Sample ID	Age	Gender	Smoking status and other features	Total tags	Unique tags
Ad-9 ⁽¹⁾	-	-	well differentiated adenocarcinoma with	35,916	15,131
			lymphoplasmatic infiltration (cancer)		
Ad-10 ⁽¹⁾	-	-	poorly differentiated adenocarcinoma	86,887	29,062
			with lymphoplastic infiltration,		
(1)		_	microdissected (cancer)		00.004
UIHMO ("	-	F	tocal fibrosis, chronic inflammation,	98,962	32,281
$C_{ald} = (2)$	48 to 76	M/F	samples from 14 smokers, pulmonary	30 070	12 188
Gold-2	40 10 70	101/1	function test consistent with moderate	50,070	12,100
			COPD		
Gold-0 (2)	45 to 76	M/F	samples from 12 smokers, pulmonary	29,273	20,193
			function test exhibit non-obstruction		
NB1 ⁽¹⁾	20	М	young adult, normal lung	88,708	24,764
NLP-1 ⁽³⁾	54 - 84	1 M/3 F	normal lung parenchyma pool of 4	66,214	17,846
NLP-2 ⁽³⁾	65 - 77	М	normal lung parenchyma pool of 4	64,214	20,003
NT12 ⁽⁴⁾	69	М	normal lung parenchyma, non-smoker	55,954	17,850
NT29 ⁽⁴⁾	17	М	normal lung parenchyma, non-smoker	79,230	22,409
IPF-1 ⁽⁴⁾	67	М	lung biopsy tissue, non-smoker ⁽⁵⁾	99,936	32,106
IPF-2 ⁽⁴⁾	56	F	lung biopsy tissue, smoker ⁽⁵⁾	70,028	22,711
IPF-6 ⁽⁴⁾	72	М	lung biopsy tissue , non-smoker ⁽⁵⁾	57,446	28,722
IPF-7 ⁽⁴⁾	57	М	lung biopsy tissue, smoker ⁽⁵⁾	96,954	31,591
IPF-8 ⁽⁴⁾	67	М	lung biopsy tissue, non-smoker ⁽⁵⁾	86,484	31,186
IPF-9 ⁽⁴⁾	70	М	lung biopsy tissue, smoker ⁽⁵⁾	97,926	29,450
IPF-3 ⁽⁴⁾	64	М	lung biopsy tissue, smoker ⁽⁶⁾	113,110	33,071
IPF-4 ⁽⁴⁾	67	F	lung biopsy tissue, smoker ⁽⁶⁾	107,624	33,987
IPF-5 ⁽⁴⁾	64	F	lung biopsy tissue, smoker ⁽⁶⁾	90,632	29,887
IPF-10 ⁽⁴⁾	67	F	lung biopsy tissue, non-smoker ⁽⁶⁾	59,428	26,977
IPF-11 ⁽⁴⁾	61	М	lung biopsy tissue, non-smoker ⁽⁶⁾	43,432	15,212
IPF-12 ⁽⁴⁾	77	М	lung biopsy tissue, smoker ⁽⁶⁾	31,932	12,858

Table S1	Summary SAC	E librarias	included in	this study
radie ST.	SUMMARY SAU	JE IIDIANES	inciuded in	Inis siuav

SAGE libraries downloaded from the SAGE Genie Website (<u>http://cgap.nci.nih.gov/SAGE</u>).
Gold SAGE libraries described by Ning *et al.*, [1].
NLP SAGE libraries described by Lonergan *et al.*, [2] and downloaded from the GEO website (<u>http://ncbi/geo/</u>).
SAGE libraries generated for this study.
And (6) respectively; progressor and relatively stable disease group as defined by clinical characteristics. Smoking status refers to a current or ever smoker.

- Ning W, Li C-J, Kaminski N, Feghali-Bostwick CA, Alber SM, et al. (2004) Comprehensive gene expression profiles reveal pathways related to the pathogenesis of chronic obstructive pulmonary disease. Proceedings of the National Academy of Sciences 101: 14895-14900.
- 2. Lonergan KM, Chari R, deLeeuw RJ, Shadeo A, Chi B, et al. (2006) Identification of Novel Lung Genes in Bronchial Epithelium by Serial Analysis of Gene Expression. Am J Respir Cell Mol Biol 35: 651-661.