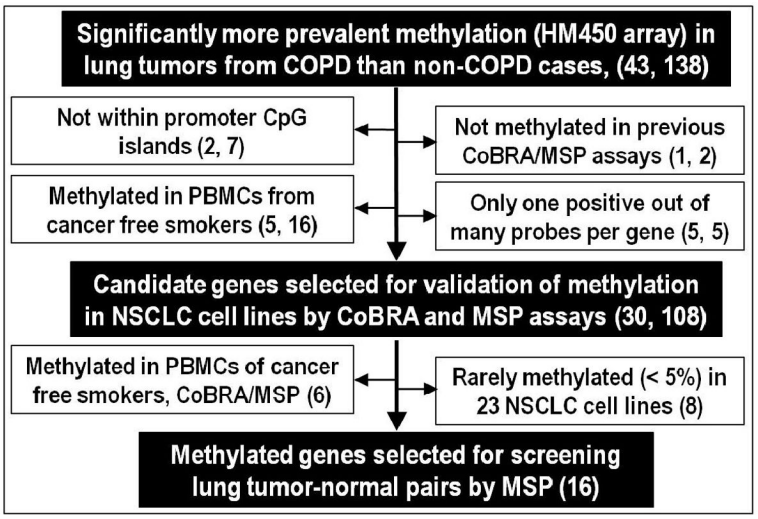

Figure 1. Strategy to identify aberrantly hypermethylated genes linking COPD to lung cancer
(n = genes, probes)

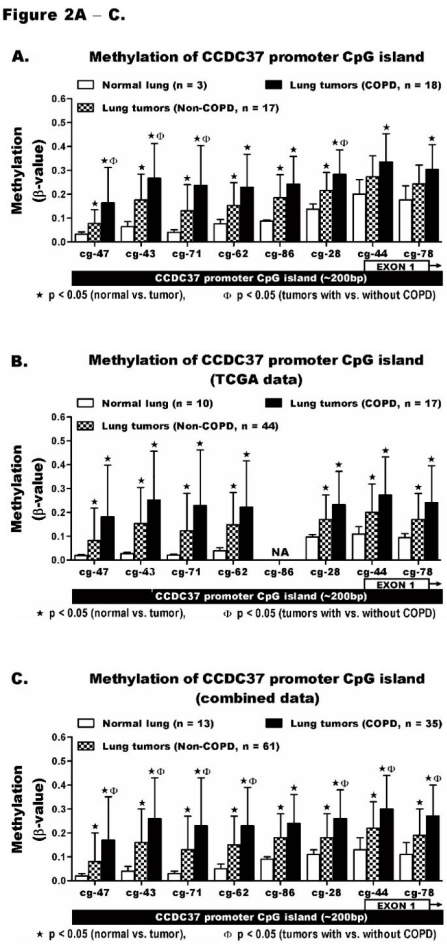

Figure 2. Quantitative methylation analysis of *CCDC37* promoter CpG islands

The HM450K array probes cg23816347, cg09686643, cg01815671, cg11229862, cg20952286, cg20312228, cg21109744, and cg00891278 interrogate the methylation levels across the promoter CpG island of *CCDC37* (~200bp) targeting CpGs within the first exon and its upstream region. The methylation level (β-value) of each probe could range from β = 0 (no methylation) to β = 1 (100% methylation). Methylation was compared between normal lung and lung tumors from COPD and non-COPD cases using **(A)** our data, **(B)** TCGA data, and **(C)** a combination of the two data sets. The results are shown as mean ± standard deviation (SD) of β-values. The methylation levels among normal lung tissues were similarly low while tumors from both COPD and non-COPD cases showed large variation (depicted by the large SD). Details of the methylation analysis including the range and prevalence for methylation of each probe are shown in supplementary Tables S4, S6, and S7. Each probe IDs on the x-axes is shortened by replacing the first 6 numbers following cg by a hyphen (−).

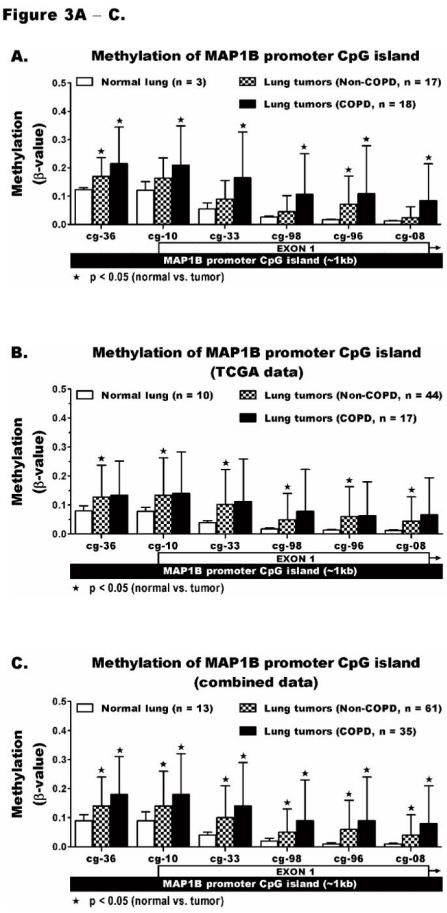

Figure 3. Quantitative methylation analysis of *MAP1B* promoter CpG islands

The methylation levels of HM450K probes cg16478236, cg02001410, cg26001333, cg15412498, cg07380496, and cg02606808 that target CpGs across the promoter CpG island of *MAP1B* were compared between normal lung and lung tumors from COPD and non-COPD cases using **(A)** our data, **(B)** TCGA data, and **(C)** a combination of the two data sets. The analysis was done exactly as described for Figure 2 and the statistical details including the range and prevalence for methylation are shown in supplementary Tables S4, S6, and S7.

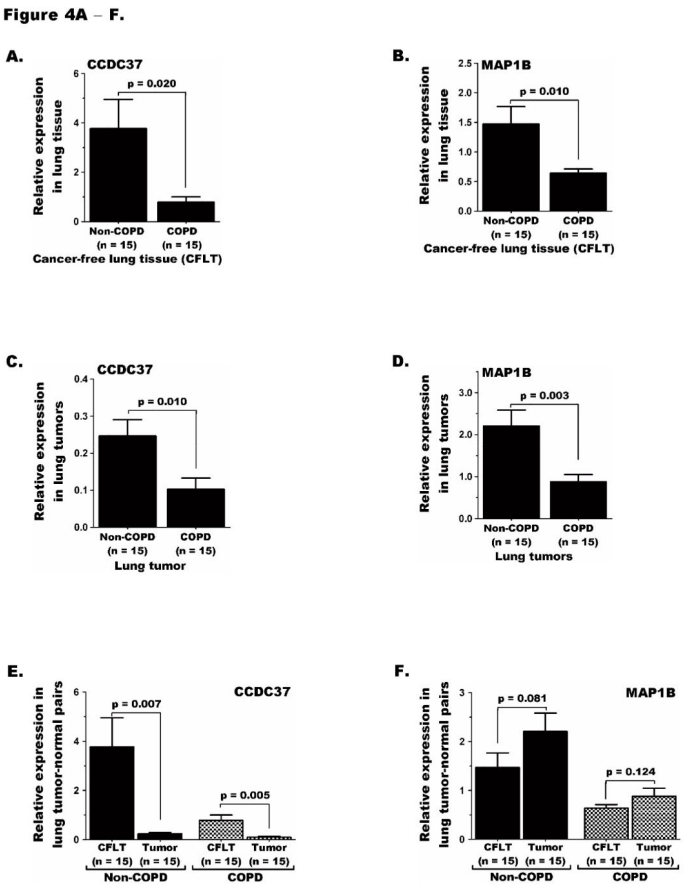

Figure 4. *CCDC37* and *MAP1B* expression in lung tumor-normal pairs

Expression of **(A)** *CCDC37* and **(B)** *MAP1B* in cancer-free lung tissue (CFLT) from COPD cases was significantly repressed than CFLT from non-COPD cases. **C and D)** Similarly, expression of both genes in lung tumors from COPD cases was significantly lower than tumors from non-COPD cases. **E)** While *CCDC37* expression in lung tumors was significantly reduced than CFLT pairs in both COPD and non-COPD groups, (**F)** *MAP1B* expression was not significantly changed between tumors and CFLT pairs in either COPD or non-COPD cases.

Table 1

Characteristics of patients whose lung samples were used in this study.

Characteristics	COPD		Non-COPD	
	Total (n = 71)	Discovery (n = 18)	Total (n = 46)	Discovery (n = 17)
Sex (F)	31 (44%)	8 (44%)	24 (52%)	11 (65%)
Age, median ± SD, (range)	65 ± 8.4 (48 — 80)	65 ± 9.0 (48 — 78)	65 ± 8.0 (49 — 80)	65 ± 9.0 (49 — 78)
Smoking status: Smokers (Current) *	15 (21%)	5 (28%)	7 (15%)	2 (12%)
Lung Cancer				
Stage IA — IB	40 (56%)	12 (67%)	20 (43%)	8 (47%)
Stage IIA — IIB	13 (18%)	2 (11%)	9 (20%)	4 (24%)
Stage IIIA — IIIB	12 (17%)	3 (17%)	13 (28%)	5 (29%)
Stage IV	7 (10%)	1 (6%)	4 (9%)	0 (0%)
Histology				
Adenocarcinoma (AC)	35 (49%)	18 (100%)	32 (70%)	17 (100%)
Squamous cell carcinoma (SCC)	36 (51%)	0 (0%)	14 (30%)	0 (0%)
Cancer-free lung tissue (CFLT)	9	3	8	3
COPD status (based on post-FEV1)				
Non-COPD (FEV1/FVC ≥ 70%)	0 (0%)	0 (0%)	46 (100%)	17 (100%)
COPD severity (FEV1/FVC < 0.70 and				
Gold 1 (FEV1 ≥ 80% predicted); Mild	4 (6%),	1 (6%),	0 (0%)	0 (0%)
Gold 2 (50% ≤ FEV1 < 80%); Moderate	41 (58%)	8 (44%)	0 (0%)	0 (0%)
Gold 3 (30% ≤ FEV1 <50%); Severe	25 (35%)	8 (44%)	0 (0%)	0 (0%)
Gold 4 (FEV1 <30%); Very severe	1 (1%)	1 (6%)	0 (0%)	0 (0%)

* All samples are from smokers.

Table 2

Characteristics of cancer-free smokers whose sputum samples were studied.

Characteristics	LSC cohort, n (%) *		PLuSS cohort, n (%) *	
	COPD (n = 100)	Non-COPD (n = 111)	COPD (n = 102)	Non-COPD (n = 103)
Sex (F)	72 (72%)	79 (71%)	64 (63%)	72 (70%)
Age, median + SD (range)	64 ± 8.3 (41 — 80)	63 ± 8.1 (41 — 76)	65 ± 5.4 (63 — 80)	64 ± 5.2 (52 — 80)
Smoking status: Smokers (Current) **	48 (48%)	56 (50%)	64 (63%)	56 (54%)
COPD status (based on post-*FEV1)*				
Non-COPD *(FEV1/FVC ≥ 70%)*	0 (0%)	111 (100%)	0 (0%)	103 (100%)
COPD: *FEV1/FVC < 0.70 and*				
Gold 1 (FEV1 ≥ 80% predicted)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Gold 2 (50% ≤ FEV1 < 80%)	63 (63%)	0 (0%)	48 (47%)	0 (0%)
Gold 3 (30% ≤ FEV1 <50%)	31 (31%)	0 (0%)	42 (41%)	0 (0%)
Gold 4 (FEV1 <30%)	6 (6%)	0 (0%)	12 (12%)	0 (0%)

* Lovelace Smokers cohort (LSC) from the Albuquerque, NM metropolitan area since 2001, Pittsburgh Lung Screening Study (PLuSS) cohort established in 2002.

** All samples are from smokers

Table 3

Prevalence for methylation of genes in lung tumors and cancer-free lung tissue (CFLT) pairs from NSCLC patients with or without COPD.

Genes	Methylation, n (%)				Fisher Exact test * (p-values)
	COPD		Non-COPD		
	CFLT (n = 22)	Tumor (n = 71)	CFLT (n = 11)	Tumor (n = 46)	
CCDC37 **	4 (18%)	54 (76%)	1 (9%)	20 (43%)	0.0013
CCDC67	1 (5%)	26 (37%)	1 (9%)	10 (22%)	0.1490
CDK5R2	0 (0%)	13 (18%)	0 (0%)	11 (24%)	0.6342
COCH	0 (0%)	14 (20%)	0 (0%)	5 (11%)	0.3064
GLB1L3	7 (32%)	30 (42%)	0 (0%)	12 (26%)	0.1157
MAP1B **	5 (23%)	48 (68%)	2 (18%)	17 (37%)	0.0035
STOX2	8 (36%)	51 (72%)	6 (55%)	29 (63%)	0.5363
WNT9B	2 (9%)	23 (32%)	1 (9%)	9 (20%)	0.1301
ZNF132	1 (5%)	25 (35%)	1 (9%)	10 (22%)	0.0988
ZNF167	3 (14%)	19 (27%)	0 (0%)	11 (24%)	0.6655

* p-values for tumors after adjustment for age, sex, smoking status, and tumor histology.

** Promoter CpG islands with significantly more prevalent methylation in lung tumors from COPD than non-COPD cases.

Table 4

Prevalence for methylation of CCDC37 and *MAP1B* promoter CpG islands in sputum samples from COPD cases and controls

Genes	Cohorts*	Methylation n (%)		Logistic regression †	
		Cases	Controls	Odds ratio (95% CI)	(p-value)
CCDC37	LSC	60/100 (60%)	40/111 (36%)	2.74 (1.55 — 4.90)	0.0006
	PLuSS	86/103 (83%)	66/107 (64%)	2.79 (1.37 — 5.66)	0.0046
	Combined	146/203 (72%)	106/218 (49%)	2.55 (1.68 — 3.87)	< 0.0001
MAP1B	LSR	24/100 (24%)	25/111 (23%)	1.10 (0.58 — 2.08)	0.78

*Lovelace Smokers Cohort (LSC), Pittsburgh Lung Screening Study (PLuSS) Cohort from the Pittsburgh Lung SPORE.

†Adjusted for age, sex, and smoking status.