A low DNA methylation epigenotype in squamous cell lung cancer and its association with idiopathic pulmonary fibrosis and poorer prognosis

(特発性肺線維症に相関する予後不良な低メチル化型肺扁平上皮癌)

千葉大学大学院医学薬学府

先端医学薬学専攻

(主任：　吉野一郎教授)

畑　敦

Table of Contents

Abstract

Rationale: Patients with idiopathic pulmonary fibrosis (IPF) have higher risk of developing lung cancers such as squamous cell carcinoma, from which prognosis is typically poor. While aberrant DNA methylation have been linked to IPF, the subsequent carcinogenesis has not been fully investigated.

Objectives: To survey genome-wide methylation of lung squamous cell carcinomas with IPF.

Methods: Using Infinium HumanMethylation450 BeadChip (Infinium 450K), we surveyed DNA methylation in 20 squamous cell carcinomas with or without IPF and in 13 surrounding non-cancerous lung tissues. Results were validated by pyrosequencing of 77 squamous cell carcinomas.

Measurements and Main Results: Unsupervised hierarchical clustering stratified squamous cell carcinomas into low- (LME) and high-methylation epigenotypes (HME). IPF was more common in the former (7/9, 78 %) than in the latter (1/6, 17 %, $P = 0.04$). Polycomb repressive complex target genes were significantly enriched in hypermethylated genes in squamous cell carcinomas. However, 'negative regulation of growth' was significantly enriched in genes specifically hypermethylated in HME. The association between IPF and LME was validated ($P = 0.003$) by pyrosequencing to measure methylation of markers common to LME and HME (*HOXA2*, *HOXA9*, and *PCDHGB6*), and of markers specific to the latter (*DLEC1*, *CFTR*, *MT1M*, *CRIP3*, and *ALDH7A1*). In addition, survival was found to be poorer among patients with LME than among all patients, and multivariate analysis revealed that LME is an independent predictor of poor prognosis.

Conclusions: Lung squamous cell carcinomas are stratified into two epigenotypes, of which LME is significantly correlated with IPF and worse prognosis.

Introduction

Patients with idiopathic pulmonary fibrosis (IPF) have 14-fold higher risk of developing lung cancer than the general population [1], with the cumulative incidence rate increasing with the duration of follow up [2]. In such patients, lung cancer generally develops near fibrotic lesions [3-5], with squamous cell carcinoma being one of the most frequent histological type [2,4-9]. Due to respiratory deterioration and aggressive cancer behavior, prognosis is typically worse for lung cancer patients with IPF than those without IPF [8,10,11]. Indeed, we previously reported that lung cancer with interstitial lung disease including IPF had poorer recurrence-free survival, whether in cohorts including patients at all pathological stages or in cohorts with stage IA patients only [5].

IPF is a persistent and chronic inflammatory disease that can be triggered by complex interactions between genetic and environmental factors such as cigarette use [12], although cancer development in fibrotic lungs occurs independently of cigarette smoking [1,13]. Recently, exome sequencing revealed APOBEC-related mutagenesis in lung cancers with IPF, suggesting viral and/or immune-related mutagenic processes [14]. Squamous cell metaplasia was also reported to frequently occur in peripheral honeycombed tissues of cancerous lungs with IPF [15]. Further, such metaplasia was linked to *TP53* mutations [16] and loss of heterozygosity at the *FHIT* gene locus [17].

On the other hand, DNA methylation is one of critical epigenetic changes induced by environmental factors, and is believed to promote IPF [18-21] as well as carcinogenesis [19]. Indeed, aberrant hypermethylation of promoters is a major driver of silencing tumor suppressor genes in most types of cancer, including lung cancer [22-25]. For example, Takenaka *et al.* reported frequent hypermethylation of the *SMAD4*

promoter and its reduced expression in lung cancer with IPF, implying loss of growth inhibition via TGF-β [26]. Although epigenetic stratification of lung malignancies and benign diseases was achieved using Infinium technology [22,27,28], DNA methylation in lung squamous cell carcinoma with IPF has not been fully investigated.

Using Infinium HumanMethylation450 BeadChip (Infinium 450K), we here conducted genome-wide DNA methylation analysis in clinical squamous cell carcinoma tissues with or without IPF. These tissues were stratified into high- and low-methylation epigenotypes (HME and LME, respectively), the latter of which is significantly correlated with IPF. The existence of methylation epigenotypes, as well as the relationship to IPF, was validated by pyrosequencing.

Materials and Methods

The study design is illustrated in Figure 1A. Additional details are provided in Supplementary Materials and Methods.

Clinical samples and cell lines

From patients who underwent lung surgery from 2004 to 2014 at Chiba University Hospital, 45 frozen and 32 formalin-fixed paraffin-embedded primary lung squamous cell carcinomas with or without IPF were obtained with informed consent. The study was approved by the institutional review board at Chiba University (No. 806). IPF was diagnosed in 35 of these tissues based on the ATS/ERS/JRS/ALAT Statement [12]. The remaining 42 tissues were free of IPF. All tissues were microscopically confirmed by two independent pathologists to contain > 40 % cancer cells, and were dissected to enrich cancer cells when necessary. Surrounding non-cancerous lung tissues (n = 20)

with or without IPF were used as control, along with normal lung tissues from non-smokers with adenocarcinoma (Figure 1A,B).

DNA was extracted from frozen samples using QIAquick DNA Mini Kit (Qiagen, Hilden, Germany) and from paraffin-embedded samples using QIAamp DNA FFPE Tissue kit (Qiagen). Normal human lung IMR-90 fibroblasts were obtained from Health Science Research Resources Bank. The composition of peripheral blood cells was extracted from The Cancer Genome Atlas. Twenty frozen squamous cell carcinoma samples, 11 surrounding non-cancerous lung tissues with (n = 6) or without (n = 5) IPF, and two control lung samples were initially surveyed on Infinium 450K chips. The full set of 77 squamous cell carcinoma samples and 20 surrounding lung tissues were used for validation by pyrosequencing.

Age, gender, smoking history, tumor location, pathologic TNM stage (p-stage), tumor grade, pleural invasion, and lymphovascular invasion were collected from our cohort (Table 1), as well as from 263 squamous cell carcinoma cases in The Cancer Genome Atlas. Tumors arising at main and segmental bronchi were considered central (Figure 1C), while tumors arising at subsegmental bronchi or at more peripheral locations were considered peripheral (Figure 1D), as previously defined [29-31].

Analysis of DNA methylation

Infinium 450K BeadChip (Illumina) contains approximately 485,000 CpG sites covering 99 % of RefSeq genes at an average of 17 sites per gene. Each CpG site was scored in β-values ranging from 0.00 to 1.00 based on signals from a methylated probe relative to the sum of signals from methylated and unmethylated probes. Bisulfite conversion was performed with Zymo EZ DNA Methylation Kit (Zymo Research, Irvine, CA), using 500 ng of genomic DNA per sample. Whole-genome amplification,

labeling, hybridization, and scanning were performed according to the manufacturer's protocols. Infinium data were submitted to the Gene Expression Omnibus under accession number GSE121849 (GSM3447735-GSM3447767).

To analyze DNA methylation, the probe nearest the transcription start site was selected if multiple probes were available against the same promoter. The CpG score for each probe was calculated as described previously [32,33], and only high- and intermediate-scoring probes were included in subsequent analysis. Genes in X and Y chromosomes were excluded to avoid gender differences. β-values > 0.2 were considered indicative of hypermethylation. Infinium 450K BeadChip (Illumina) data from paraffin-embedded samples were integrated into similar data from frozen samples, as described in detail in Supplementary Materials and Methods. DNA methylation data for 263 squamous carcinoma samples were also obtained from The Cancer Genome Atlas.

Pyrosequencing

Methylation status of putative marker genes was quantified by pyrosequencing in PyroMark Q96 (Qiagen), as previously described [32]. Methylation > 20 % was considered hypermethylation. Primers were designed in Pyro Q-CpG Software (Qiagen, Supplementary Table S1).

Immunohistochemistry

Protein expression of TP53 and $p16^{INK4A}$ was analyzed by immunohistochemistry in a BenchMark ULTRA automated staining system (Roche, Basel, Switzerland), using mouse monoclonal anti-TP53 (clone DO-7) and anti-p16 (CINtec p16 Histology) from Roche. Samples with strong and uniform nuclear staining for TP53 were considered to harbor activating mutations, while those without nuclear staining were considered to

harbor null mutations. Sporadic nuclear staining in surrounding non-cancerous cells, such as inflammatory cells, were used as wild-type internal controls. Finally, samples with nuclear and cytoplasmic immunoreactivity for p16 in > 70 % of tumor cells were considered to express p16.

Gene Ontology analysis

Enrichment of Gene Ontology terms for biologic process was analyzed using Functional Annotation at DAVID Bioinformatics Resources (http://david.abcc.ncifcrf.gov/).

Statistical analysis

Unsupervised two-way hierarchical clustering was conducted in Cluster 3.0 (http://bonsai.hgc.jp/~mdehoon/software/cluster/) or in R (www.r-project.org/). Heatmaps were drawn in Java TreeView (http://jtreeview.sourceforge.net/). Association between DNA methylation and clinicopathological factors, including presence or absence of IPF, was analyzed by Fisher's exact test, χ^2 test, or Kolmogorov-Smirnov test for categorical variables, and by Mann-Whitney test for continuous data. Cancer-specific overall survival was measured from enrollment date until date of cancer-related death. Patients who died from causes unrelated to cancer were censored. Survival curves were constructed by the Kaplan-Meier method, and *P* values were calculated by log-rank test. Multivariate analysis was performed using Cox proportional hazards. Analyses were performed in JMP pro 12.0.1 (SAS Institute Inc., Cary, NC), with $P < 0.05$ considered statistically significant.

Results

Patient characteristics

Age, tumor location, and p16 expression were significantly different between squamous cell carcinoma patients with or without IPF. In contrast, these patients were comparable in terms of gender, smoking status, tumor stage, pleural and lymphovascular invasion, and tumor grade. Most squamous cell carcinomas with IPF (97.1 %) were observed in the periphery of the lung parenchyma (P = 0.04), and mostly near fibrotic areas (75.8 %). While activating TP53 mutations were frequently observed in both groups, with incidence 82.9 % in carcinomas with IPF and 88.1 % in carcinomas without IPF (P = 0.8), p16 expression was significantly more frequent in the former (35.3 %) than in the latter (14.0 %, P = 0.03).

Analysis of DNA methylation

The standard deviation of β-values was calculated for probes nearest the transcription start sites at each promoter in 20/77 squamous cell carcinomas, 11/20 surrounding lung tissues, and 2/2 normal tissues. Subsequently, 570 probes with standard deviation > 0.13 were extracted, and analyzed by unsupervised two-way hierarchical clustering (Figure 2A). This analysis revealed distinct low- (LME) and high-methylation epigenotypes (HME), which were found in nine and six squamous cell carcinomas, respectively. Remarkably, IPF was significantly associated with the former (7/9, 78 %), but not with the latter (1/6, 17 %) (P = 0.04 by Fisher's exact test). This trend was also observed in another 24 formalin-fixed paraffin-embedded samples, as well as in 390 squamous cell carcinomas in The Cancer Genome Atlas (Supplementary Materials and Methods, and Supplementary Figures E1 and E2). Accordingly, aberrant hypermethylation was less frequently observed in squamous cell carcinomas with IPF (532 genes) than in those without IPF (698 genes, Figure 2B). However, aberrant hypermethylation was more frequently observed in surrounding non-cancerous lung tissues with IPF (192 genes)

than in those without IPF (98 genes, Figure 2B).

Extraction of marker genes

We found 464 genes that are aberrantly hypermethylated in both epigenotypes (common markers), and 287 genes hypermethylated specifically in HME (Figure 3A). Genes targeted by polycomb repressive complex in embryonic stem cells [34] were significantly enriched in both groups of markers (Figure 3B), along with genes related to transcription (*HOXA2*, *HOXA9*, etc.) or cell adhesion (*PCDHGB2*, *PCDHGB6*, etc.) (Figure 3C). In contrast, genes that suppress growth (*MT1M*, *MT1G*, etc.) are significantly enriched in HME markers (Figure 3D).

Validation of DNA methylation by pyrosequencing

Five of 287 HME markers (*DLEC1*, *CFTR*, *CRIP3*, *ALDH7A1*, and *MT1M*) and three of 464 common markers (*HOXA2*, *HOXA9*, and *PCDHGB6*) were validated by pyrosequencing. Methylation levels at several CpG sites around the Infinium probe were quantified precisely, and methylation profiles were confirmed to strongly correlate to those obtained by Infinium assay (Figure 4A). In additional squamous cell carcinoma specimens and surrounding lung tissues (Figure 4B), carcinomas without IPF were significantly associated with frequent hypermethylation, in comparison to those with IPF ($P = 0.01$ by Kolmogorov-Smirnov test). On the other hand, methylation levels of the HME markers *DLEC1*, *CRIP3*, and *ALDH7A1* were significantly lower in carcinomas with IPF than in carcinomas without IPF ($P = 0.03$, 0.049, 0.02, respectively, Figure 4C).

Correlation of IPF, clinicopathological factors, and DNA methylation status

Squamous carcinoma samples hypermethylated (methylation level > 20 %) in four or all five HME markers were classified as HME (Figure 5A). Accordingly, 23 of the 77

carcinomas in our cohort were classified as HME, and the remaining 54 were classified as LME. Analysis of clinicopathological factors revealed that squamous cell carcinomas with IPF are significantly associated with LME (P = 0.003), although other clinicopathological factors did not correlate with methylation status (Figure 5B and Table 2).

Analysis of survival

Analysis of cancer-related deaths indicated that LME is significantly associated with worse prognosis (P = 0.01, Figure 5C). This association was also observed in stage I and II patients (P = 0.02, Figure 5D), as well as in 363 cases in The Cancer Genome Atlas (Figure 5E). Notably, prognosis is significantly worse in LME tumors without IPF than in HME tumors with IPF, although LME is significantly associated with IPF (P = 0.04 for all patients and P = 0.05 for stage I and II patients, Figure 5F,G). Unadjusted and adjusted hazards ratios for various prognostic factors are listed in Table 3. IPF, epigenotype (LME), pathologic stage, N stage, and lymphovascular invasion were found to be significant prognostic factors in univariate analysis. However, multivariate analysis identified LME and tumor stage as independent predictors of poor prognosis, at least in our cohort.

Discussion

While aberrant DNA methylation has been linked to IPF [21], its possible association with subsequent squamous lung carcinogenesis has not been investigated. Using Infinium technology to survey genome-wide methylation of promoters in lung squamous cell carcinomas with or without IPF, we found that these carcinomas can be

classified as HME or LME, that carcinomas with IPF are significantly associated with the latter, and that LME is significantly associated with poor prognosis. A novel panel of five HME markers was also established.

Shi *et al.* identified more than 400 differentially methylated genes with altered expression in lung squamous cell carcinomas, many of which are related to development and cell adhesion, and some were proposed as potential diagnostic markers [35]. In addition, analysis of genome-wide methylation data in The Cancer Genome Atlas revealed several epigenotypes in squamous cell carcinomas [22], although the prognostic value of each, as well as their relationship to IPF, was not investigated. Indeed, only *SMAD4* was previously reported as an aberrantly hypermethylated gene in lung cancers with IPF [26]. We have now identified 464 genes that are aberrantly methylated in squamous cell carcinomas with or without IPF. Gene Ontology terms related to transcription and cell adhesion are enriched in these genes, in agreement with previous studies [24,35].

Aberrant promoter methylation in cancer is observed more frequently in genes targeted by the polycomb repressor complex in embryonic stem cells than in all other genes, suggesting that cancers arise from stem cells, via a mechanism in which reversible gene repression is replaced by permanent gene silencing [36]. Strikingly, genes targeted by both forms of polycomb repressor complex, *i.e.*, PRC1 and PRC2 [34], are significantly enriched in HME-specific markers and in markers common to both HME and LME, although PRC1 target genes are further enriched in common markers. Common markers also include genes related to transcription and cell adhesion, while HME markers include genes related to tumor suppression. Indeed, hypermethylation in HME may silence tumor suppressors and thereby trigger carcinogenesis.

Since most patients in our cohort are smokers (94.8 %), with mean Brinkman Index above 40 pack-years, cigarette smoking should have had substantial impact in our study due to its strong association with both squamous cell carcinoma and IPF. Recently, Vaz *et al.* [37] reported that smoking elicits epigenetic alterations in normal human bronchial epithelial cells exposed to cigarette smoke from 10 days to 15 months. Binding to chromatin by DNMT1 and EZH2, a component of PRC2, was initially observed, followed by aberrant methylation of promoters driving homeobox-related genes and transcription factors, highlighting a pathway from PRC2 to DNA methylation. These genes are similar to markers that are common to both HME and LME, suggesting that methylation of these genes might be linked to smoking.

DLEC1, *CFTR*, *MT1M*, *CRIP3*, and *ALDH7A1* were established as HME-specific markers, of which *DLEC1*, *CFTR*, and *MT1M* were previously reported as tumor suppressors [38-40]. Hypermethylation of *CRIP3* and *ALDH7A1* was not described previously for lung cancer, while hypermethylation of *CRIP3* has been reported in prostate cancers [41], although the physiological significance of this epigenetic lesion is unknown. *ALDH7A1* encodes an aldehyde dehydrogenase that promotes bone metastasis in prostate cancer cell lines [42]. On the other hand, hypermethylation of *HOXA2*, *HOXA9*, and *PCDHGB6*, all of which are markers common to HME and LME, was previously reported in non-small cell lung cancer [24,43,44]. *PCDHGB6* encodes a cadherin-type protein that facilitates cell adhesion and formation of cell-cell junctions [45].

Survival analysis in our cohort and in similar cases in The Cancer Genome Atlas revealed that LME is significantly associated with worse prognosis than HME. Better prognosis in higher methylation subtypes was also observed in other tumors such as

colon cancer, breast cancer, and glioma [46-49]. Previously, genomic and epigenomic surveys of lung squamous cell cancers in The Cancer Genome Atlas demonstrated that these tumors can be classified into four epigenotypes [22], which, in turn, are associated with transcriptional subtypes [50]. Interestingly, the epigenotype corresponding to the highest levels of DNA methylation is predominantly associated with the 'classical' type, the most typical form of lung squamous cell cancer. Other phenotypes, including the most aggressive 'primitive' type, were found to have the most normal-like epigenotype, with the lowest levels of hypermethylation. These are consistent with the observed relationship between LME and poor prognosis, although further analysis is necessary to confirm the association between methylation epigenotypes and transcription subtypes.

Although there have only been a few surveys of lung cancers with IPF [15,16,51,52], these surveys found that most are generated in peripheral tissues, and typically in the lower lobe [4,9]. In addition, 65 % of cases are associated with peripheral honeycombed areas or in the border between honeycombed and non-fibrotic areas [15]. The cancers in our cohort are consistent with these characteristics. On the other hand, Kawasaki *et al.* observed frequent *TP53* mutations in squamous metaplasia in fibrotic peripheral tissues [16]. In our cohort, over 80 % of carcinomas with IPF also harbored activating *TP53* mutations, as assessed by immunohistochemistry, implying that these mutations are crucial to the development of carcinomas with IPF, although there was no significant difference in incidence of *TP53* mutations compared to carcinomas without IPF. Recently, Hwang *et al.* found that BRAF mutations are more frequently observed (17.1 %) in patients with IPF-associated lung cancer than in the general population, and found APOBEC3-related mutational signatures [14]. Of note, APOBEC3 is known to be induced by exogenous factors such as viral infection, while

Folcik *et al.* reported that lungs with IPF are frequently infected by herpesvirus saimiri [53]. Similarly, we observed that squamous cell carcinomas with IPF frequently express p16, a marker of infection by human papilloma virus, which was reported to inactivate RB1 in cervical squamous cell carcinoma [54]. Thus, squamous cell carcinomas with IPF may be linked to viruses that inactivate RB1, although this hypothesis requires validation.

We note that we did not measure gene expression, although aberrant hypermethylation of promoters should repress genes in theory. In addition, analysis of surrounding lung tissues could have been affected by stromal cells, so microdissection of pre-cancerous lesions such as squamous metaplasia may be needed to confirm results. Third, patients with advanced cancer and who were provided non-surgical care were excluded. Fourth, features other than LME, *e.g.*, genetic alterations, may be required to generate poor prognosis. For example, Shinjo *et al*. described the correlation of LME and mutations in epidermal growth factor receptor in adenocarcinomas [55]. Hence, further investigation of potential genetic alterations should be performed for LME squamous cell lung carcinomas. Nonetheless, we anticipate that genome-wide epigenomic profiles of squamous cell carcinomas with IPF, as assessed for the first time by Infinium 450K BeadChip technology, will prove clinically useful.

In summary, genome-wide methylation analysis of squamous cell carcinomas revealed for the first time that a low-methylation epigenotype is associated with squamous cell carcinomas with IPF. Moreover, genes targeted by polycomb repressor complex are methylated in squamous cell carcinomas with IPF, probably due to smoking, although additional events may ultimately be required to trigger carcinogenesis.

Acknowledgments

We thank Eriko Ikeda and Keiko Iida for technical assistance, and Editage (www.editage.jp) for English language editing.

Reference

1. Hubbard R, Venn A, Lewis S, Britton J. Lung cancer and cryptogenic fibrosing alveolitis. A population-based cohort study. Am J Respir Crit Care Med 2000;161:5-8.
2. Ozawa Y, Suda T, Naito T, Enomoto N, Hashimoto D, Fujisawa T, Nakamura Y, Inui N, Nakamura H, Chida K. Cumulative incidence of and predictive factors for lung cancer in IPF. Respirology 2009;14:723-728.
3. Kishi K, Homma S, Kurosaki A, Motoi N, Yoshimura K. High-resolution computed tomography findings of lung cancer associated with idiopathic pulmonary fibrosis. J Comput Assist Tomogr 2006;30:95-99.
4. Lee HJ, Im JG, Ahn JM, Yeon KM. Lung cancer in patients with idiopathic pulmonary fibrosis: CT findings. J Comput Assist Tomogr 1996;20:979-982.
5. Hata A, Suzuki H, Nakajima T, Tanaka K, Fujiwara T, Wada H, Iwata T, Yoshida S, Yoshino I. Concomitant interstitial lung disease is a risk factor for pleural invasion in lung cancer. Ann Thorac Surg 2017;103:967-974.
6. Kawasaki H, Nagai K, Yokose T, Yoshida J, Nishimura M, Takahashi K, Suzuki K, Kakinuma R, Nishiwaki Y. Clinicopathological characteristics of surgically resected lung cancer associated with idiopathic pulmonary fibrosis. J Surg Oncol 2001;76:53-57.
7. Omori T, Tajiri M, Baba T, Ogura T, Iwasawa T, Okudela K, Takemura T, Oba MS, Maehara T, Nakayama H, Tsuboi M, Masuda M. Pulmonary resection for lung cancer in patients with idiopathic interstitial pneumonia. Ann Thorac Surg 2015;100:954-960.
8. Tomassetti S, Gurioli C, Ryu JH, Decker PA, Ravaglia C, Tantalocco P, Buccioli

M, Piciucchi S, Sverzellati N, Dubini A, Gavelli G, Chilosi M, Poletti V. The impact of lung cancer on survival of idiopathic pulmonary fibrosis. Chest 2015;147:157-164.

9. Lee T, Park JY, Lee HY, Cho YJ, Yoon HI, Lee JH, Jheon S, Lee CT, Park JS. Lung cancer in patients with idiopathic pulmonary fibrosis: clinical characteristics and impact on survival. Respir Med 2014;108:1549-1555.
10. Watanabe A, Higami T, Ohori S, Koyanagi T, Nakashima S, Mawatari T. Is lung cancer resection indicated in patients with idiopathic pulmonary fibrosis? J Thorac Cardiovasc Surg 2008;136:1357-1363.
11. Sato T, Watanabe A, Kondo H, Kanzaki M, Okubo K, Yokoi K, Matsumoto K, Marutsuka T, Shinohara H, Teramukai S, Kishi K, Ebina M, Sugiyama Y, Meinoshin O, Date H, Japanese Association for Chest Surgery. Long-term results and predictors of survival after surgical resection of patients with lung cancer and interstitial lung diseases. J Thorac Cardiovasc Surg 2015;149:64-69.
12. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, Lynch DA, Ryu JH, Swigris JJ, Wells AU, Ancochea J, Bouros D, Carvalho C, Costabel U, Ebina M, Hansell DM, Johkoh T, Kim DS, King TE Jr, Kondoh Y, Myers J, Müller NL, Nicholson AG, Richeldi L, Selman M, Dudden RF, Griss BS, Protzko SL, Schünemann HJ, ATS/ERS/JRS/ALAT Committee on Idiopathic Pulmonary Fibrosis. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Respir Crit Care Med 2011;183:788-824.
13. Samet JM. Does idiopathic pulmonary fibrosis increase lung cancer risk? Am J

Respir Crit Care Med 2000;161:1-2.

14. Hwang JA, Kim D, Chun SM, Bae S, Song JS, Kim MY, Koo HJ, Song JW, Kim WS, Lee JC, Kim HR, Choi CM, Jang SJ. Genomic profiles of lung cancer associated with idiopathic pulmonary fibrosis. J Pathol 2018;244:25-35.
15. Hironaka M, Fukayama M. Pulmonary fibrosis and lung carcinoma: a comparative study of metaplastic epithelia in honeycombed areas of usual interstitial pneumonia with or without lung carcinoma. Pathol Int 1999;49:1060-1066.
16. Kawasaki H, Ogura T, Yokose T, Nagai K, Nishiwaki Y, Esumi H. p53 gene alteration in atypical epithelial lesions and carcinoma in patients with idiopathic pulmonary fibrosis. Hum Pathol 2001;32:1043-1049.
17. Sozzi G, Sard L, De Gregorio L, Marchetti A, Musso K, Buttitta F, Tornielli S, Pellegrini S, Veronese ML, Manenti G, Incarbone M, Chella A, Angeletti CA, Pastorino U, Huebner K, Bevilaqua G, Pilotti S, Croce CM, Pierotti MA. Association between cigarette smoking and FHIT gene alterations in lung cancer. Cancer Res 1997;57:2121-2123.
18. Sanders YY, Ambalavanan N, Halloran B, Zhang X, Liu H, Crossman DK, Bray M, Zhang K, Thannickal VJ, Hagood JS. Altered DNA methylation profile in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2012;186:525-535.
19. Weisenberger DJ. Characterizing DNA methylation alterations from The Cancer Genome Atlas. J Clin Invest 2014;124:17-23.
20. Yang IV. Epigenomics of idiopathic pulmonary fibrosis. Epigenomics 2012;4:195-203.
21. Yang IV, Pedersen BS, Rabinovich E, Hennessy CE, Davidson EJ, Murphy E, Guardela BJ, Tedrow JR, Zhang Y, Singh MK, Correll M, Schwarz MI, Geraci M,

Sciurba FC, Quackenbush J, Spira A, Kaminski N, Schwartz DA. Relationship of DNA methylation and gene expression in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2014;190:1263-1272.

22. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature 2012;489:519-525.
23. Carvalho RH, Hou J, Haberle V, Aerts J, Grosveld F, Lenhard B, Philipsen S. Genomewide DNA methylation analysis identifies novel methylated genes in non-small-cell lung carcinomas. J Thorac Oncol 2013;8:562-573.
24. Heller G, Babinsky VN, Ziegler B, Weinzierl M, Noll C, Altenberger C, Müllauer L, Dekan G, Grin Y, Lang G, End-Pfützenreuter A, Steiner I, Zehetmayer S, Döme B, Arns BM, Fong KM, Wright CM, Yang IA, Klepetko W, Posch M, Zielinski CC, Zöchbauer-Müller S. Genome-wide CpG island methylation analyses in non-small cell lung cancer patients. Carcinogenesis 2013;34:513-521.
25. Kenmotsu H, Serizawa M, Koh Y, Isaka M, Takahashi T, Taira T, Ono A, Maniwa T, Takahashi S, Mori K, Endo M, Abe M, Hayashi I, Nakajima T, Ohde Y, Yamamoto N. Prospective genetic profiling of squamous cell lung cancer and adenosquamous carcinoma in Japanese patients by multitarget assays. BMC Cancer 2014;14:786.
26. Takenaka K, Gemma A, Yoshimura A, Hosoya Y, Nara M, Hosomi Y, Okano T, Kunugi S, Koizumi K, Fukuda Y, Uematsu K, Shimizu K, Kudoh S. Reduced transcription of the Smad4 gene during pulmonary carcinogenesis in idiopathic pulmonary fibrosis. Mol Med Rep 2009;2:73-80.
27. Sato T, Arai E, Kohno T, Takahashi Y, Miyata S, Tsuta K, Watanabe S, Soejima K, Betsuyaku T, Kanai Y. Epigenetic clustering of lung adenocarcinomas based on

DNA methylation profiles in adjacent lung tissue: its correlation with smoking history and chronic obstructive pulmonary disease. Int J Cancer 2014;135:319-334.

28. Sundar IK, Yin Q, Baier BS, Yan L, Mazur W, Li D, Susiarjo M, Rahman I. DNA methylation profiling in peripheral lung tissues of smokers and patients with COPD. Clin Epigenetics 2017;9:38.
29. Funai K, Yokose T, Ishii G, Araki K, Yoshida J, Nishimura M, Nagai K, Nishiwaki Y, Ochiai A. Clinicopathologic characteristics of peripheral squamous cell carcinoma of the lung. Am J Surg Pathol 2003;27:978-984.
30. Sakurai H, Asamura H, Watanabe S, Suzuki K, Tsuchiya R. Clinicopathologic features of peripheral squamous cell carcinoma of the lung. Ann Thorac Surg 2004;78:222-227.
31. Shimmyo T, Hashimoto T, Kobayashi Y, Miyagi Y, Ishikawa Y, Nakagawa K, Osada H, Tsuchiya E. p53 mutation spectra for squamous cell carcinomas at different levels of human bronchial branches. Int J Cancer 2006;119:501-507.
32. Matsusaka K, Kaneda A, Nagae G, Ushiku T, Kikuchi Y, Hino R, Uozaki H, Seto Y, Takada K, Aburatani H, Fukayama M. Classification of Epstein-Barr virus-positive gastric cancers by definition of DNA methylation epigenotypes. Cancer Res 2011;71:7187-7197.
33. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M, Schübeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007;39:457-466.
34. Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, Adli M, Kasif S, Ptaszek LM, Cowan CA, Lander

ES, Koseki H, Bernstein BE. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet 2008;4:e1000242.

35. Shi YX, Wang Y, Li X, Zhang W, Zhou HH, Yin JY, Liu ZQ. Genome-wide DNA methylation profiling reveals novel epigenetic signatures in squamous cell lung cancer. BMC Genomics 2017;18:901.
36. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, Laird PW. Epigenetic stem cell signature in cancer. Nat Genet 2007;39:157-158.
37. Vaz M, Hwang SY, Kagiampakis I, Phallen J, Patil A, O'Hagan HM, Murphy L, Zahnow CA, Gabrielson E, Velculescu VE, Easwaran HP, Baylin SB. Chronic cigarette smoke-induced epigenomic changes precede sensitization of bronchial epithelial cells to single-step transformation by KRAS mutations. Cancer Cell 2017;32:360-376.
38. Daigo Y, Nishiwaki T, Kawasoe T, Tamari M, Tsuchiya E, Nakamura Y. Molecular cloning of a candidate tumor suppressor gene, DLC1, from chromosome 3p21.3. Cancer Res 1999;59:1966-1972.
39. Than BL, Linnekamp JF, Starr TK, Largaespada DA, Rod A, Zhang Y, Bruner V, Abrahante J, Schumann A, Luczak T, Walter, Niemczyk A, O'Sullivan MG, Medema JP, Fijneman RJ, Meijer GA, Van den Broek E, Hodges CA, Scott PM, Vermeulen L, Cormier RT. CFTR is a tumor suppressor gene in murine and human intestinal cancer. Oncogene 2016;35:4179-4187.
40. Mao J, Yu H, Wang C, Sun L, Jiang W, Zhang P, Xiao Q, Han D, Saiyin H, Zhu J, Chen T, Roberts LR, Huang H, Yu L. Metallothionein MT1M is a tumor suppressor of human hepatocellular carcinomas. Carcinogenesis

2012;33:2568-2577.

41. Zhao F, Olkhov-Mitsel E, van der Kwast T, Sykes J, Zdravic D, Venkateswaran V, Zlotta AR, Loblaw A, Fleshner NE, Klotz L, Vesprini D, Bapat B. Urinary DNA methylation biomarkers for noninvasive prediction of aggressive disease in patients with prostate cancer on active surveillance. J Urol 2017;197:335-341.
42. van den Hoogen C, van der Horst G, Cheung H, Buijs JT, Pelger RC, van der Pluijm G. The aldehyde dehydrogenase enzyme 7A1 is functionally involved in prostate cancer bone metastasis. Clin Exp Metastasis 2011;28:615-625.
43. Hwang JA, Lee BB, Kim Y, Hong SH, Kim YH, Han J, Shim YM, Yoon CY, Lee YS, Kim DH. HOXA9 inhibits migration of lung cancer cells and its hypermethylation is associated with recurrence in non-small cell lung cancer. Mol Carcinog 2015;54 Suppl 1:E72-E80.
44. Sandoval J, Mendez-Gonzalez J, Nadal E, Chen G, Carmona FJ, Sayols S, Moran S, Heyn H, Vizoso M, Gomez A, Sanchez-Cespedes M, Assenov Y, Müller F, Bock C, Taron M, Mora J, Muscarella LA, Liloglou T, Davies M, Pollan M, Pajares MJ, Torre W, Montuenga LM, Brambilla E, Field JK, Roz L, Lo Iacono M, Scagliotti GV, Rosell R, Beer DG, Esteller M. A prognostic DNA methylation signature for stage I non-small-cell lung cancer. J Clin Oncol 2013;31:4140-4147.
45. Powrózek T, Krawczyk P, Kuźnar-Kamińska B, Batura-Gabryel H, Milanowski J. Analysis of RTEL1 and PCDHGB6 promoter methylation in circulating-free DNA of lung cancer patients using liquid biopsy: a pilot study. Exp Lung Res 2016;42:307-313.
46. Fang F, Turcan S, Rimner A, Kaufman A, Giri D, Morris LG, Shen R, Seshan V, Mo Q, Heguy A, Baylin SB, Ahuja N, Viale A, Massague J, Norton L, Vahdat LT,

Moynahan ME, Chan TA. Breast cancer methylomes establish an epigenomic foundation for metastasis. Sci Transl Med 2011;3:75ra25.

47. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, Pan F, Pelloski CE, Sulman EP, Bhat KP, Verhaak RG, Hoadley KA, Hayes DN, Perou CM, Schmidt HK, Ding L, Wilson RK, Van Den Berg D, Shen H, Bengtsson H, Neuvial P, Cope LM, Buckley J, Herman JG, Baylin SB, Laird PW, Aldape K, Cancer Genome Atlas Research Network. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 2010;17:510-522.
48. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, Barker M, Leggett B, Levine J, Kim M, French AJ, Thibodeau SN, Jass J, Haile R, Laird PW. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787-793.
49. Yagi K, Akagi K, Hayashi H, Nagae G, Tsuji S, Isagawa T, Midorikawa Y, Nishimura Y, Sakamoto H, Seto Y, Aburatani H, Kaneda A. Three DNA methylation epigenotypes in human colorectal cancer. Clin Cancer Res 2010;16:21-33.
50. Wilkerson MD, Yin X, Hoadley KA, Liu Y, Hayward MC, Cabanski CR, Muldrew K, Miller CR, Randell SH, Socinski MA, Parsons AM, Funkhouser WK, Lee CB, Roberts PJ, Thorne L, Bernard PS, Perou CM, Hayes DN. Lung squamous cell carcinoma mRNA expression subtypes are reproducible, clinically important, and correspond to normal cell types. Clin Cancer Res 2010;16:4864-4875.

51. Takahashi T, Munakata M, Ohtsuka Y, Nisihara H, Nasuhara Y, Kamachi-Satoh A, Dosaka-Akita H, Homma Y, Kawakami Y. Expression and alteration of ras and p53 proteins in patients with lung carcinoma accompanied by idiopathic pulmonary fibrosis. Cancer 2002;95:624-633.

52. Uematsu K, Yoshimura A, Gemma A, Mochimaru H, Hosoya Y, Kunugi S, Matsuda K, Seike M, Kurimoto F, Takenaka K, Koizumi K, Fukuda Y, Tanaka S, Chin K, Jablons DM, Kudoh S. Aberrations in the fragile histidine triad (FHIT) gene in idiopathic pulmonary fibrosis. Cancer Res 2001;61:8527-8533.

53. Folcik VA, Garofalo M, Coleman J, Donegan JJ, Rabbani E, Suster S, Nuovo A, Magro CM, Di Leva G, Nuovo GJ. Idiopathic pulmonary fibrosis is strongly associated with productive infection by herpesvirus saimiri. Mod Pathol 2014;27:851-862.

54. Nakagawa T, Matsusaka K, Misawa K, Ota S, Takane K, Fukuyo M, Rahmutulla B, Shinohara KI, Kunii N, Sakurai D, Hanazawa T, Matsubara H, Nakatani Y, Okamoto Y, Kaneda A. Frequent promoter hypermethylation associated with human papillomavirus infection in pharyngeal cancer. Cancer Lett 2017;407:21-31.

55. Shinjo K, Okamoto Y, An B, Yokoyama T, Takeuchi I, Fujii M, Osada H, Usami N, Hasegawa Y, Ito H, Hida T, Fujimoto N, Kishimoto T, Sekido Y, Kondo Y. Integrated analysis of genetic and epigenetic alterations reveals CpG island methylator phenotype associated with distinct clinical characters of lung adenocarcinoma. Carcinogenesis 2012;33:1277-1285.

Table 1. Clinicopathological characteristics of squamous cell carcinomas with or without IPF.

Variable	with IPF (n = 35)	without IPF (n = 42)	*P* value
Age (year ± SD)	72.2 ± 6.4	68.9 ± 7.0	0.03*
Gender			
Male	31 (91.2 %)	38 (88.4 %)	1
Brinkman Index (pack-years)	49.8 ± 24.2	43.0 ± 26.8	0.3
Pathological stage			0.3
IA	8 (23.5 %)	10 (23.3 %)	
IB	11 (32.6 %)	9 (20.9 %)	
IIA	2 (5.9 %)	6 (14.0 %)	
IIB	6 (17.7 %)	10 (23.3 %)	
IIIA	5 (14.7 %)	6 (14.0 %)	
IIIB	0 (0 %)	2 (4.7 %)	
IV	2 (5.9 %)	0 (0 %)	
T stage			0.2
T1	11 (32.6 %)	13 (30.2 %)	
T2	20 (58.8 %)	18 (41.9 %)	
T3	3 (8.8 %)	11 (25.6 %)	
T4	0 (0 %)	1 (2.3 %)	
N stage			0.8
N0	24 (70.6 %)	33 (76.7 %)	
N1	4 (11.8 %)	5 (11.6 %)	
N2	6 (17.7 %)	5 (11.6 %)	
M stage			0.2
M1a	2 (5.9 %)	0 (0 %)	
Pathological pleural invasion (+)	15 (44.1 %)	15 (34.9 %)	0.5
Lymphovascular invasion (+)	21 (61.8 %)	27 (62.8 %)	1.0
Grade			0.5
G1	9 (26.5 %)	7 (16.3 %)	
G2	17 (50 %)	27 (62.8 %)	
G3	8 (23.5%)	9 (20.9 %)	
Tumor location in lobe			0.8
Upper	13 (38.2 %)	19 (44.2 %)	
Middle	0 (0 %)	1 (2.3 %)	
Lower	21 (61.8 %)	23 (53.5 %)	
Tumor location in lung			0.04*
Central	1 (2.9 %)	9 (20.9 %)	
Peripheral	33 (97.1 %)	34 (79.1 %)	
TP53 mutation			0.8
nuclear staining (+)	24 (70.6 %)	28 (65.1 %)	
null	5 (14.7 %)	9 (20.9 %)	
p16 (+)	12 (35.3 %)	6 (14.0 %)	0.03*

*, $P < 0.05$.

Table 2. Clinical and pathological characteristics of high- (HME) and low-methylation epigenotypes (LME).

Variable	HME (n = 23)	LME (n = 54)	*P* value
Age (year ± SD)	70.1 ± 7.3	70.4 ± 6.8	0.9
Gender			
Male	20 (87.0 %)	49 (90.7 %)	0.7
Brinkman Index (pack-years)	51.7 ± 28.4	43.6 ± 24.4	0.2
IPF (+)	4 (17.4 %)	30 (55.6 %)	0.003*
Pathological stage			0.4
IA	8 (34.8 %)	10 (18.5 %)	
IB	3 (13.0 %)	17 (31.5 %)	
IIA	3 (13.0 %)	5 (9.3 %)	
IIB	4 (17.4 %)	12 (22.2 %)	
IIIA	4 (17.4 %)	7 (13.0 %)	
IIIB	1 (4.4 %)	1 (1.9 %)	
IV	0 (0 %)	2 (3.7 %)	
T stage			0.1
T1	10 (43.5 %)	14 (25.9 %)	
T2	10 (43.5 %)	28 (51.9 %)	
T3	2 (8.7 %)	12 (22.2 %)	
T4	1 (4.4 %)	0 (0 %)	
N stage			0.8
N0	16 (69.6 %)	41 (75.9 %)	
N1	3 (13.0 %)	6 (11.1 %)	
N2	4 (17.4 %)	7 (13.0 %)	
M stage			1.0
M1a	0 (0 %)	2 (3.7 %)	
Pathological pleural invasion (+)	6 (26.1 %)	24 (44.4 %)	0.2
Lymphovascular invasion (+)	16 (69.6 %)	32 (59.3 %)	0.4
Grade			0.7
G1	4 (17.4 %)	12 (22.2 %)	
G2	15 (65.2 %)	29 (53.7 %)	
G3	4 (17.4 %)	13 (24.1 %)	
Tumor location in lobe			0.6
Upper	8 (34.8 %)	24 (44.4 %)	
Middle	0 (0 %)	1 (1.9 %)	
Lobar	15 (65.2 %)	29 (53.7 %)	
Tumor location			0.5
Central	4 (17.4 %)	6 (11.1 %)	
Peripheral	19 (82.6 %)	48 (88.9 %)	
TP53 mutation			0.5
nuclear staining (+)	18 (78.3 %)	34 (63.0 %)	
null	3 (13.0 %)	11 (20.4 %)	
p16 (+)	8 (34.8 %)	10 (18.5 %)	0.1

*, $P < 0.05$.

Table 3. Analysis of prognostic factors using Cox proportional hazards.

Variable	Univariate analysis			Multivariate analysis		
	Hazard ratio	**95 % CI**	***P* value**	**Hazard ratio**	**95 % CI**	***P* value**
IPF (+)	2.38	0.96-6.20	0.06	1.95	0.62-6.53	0.3
Age						
< 70	Reference			-	-	-
≥ 70	1.56	0.63-4.19	0.3	-	-	-
Brinkman index	1.75	0.36-31.6	0.6	-	-	-
Epigenotype						
HME	Reference			Reference		
LME	8.85	1.82-159.1	0.003	10.4	1.88-195	0.005
Tumor location						
Central	Reference			-	-	-
Peripheral	0.96	0.27-6.14	1.0	-	-	-
Pathologic stage						
I	Reference			Reference		
II	0.48	0.44-5.14	0.5	1.04	0.24-3.93	1.0
III	5.23	1.62-16.9	0.006	6.8	0.63-71.3	0.1
IV	75.1	8.87-534	< 0.001	55.23	3.84-714	0.007
T stage						
1	Reference			-	-	-
2	1.85	0.60-6.84	0.3	-	-	-
3	2.39	0.63-9.69	0.2	-	-	-
4	13.3	0.66-98.5	0.08	-	-	-
Pathological pleural invasion (+)	1.87	0.75-4.72	0.2	-	-	-
N stage						
0	Reference			Reference		
1	4.3	1.15-13.4	0.03	3.84	0.53-22.6	0.2
2	4.91	1.5-14.2	0.01	0.36	0.03-4.79	0.4
Grade						
G1	Reference			-	-	-
G2	0.71	0.24-2.30	0.5	-	-	-
G3	1.12	0.31-4.05	0.9	-	-	-
Lymphovascular invasion (+)	2.67	1.01-8.36	0.05	5.98	1.62-29.4	0.1

Figure legends

Figure 1. Study design and representative radiological and pathological findings. (A) Flow chart of the study. (B) Representative pathological findings following hematoxylin/eosin staining. (C,D) Typical radiograph of squamous cell carcinomas with IPF at peripheral lung tissues, and of (D) squamous cell carcinomas without IPF at central lung tissues.

Figure 2. Infinium 450K analysis of squamous cell carcinomas with or without IPF, and of surrounding non-cancerous lung tissues. (A) Unsupervised two-way hierarchical clustering (*top*) of 570 high- and intermediate-scoring CpG probes from 20/77 squamous cell carcinomas, 11/20 surrounding non-cancerous tissues, and 2/2 control normal samples. β-values from peripheral blood cells and IMR-90 fibroblasts are also shown, along with those of genes targeted by polycomb repressor in embryonic cells (black bars in *top right*; ref. 34). *Bottom*, source of tissue and IPF status. *Orange*, squamous cell carcinoma; *blue*, surrounding non-cancerous tissues; *purple*, control normal lung samples. *Black bar*, with IPF; *white bar*, without IPF. While lung squamous cell carcinomas generally showed aberrant hypermethylation compared to control lung tissues and surrounding non-cancerous tissues, these carcinomas are classifiable into low- (LME) and high-methylation epigenotypes (HME). Squamous cell carcinomas with IPF are significantly associated with LME ($P = 0.04$). (B) Aberrantly hypermethylated genes are more frequently observed in squamous cell carcinomas without IPF than in tumors with IPF. However, aberrant hypermethylation is observed in more genes in surrounding non-cancerous lung tissues with IPF than in those without

IPF.

Figure 3. Genes differentially methylated between lung tissues with or without IPF and between squamous cell carcinoma epigenotypes. (A) Extraction of methylation marker genes. β-value > 0.2 was considered indicative of hypermethylation. Group #1: genes hypermethylated in all HME and LME squamous cell carcinomas, as well as in lung tissues with or without IPF. Group #2: genes unmethylated in control lungs but aberrantly hypermethylated in HME and LME squamous cell carcinomas and in surrounding lung tissues with or without IPF. Group #3: genes aberrantly hypermethylated in HME and LME squamous cell carcinomas, and in surrounding lung tissues with IPF. Group #4: genes aberrantly hypermethylated in HME and LME squamous cell carcinomas. Group #5: genes aberrantly hypermethylated in HME squamous cell carcinomas. Group #6: genes unmethylated in all groups. Genes hypermethylated in LME and HME (groups #2, #3, and #4, n = 464) were defined as common markers. Genes hypermethylated specifically in HME (group #5, n = 287) were defined as HME markers. (B) Ratio of PRC1 and PRC2 target genes in markers and in all genes. *Black box*, PRC1 or PRC2 target genes. Both PRC1 and PRC2 target genes were significantly enriched in HME and common markers, although PRC1 target genes were significantly more enriched in common markers than in HME markers. *, $P < 0.05$; **, $P < 0.001$. (C) Genes related to transcription and cell adhesion were significantly enriched in common markers. (D) Genes related to negative regulation of growth were significantly enriched in HME markers.

Figure 4. Validation of DNA methylation by pyrosequencing. (A) Methylation status of

five HME markers and three common markers identified by Infinium (*top*) was validated by precise measurement of methylation levels at multiple consecutive CpG sites around the Infinium probe using pyrosequencing (*middle*). *Black bar*, surrounding lung tissue with IPF; *white bar*, surrounding lung tissue without IPF (*bottom*). (B) The full repertoire of clinical specimens of squamous cell carcinomas (n = 77) and surrounding non-cancerous tissues (n = 20) were also pyrosequenced. Specimens are aligned left to right according to a descending scale of average methylation at markers. Squamous cell carcinomas with IPF are clustered on the right (P = 0.01 by Kolmogorov-Smirnov test), implying that these carcinomas are associated with frequent hypomethylation. (C) Correlation between squamous cell carcinomas and methylation status of five HME markers. Methylation of *DLEC1*, *CRIP3*, and *ALDH7A1* was significantly lower in squamous cell carcinomas with IPF, while *CFTR* and *MT1M* were relatively hypomethylated.

Figure 5. Impact of epigenotype on survival from squamous cell carcinomas. (A) Squamous cell carcinoma samples hypermethylated (methylation level > 20 %, β-value > 0.2 in Infinium analysis) in four or five HME markers were considered HME. All others were considered LME. (B) Squamous cell carcinomas with IPF are significantly associated with LME (P = 0.003). (C-G) LME was significantly associated with worse prognosis in (C) the full cohort, (D) in patients with stage I and II cancer, (E) in a cohort that includes our cases and comparable cases in The Cancer Genome Atlas, (F) in patients with LME squamous cell carcinomas without IPF, and (G) in patients with stage I and II squamous cell carcinomas without IPF.

Figures

Figure 1.

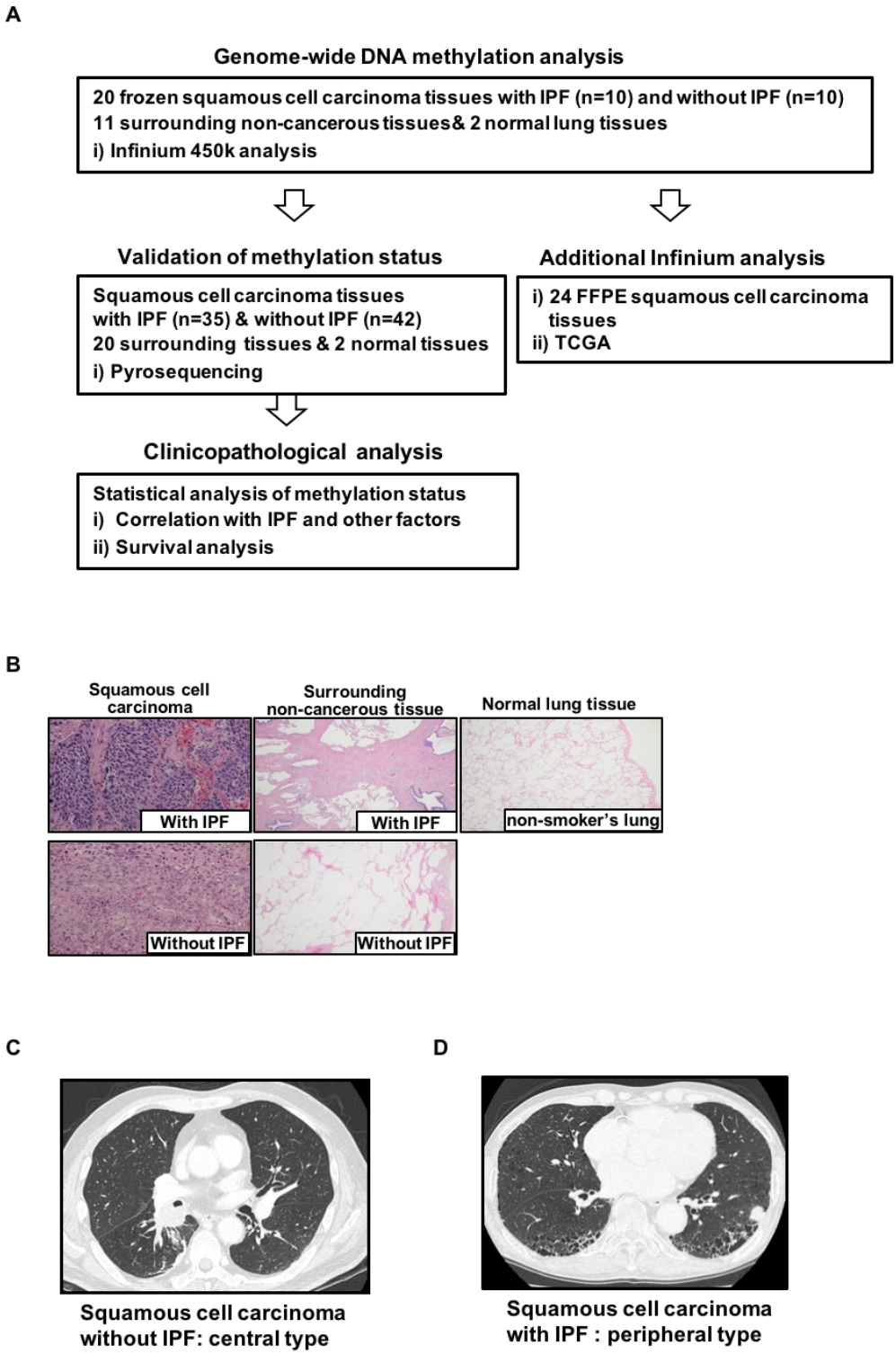
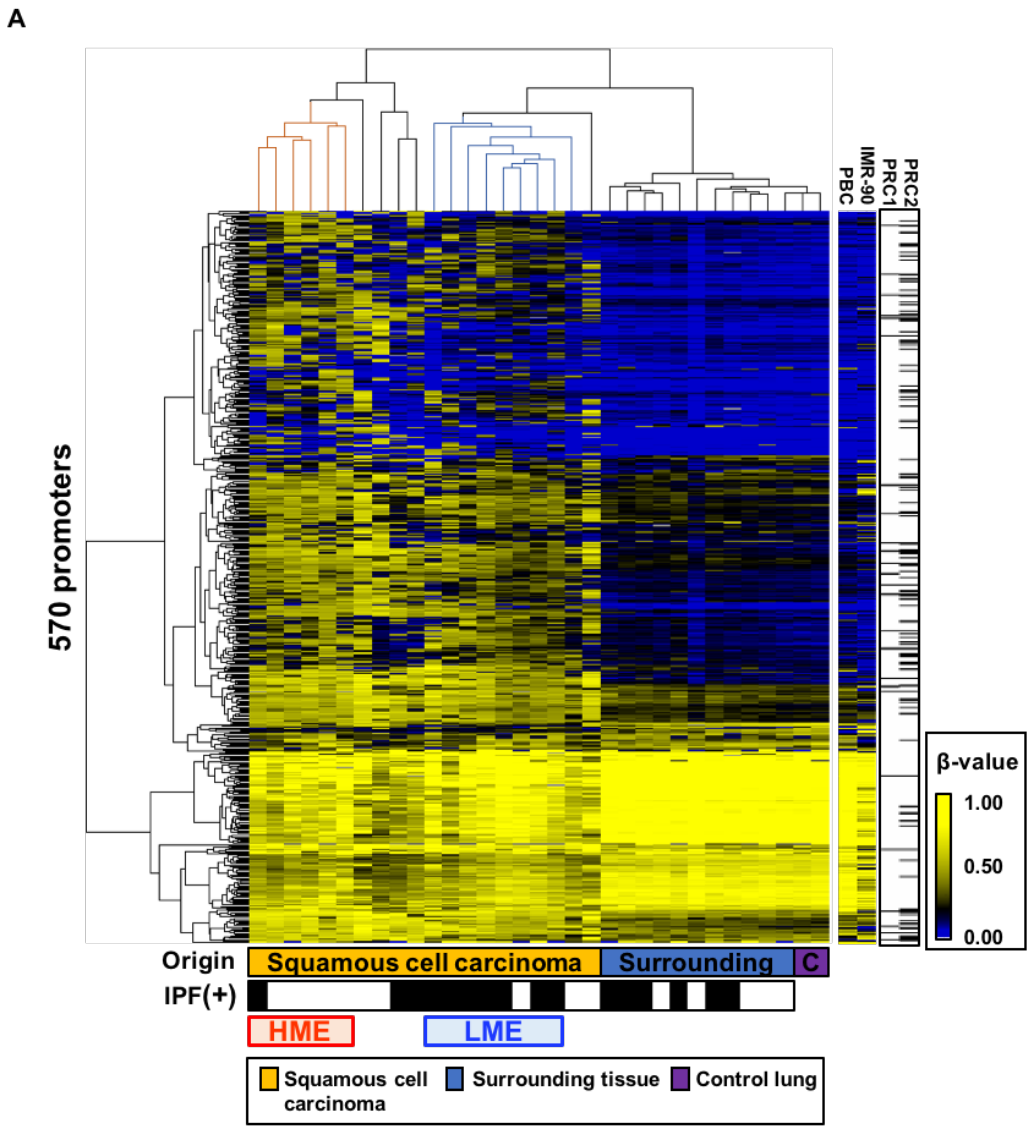


Figure 2.


A
570 promoters
PBC
IMR-90
PRC1
PRC2
β-value
1.00
0.50
0.00
Origin
Squamous cell carcinoma
Surrounding
C
IPF(+)
HME
LME
Squamous cell carcinoma
Surrounding tissue
Control lung

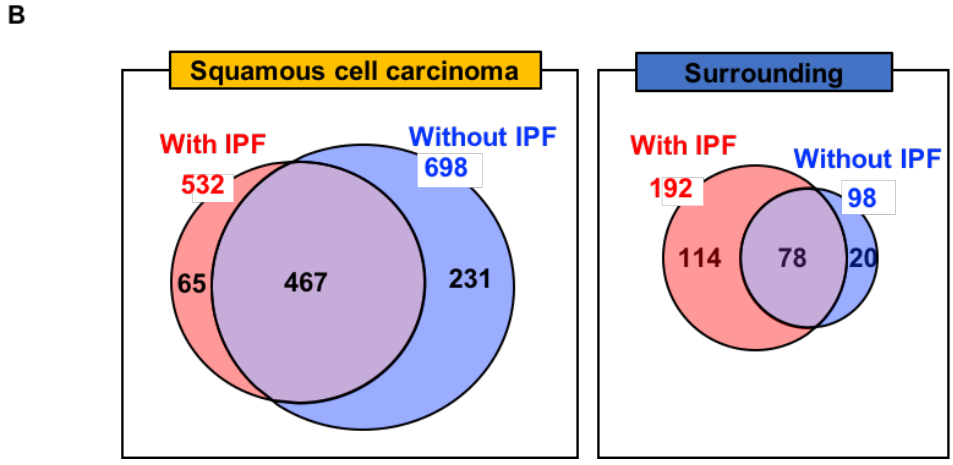

B
Squamous cell carcinoma
With IPF
Without IPF
532
698
65
467
231
Surrounding
With IPF
Without IPF
192
98
114
78
20

Figure 3.

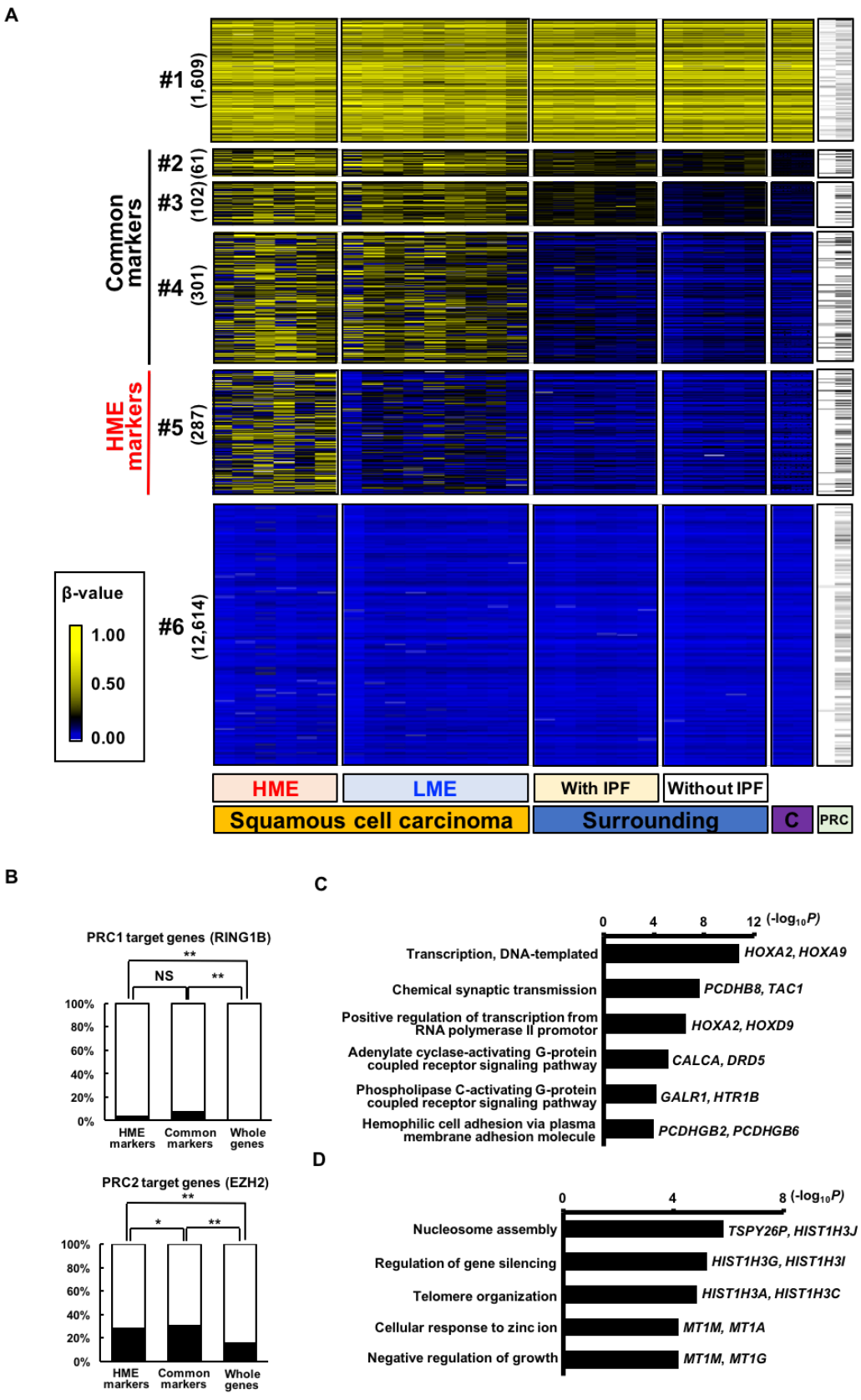

A
#1 (1,609)
#2 (61)
#3 (102)
#4 (301)
#5 (287)
#6 (12,614)
Common markers
HME markers
β-value
1.00
0.50
0.00
HME
LME
With IPF
Without IPF
Squamous cell carcinoma
Surrounding
C
PRC
B
PRC1 target genes (RING1B)
NS
**
100%
80%
60%
40%
20%
0%
HME markers
Common markers
Whole genes
PRC2 target genes (EZH2)
*
C
0
4
8
12 (-log10P)
Transcription, DNA-templated
HOXA2, HOXA9
Chemical synaptic transmission
PCDHB8, TAC1
Positive regulation of transcription from RNA polymerase II promotor
HOXA2, HOXD9
Adenylate cyclase-activating G-protein coupled receptor signaling pathway
CALCA, DRD5
Phospholipase C-activating G-protein coupled receptor signaling pathway
GALR1, HTR1B
Hemophilic cell adhesion via plasma membrane adhesion molecule
PCDHGB2, PCDHGB6
D
8 (-log10P)
Nucleosome assembly
TSPY26P, HIST1H3J
Regulation of gene silencing
HIST1H3G, HIST1H3I
Telomere organization
HIST1H3A, HIST1H3C
Cellular response to zinc ion
MT1M, MT1A
Negative regulation of growth
MT1M, MT1G

Figure 4.

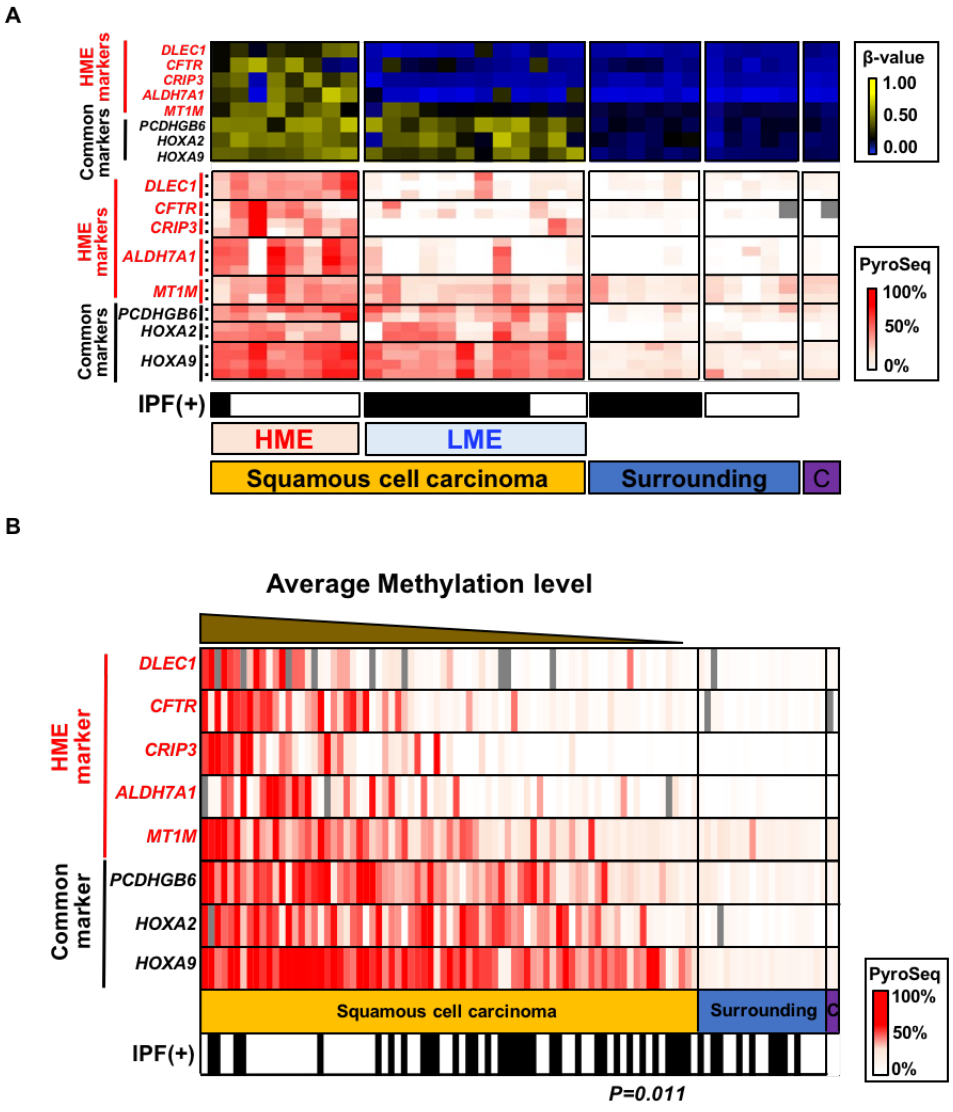
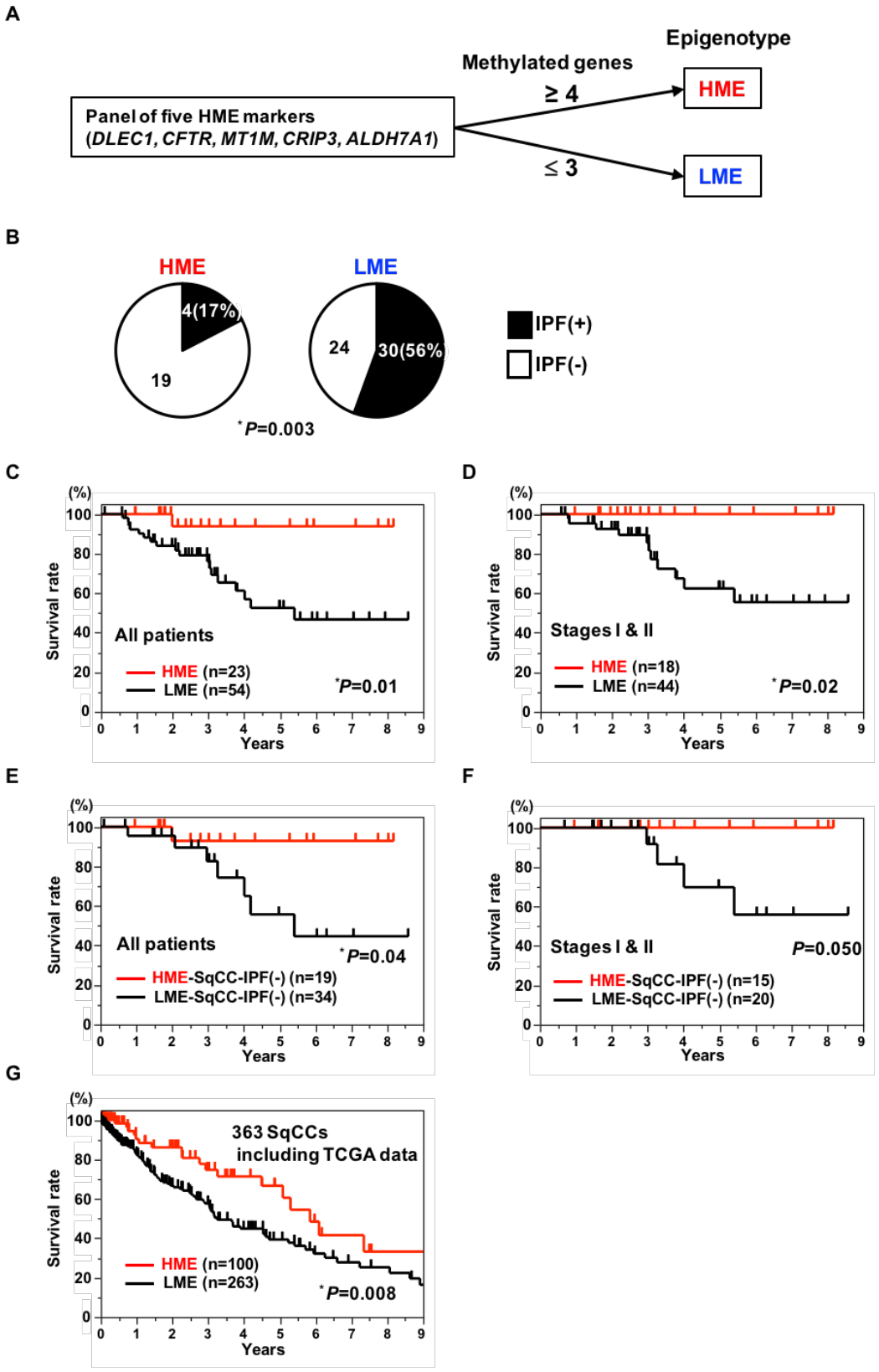

A
HME markers
Common markers
DLEC1
CFTR
CRIP3
ALDH7A1
MT1M
PCDHGB6
HOXA2
HOXA9
β-value
1.00
0.50
0.00
PyroSeq
100%
50%
0%
IPF(+)
HME
LME
Squamous cell carcinoma
Surrounding
C
B
Average Methylation level
HME marker
Common marker
Squamous cell carcinoma
Surrounding
C
IPF(+)
P=0.011
C
DLEC1
CFTR
CRIP3
ALDH7A1
MTM1
*P=0.03
P=0.1
*P=0.049
*P=0.02
P=0.1
SqCC IPF(+)
SqCC IPF(–)

Figure 5.


A
Epigenotype
Methylated genes
≥ 4
HME
Panel of five HME markers
(DLEC1, CFTR, MT1M, CRIP3, ALDH7A1)
≤ 3
LME
B
HME
LME
4(17%)
19
24
30(56%)
IPF(+)
IPF(-)
*P=0.003
C
(%)
Survival rate
All patients
HME (n=23)
LME (n=54)
*P=0.01
Years
D
Stages I & II
HME (n=18)
LME (n=44)
*P=0.02
E
All patients
*P=0.04
HME-SqCC-IPF(-) (n=19)
LME-SqCC-IPF(-) (n=34)
F
Stages I & II
P=0.050
HME-SqCC-IPF(-) (n=15)
LME-SqCC-IPF(-) (n=20)
G
363 SqCCs
including TCGA data
HME (n=100)
LME (n=263)
*P=0.008

Supplementary Materials and Methods

Detailed Materials and Methods are provided below.

Clinical samples and cell lines

From patients who underwent lung surgery from 2004 to 2014 at Chiba University Hospital, 45 frozen and 32 formalin-fixed paraffin-embedded primary lung squamous cell carcinomas with or without IPF were obtained with informed consent. The study was approved by the institutional review board at Chiba University (No. 806). IPF was diagnosed in 35 of these tissues based on the ATS/ERS/JRS/ALAT Statement [E1]. The remaining 42 tissues were free of IPF. All tissues were microscopically confirmed by two independent pathologists to contain > 40 % cancer cells, and were dissected to enrich cancer cells when necessary. Surrounding non-cancerous lung tissues (n = 20) with or without IPF were used as control, along with normal lung tissues from non-smokers with adenocarcinoma. Patients without IPF were eligible if they have forced expiratory volume in one second ($FEV_{1.0}$) / forced vital capacity (FVC) >70%, in order to minimize the impact of chronic obstructive pulmonary disease.

DNA was extracted from frozen samples using QIAquick DNA Mini Kit (Qiagen, Hilden, Germany) and from paraffin-embedded samples using QIAamp DNA FFPE Tissue kit (Qiagen). Normal human lung IMR-90 fibroblasts were obtained from Health Science Research Resources Bank. The composition of peripheral blood cells was extracted from The Cancer Genome Atlas. Twenty frozen squamous cell carcinoma samples, 11 surrounding non-cancerous lung tissues with (n = 6) or without (n = 5) IPF, and two control lung samples were initially surveyed on Infinium 450K chips. The full set of 77 squamous cell carcinoma samples and 20 surrounding lung tissues were

used for validation by pyrosequencing.

Age, gender, smoking history, tumor location, pathologic TNM stage (p-stage), tumor grade, pleural invasion, and lymphovascular invasion were collected from our cohort (Table 1), as well as from 263 squamous cell carcinoma cases in The Cancer Genome Atlas. Tumors arising at main and segmental bronchi were considered central (Figure 1C), while tumors arising at subsegmental bronchi or at more peripheral locations were considered peripheral (Figure 1D), as previously defined [E2-4].

Analysis of DNA methylation

Infinium 450K BeadChip (Illumina) contains approximately 485,000 CpG sites covering 99 % of RefSeq genes at an average of 17 sites per gene. Each CpG site was scored in β-values ranging from 0.00 to 1.00 based on signals from a methylated probe relative to the sum of signals from methylated and unmethylated probes. Bisulfite conversion was performed with Zymo EZ DNA Methylation Kit (Zymo Research, Irvine, CA), using 500 ng of genomic DNA per sample. Whole-genome amplification, labeling, hybridization, and scanning were performed according to the manufacturer's protocols. Infinium data were submitted to the Gene Expression Omnibus under accession numbers (GSM3447735-GSM3447767).

Genes targeted by polycomb repressive complex were identified using available ChIP-sequencing data from embryonic stem cells [E5]. Only probes nearest to transcription start sites were analyzed for methylation if multiple probes were available against a promoter. CpG sites were scored based on previous reports [E6,7], and only high- and intermediate-scoring probes with CpG score > 0.48 were included in subsequent analyses. Genes in X and Y chromosomes were also

excluded to avoid gender differences. Genes with β-values > 0.2 were considered hypermethylated.

Additional formalin-fixed, paraffin-embedded specimens were analyzed for DNA methylation using Infinium 450K BeadArray (Illumina). Bisulfite conversion was achieved using Zymo EZ DNA Methylation Kit (Zymo Research, Irvine, CA), and using 500 ng genomic DNA per sample. Whole genome amplification, labeling, hybridization, and scanning were performed according to the manufacturer's protocols. Probes with comparable β-values in both frozen and formalin-fixed paraffin-embedded specimens were previously selected [E8], and used for subsequent clustering analysis. For comparison, DNA methylation data for 263 frozen samples of squamous cell carcinoma were obtained from The Cancer Genome Atlas.

Pyrosequencing

Genomic DNA (500 ng) from the 77 squamous cell carcinomas in the cohort was bisulfite converted using EZ DNA Methylation Kit (Zymo Research), and suspended in 40 μL distilled water. Primers to amplify promoters with several CpG sites were designed in Pyro Q-CpG Software (Qiagen) to ensure products < 120 bp, and to include zero or one CpG site per primer (Supplementary Table S1). For cytosine residues in CpG sites within a primer, a nucleotide that does not anneal to C or U was chosen, such as adenosine (A). Biotinylated PCR products were bound to Streptavidin Sepharose High Performance Beads (Amersham Biosciences, Uppsala, Sweden), washed, and denatured using 0.2 M NaOH. After addition of 0.3 μM sequencing primer, samples were pyrosequenced in PyroMark Q96 (Qiagen) [E6] according to the manufacturer's instructions. Methylation control samples (0 %, 25 %, 50 %, 75 %, and 100 %) were prepared as described previously [E9], and used to confirm the accuracy of pyrosequencing assays. Methylation levels > 20 % were considered

indicative of hypermethylation.

Immunohistochemistry

Protein expression of TP53 and $p16^{INK4A}$ was analyzed by immunohistochemistry in a BenchMark ULTRA automated staining system (Roche, Basel, Switzerland), using mouse monoclonal anti-TP53 (clone DO-7) and anti-p16 (CINtec p16 Histology) from Roche. Samples with strong and uniform nuclear staining for TP53 were considered to harbor activating mutations, while those without nuclear staining were considered to harbor null mutations. Sporadic nuclear staining in surrounding non-cancerous cells, such as inflammatory cells, were used as wild-type internal controls. Finally, samples with nuclear and cytoplasmic immunoreactivity for p16 in > 70 % of tumor cells were considered to express p16.

Gene Ontology analysis

Enrichment of Gene Ontology terms (biologic process: BP_ALL) was analyzed using Functional Annotation at DAVID Bioinformatics Resources (http://david.abcc.ncifcrf.gov/).

Statistical Analysis

Unsupervised two-way hierarchical clustering was conducted in Cluster 3.0 (http://bonsai.hgc.jp/~mdehoon/software/cluster/) or in R (www.r-project.org/). Heatmaps were drawn in Java TreeView (http://jtreeview.sourceforge.net/). Association between DNA methylation and clinicopathological factors, including presence or absence of IPF, was analyzed by Fisher's exact test, χ^2 test, or Kolmogorov-Smirnov test for categorical variables, and by Mann-Whitney test for

continuous data. Cancer-specific overall survival was measured from enrollment date until date of cancer-related death. Patients who died from causes unrelated to cancer were censored. Survival curves were constructed by the Kaplan-Meier method, and *P* values were calculated by log-rank test. Multivariate analysis was performed using Cox proportional hazards. Analyses were performed in JMP pro 12.0.1 (SAS Institute Inc., Cary, NC), with $P < 0.05$ considered statistically significant.

Supplementary Table E1. Primers for pyrosequencing.

Gene	strand	Primer type	Primer sequence	Anneal (°C)	Product (bp)	Probe position of Infinium
HME markers						
DLEC1	top	Fwd	GGATTAAAGAGTGTTAGGGGATAATGT	55	85	-95
		Rev*	AACTTCCAAATAAACTAACTAAAACTACT			
		Seq	GTTAGGGGATAATGTGG			
CFTR	top	Fwd	ATTTGGGGTAGGATTAGGTAGTAT	55	82	53
		Rev*	CACACCACCCCTTCCTTTTA			
		Seq	GGGGTAGGATTAGGTAGTATT			
CRIP3	top	Fwd	AGGGAGGTTTAAAGTTGAGT	55	120	43
		Rev*	AACAAATCCAACTCATAACTCC			
		Seq	GAGGGGGAGGAATTGA			
ALDH7A1	top	Fwd	GGGAGAATTTATTGGGTAGAT	55	99	-81
		Rev*	AAATTCAACCCCAAACTCCTCC			
		Seq	TGGGTAGATTAGATTTTTTTTAT			
MT1M	top	Fwd	ATTAGTGGTGTAGGAGTAGTTG	55	73	-107
		Rev*	CTCCATTTATCCCTTAAAATCTCCAACCTT			
		Seq	AGTAGTTGGGGTTTATT			
Common markers						
PCDHGB6	top	Fwd	GTTGGTAATTAGTTGTGTTGGGAAAGG	55	70	146
		Rev*	AAAAAAAAACCAACTCCCACACAAAAC			
		Seq	GTTGGGAAAGGTTTG			
HOXA2	top	Fwd	TGGGTTTTAATTAAAAGAAGTTATTGT	55	91	33
		Rev*	AAATAAAACCTCCCTAACCCCCAATAAA			
		Seq	TTATTGTTTGTTTGATTGATTTAA			
HOXA9	bottom	Fwd*	TTAGGGTTTTAGTGGTGGTTATTA	55	111	-35
		Rev	CCCCCCCATACCACCAAATTATTACATA			
		Seq	ATTACATAAAATCTACAATTTCA			

*Primers with 5'-biotin tag. *Probe position*, position of the CpG site analyzed in Infinium assay, with the transcription start site marked as +1. *Top*/*bottom*, the strand used as template in bisulfite PCR. *Fwd*, forward primer. *Rev*, reverse primer. *Seq*, primer sequence. Primers were designed to determine methylation status at multiple CpG sites, including the C at the probe position.

Supplementary Reference

E1. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, Colby TV, Cordier JF, Flaherty KR, Lasky JA, Lynch DA, Ryu JH, Swigris JJ, Wells AU, Ancochea J, Bouros D, Carvalho C, Costabel U, Ebina M, Hansell DM, Johkoh T, Kim DS, King TE Jr, Kondoh Y, Myers J, Müller NL, Nicholson AG, Richeldi L, Selman M, Dudden RF, Griss BS, Protzko SL, Schünemann HJ, ATS/ERS/JRS/ALAT Committee on Idiopathic Pulmonary Fibrosis. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Respir Crit Care Med 2011;183:788-824.

E2. Funai K, Yokose T, Ishii G, Araki K, Yoshida J, Nishimura M, Nagai K, Nishiwaki Y, Ochiai A. Clinicopathologic characteristics of peripheral squamous cell carcinoma of the lung. Am J Surg Pathol 2003;27:978-984.

E3. Sakurai H, Asamura H, Watanabe S, Suzuki K, Tsuchiya R. Clinicopathologic features of peripheral squamous cell carcinoma of the lung. Ann Thorac Surg 2004;78:222-227.

E4. Shimmyo T, Hashimoto T, Kobayashi Y, Miyagi Y, Ishikawa Y, Nakagawa K, Osada H, Tsuchiya E. p53 mutation spectra for squamous cell carcinomas at different levels of human bronchial branches. Int J Cancer 2006;119:501-507.

E5. Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, Presser A, Nusbaum C, Xie X, Chi AS, Adli M, Kasif S, Ptaszek LM, Cowan CA, Lander ES, Koseki H, Bernstein BE. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet 2008;4:e1000242.

E6. Matsusaka K, Kaneda A, Nagae G, Ushiku T, Kikuchi Y, Hino R, Uozaki H, Seto Y, Takada K,

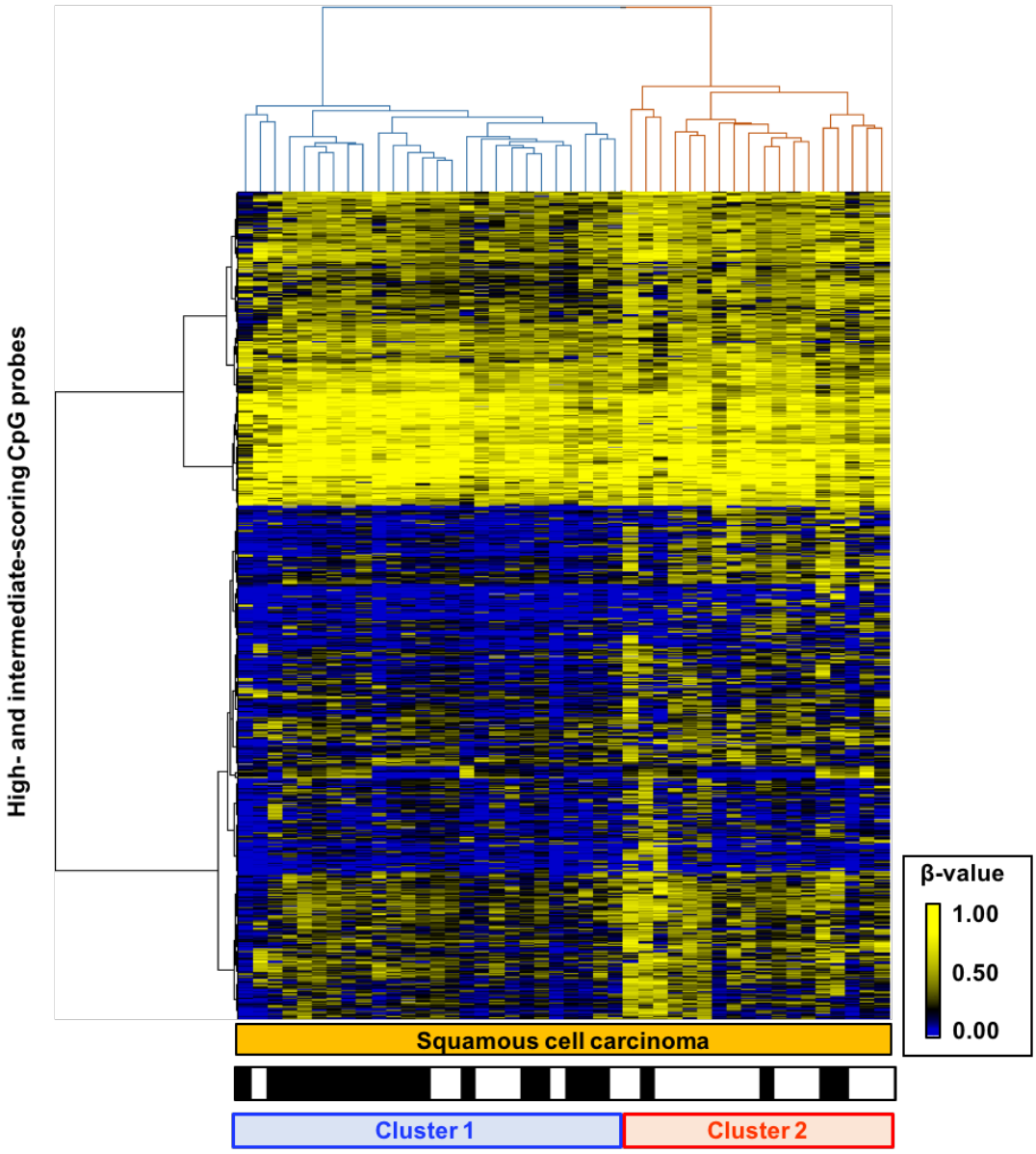
Aburatani H, Fukayama M. Classification of Epstein-Barr virus-positive gastric cancers by definition of DNA methylation epigenotypes. Cancer Res 2011;71:7187-7197.

E7. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M, Schübeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007;39:457-466.

E8. Takane K, Fukuyo M, Matsusaka K, Ota S, Rahmutulla B, Matsushita K, Miyauchi H, Nakatani Y, Matsubara H, Kaneda A. The frequency of promoter DNA hypermethylation is decreased in colorectal neoplasms of familial adenomatous polyposis. Oncotarget 2018;9:32653-32666

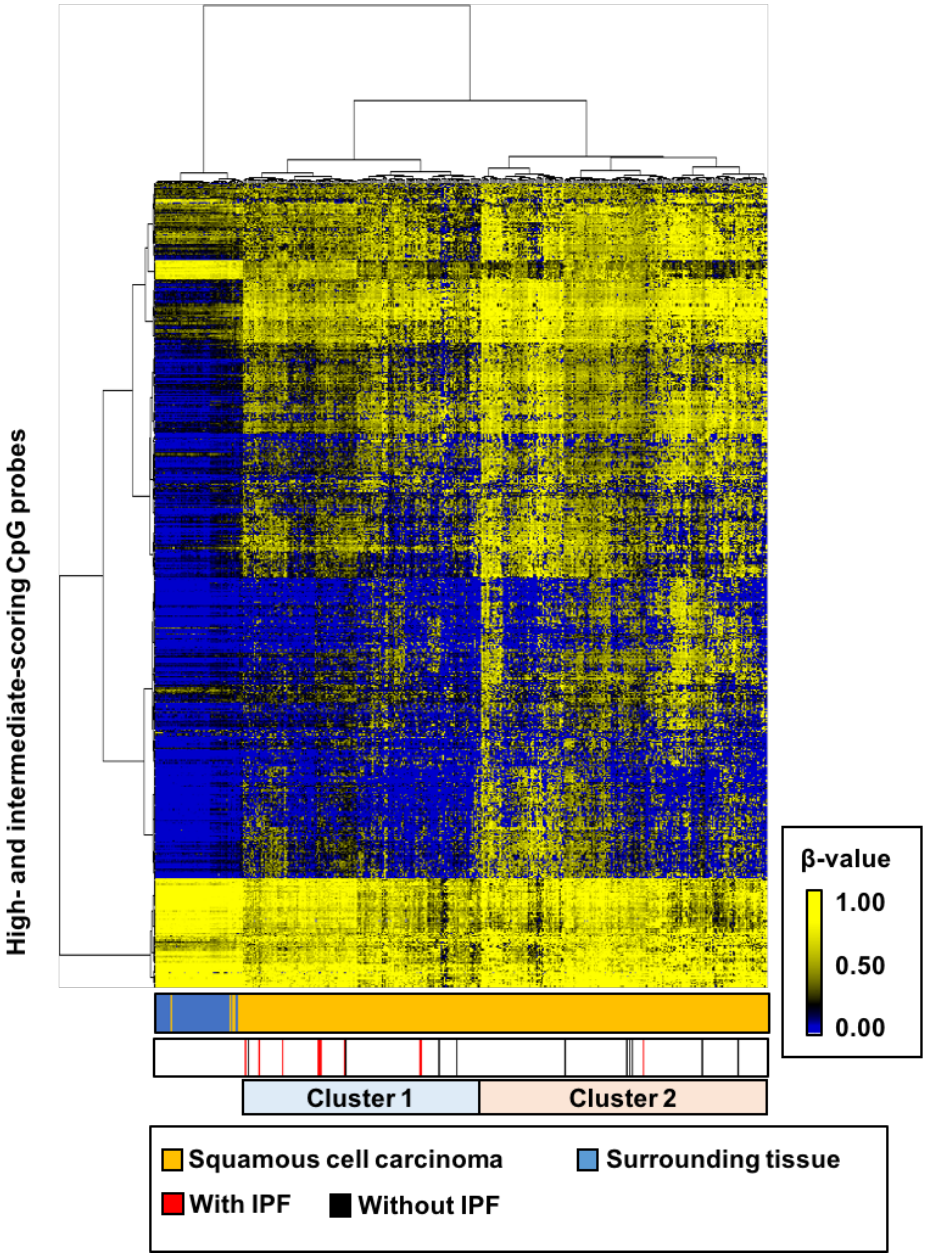
E9. Yagi K, Akagi K, Hayashi H, Nagae G, Tsuji S, Isagawa T, Midorikawa Y, Nishimura Y, Sakamoto H, Seto Y, Aburatani H, Kaneda A. Three DNA methylation epigenotypes in human colorectal cancer. Clin Cancer Res 2010;16:21-33.

Figure E1.



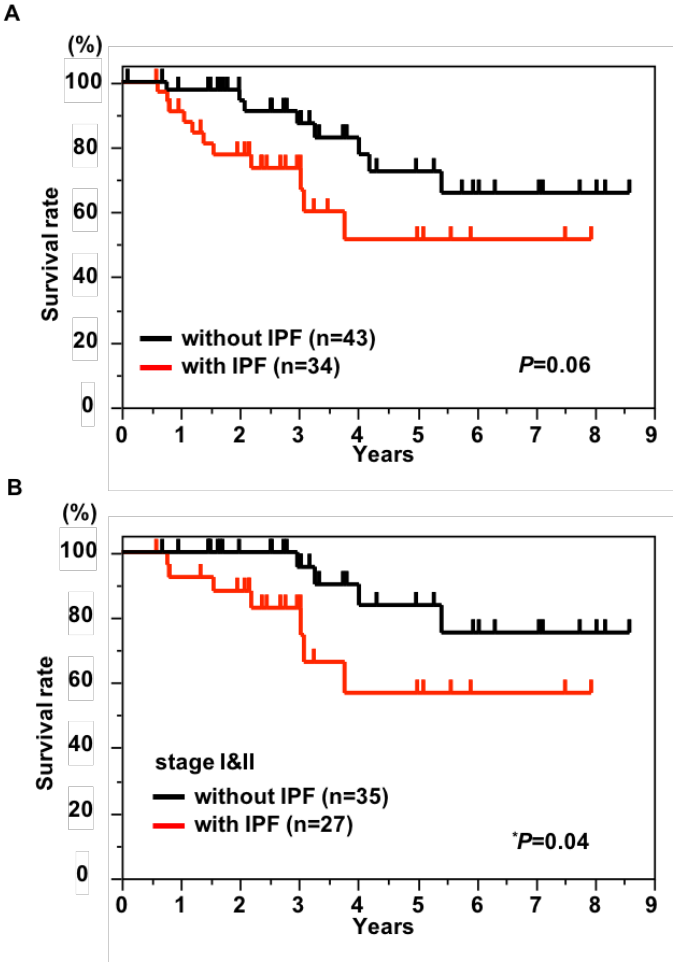
Supplementary Figure E1. Infinium analysis of 20 frozen and 24 formalin-fixed, paraffin-embedded squamous cell carcinomas with (n = 22) or without IPF (n = 22). Samples were clustered by unsupervised two-way hierarchical clustering. *Top*, 783 high- and intermediate-scoring CpG probes were used for clustering. *Bottom*, IPF status. *Black bar*, squamous cell carcinomas with IPF; *white bar*, squamous cell carcinomas without IPF. Specimens were classified into two clusters, of which one was enriched with squamous cell carcinomas with IPF (Cluster 1) with relatively lower methylation levels ($P = 0.005$).

Figure E2.



Supplementary Figure E2. Infinium 450K analysis of squamous cell carcinomas with or without IPF in the current cohort, and of cases in The Cancer Genome Atlas. In total, 390 squamous cell carcinomas and 55 surrounding non-cancerous lung tissues were clustered by unsupervised two-way hierarchical clustering. *Top*, 591 high- and intermediate-scoring CpG probes were used for clustering. *Bottom*, source of samples and IPF status. *Orange*, squamous cell carcinoma; *blue*, surrounding non-cancerous tissues. *Red bar*, squamous cell carcinoma with IPF; *black bar*, squamous cell carcinoma without IPF. Carcinomas were classified into two methylation subtypes. Most of carcinomas with IPF (9/10, 90 %) are classified in Cluster 1, which has lower methylation levels than Cluster 2. In contrast, only 4/10 (40 %) of carcinomas without IPF are in this cluster ($P = 0.06$).

Figure E3.



Supplementary Figure E3. Impact of IPF on survival from squamous cell carcinomas. Squamous cell carcinomas with IPF are modestly associated with worse prognosis in (A) all patients, and in (B) patients with stage I and II cancers.