



Supplementary Figure 1:

Representative sequencing traces from a tumor (T) and adjacent, non-cancerous (A) sample from a patient with the variant allele (18T and A) and a patient without the variant allele (7C). Solid arrows point to heterozygosity (T/G) at the fourth nucleotide of LCS6. The double arrow points to the homozygous T allele.

Patient	Sex	Population	Age	Pack-Year	Cancer Type	LCS6 Variant
1	F	Caucasian	64	150	Adenocarcinoma	N
2	M	Caucasian	73	20	Adenosquamous	N
3	M	Caucasian	64	50	Large cell	N
4	F	Caucasian	76	unknown	Adenocarcinoma	N
5	M	Hispanic	54	70	Squamous cell	N
6	M	Caucasian	64	10	Squamous cell	Y
7	M	Caucasian	86	60	Squamous cell	N
8	F	Caucasian	54	0	Adenocarcinoma	N
9	F	Caucasian	58	40	Adenosquamous	N
10	F	Caucasian	65	150	Adenocarcinoma	N
11	M	Caucasian	64	17	Adenocarcinoma	Y
12	F	Hispanic	65	20	Adenocarcinoma	N
13	F	Caucasian	89	unknown	Adenocarcinoma	N
14	F	Caucasian	47	0	Large cell	N
15	M	Caucasian	48	45	Squamous cell	Y
16	M	Caucasian	85	70	Adenocarcinoma	N
17	M	Caucasian	86	75	Adenosquamous	N
18	F	Caucasian	49	15	Squamous cell, Adenocarcinoma	Y
19	M	Asian	55	40	Squamous cell	N
20	F	Caucasian	74	0	Adenocarcinoma	N
21	F	Caucasian	58	2	Adenocarcinoma	Y
22	F	AA	40	20	Adenocarcinoma	N
23	F	Caucasian	52	30	Adenocarcinoma, BAC	N
24	M	Caucasian	50	40	Large cell	N
25	M	Caucasian	69	105	Adenocarcinoma	N
26	F	Caucasian	75	50	Adenocarcinoma	N
27	M	Caucasian	83	60	Large cell	N
28	M	Caucasian	42	20	Adenocarcinoma	Y
29	F	Caucasian	52	35 years	Adenocarcinoma	N
30	M	Caucasian	71	70	Adenocarcinoma, Squamous cell	N
31	F	AA	69	50	Adenocarcinoma	N
32	M	Caucasian	44	50	Adenocarcinoma	N
33	F	Caucasian	66	2 nd hand	Adenocarcinoma	Y
34	F	Caucasian	73	75	Adenocarcinoma	N
35	M	AA	72	30	Adenocarcinoma	N
36	F	Caucasian	72	120	Squamous cell	Y
37	F	Caucasian	62	50	Adenocarcinoma	N
38	M	Caucasian	74	unknown	Adenocarcinoma	N
39	M	Caucasian	78	20	Adenocarcinoma	N
40	M	Caucasian	68	32	Adenocarcinoma	N
41	M	Caucasian	66	40	Adenosquamous	N
42	M	Caucasian	88	40	Squamous cell	Y
43	F	Caucasian	63	60	Adenocarcinoma	N
44	F	AA	69	0	Adenocarcinoma	N

45	M	Caucasian	60	17	Squamous cell	Y
46	F	AA	49	50	Adenocarcinoma	N
47	F	Caucasian	65	60	Adenosquamous	N
48	M	Caucasian	63	45	Adenocarcinoma	N
49	F	Asian	67	0	Adenocarcinoma	N
50	M	Caucasian	60	100	Adenocarcinoma	Y
51	M	Caucasian	65	125	Carcinoma	N
52	M	Caucasian	52	80	Squamous cell	N
53	M	Caucasian	62	90	Squamous cell	N
54	F	Caucasian	69	140	Large cell	N
55	F	Caucasian	61	40	Squamous cell	N
56	F	Caucasian	80	15	Squamous cell	N
57	M	Caucasian	73	60	Squamous cell	Y
58	F	Caucasian	57	120	Adenocarcinoma	N
59	F	Caucasian	56	15	Adenocarcinoma	N
60	M	Caucasian	43	40	Adenocarcinoma	Y
61	F	Asian	47	10	Adenocarcinoma	N
62	F	Caucasian	39	2 nd hand	BAC	N
63	M	Caucasian	76	55	Adenocarcinoma	N
64	F	Caucasian	43	33	Adenocarcinoma	Y
65	F	Caucasian	50	30	Squamous cell	N
66	M	Caucasian	70	100	Squamous cell, Adenocarcinoma	N
67	M	Caucasian	70	56	Squamous cell	N
68	F	Caucasian	73	94	Adenocarcinoma	N
69	F	AA	62	40	Adenocarcinoma	N
70	M	Caucasian	58	80	Adenocarcinoma	N
71	M	Caucasian	71	55	Adenocarcinoma	N

Supplementary Table 1: Yale NSCLC patient characteristics.

Proportion male (M) (49.3%) and female (F) (50.7%), proportion Caucasian (84.5%), African American (AA)(8.6%), Hispanic (2.8%) and Asian (4.2%). For patients with multiple simultaneously diagnosed lung cancers both cancer types are listed. The presence or absence of the variant allele at the LCS6 SNP is denoted as Yes (Y) or No (N).

Variables	Controls	Cases	P-value
N	325	218	
Age	64.8 ± 9.0	65.1 ± 9.0	0.72 ^a
Sex (male, %)	68.9	73.9	0.22 ^b
Ethnicity (%)			0.01 ^b
White	67.4	75.7	
Hispanic	24.9	14.7	
Others	7.7	9.6	
Current smoking status (current smoker, %)	32.9	37.2	0.31 ^b
Pack-years	41.4 ± 28.5	56.9 ± 32.4	<0.0001 ^c
Family history of cancer (yes, %)	44.6	59.4	0.0008 ^b
Histology (%)			
Adenocarcinoma		45.9	
Squamous cell carcinoma		24.8	
Others ^d		29.4	

Supplementary Table 2A. Demographic New Mexico Lung Cancer Case-Control

Data. Cases indicate patients with lung cancer and controls are non-cancerous patients.

^aTwo-sided two-sample *t*-test between cases and controls. ^b χ^2 test for differences between cases and controls. ^cTwo-sided Wilcoxon rank sum test between cases and controls.

^dOthers included large cell lung cancer, poorly differentiated and other non-small cell lung cancer.

Supplementary Table 2B. Demographic characteristics among Boston NSCLC cases and controls

Characteristics	Overall			Male			Female		
	Cases (n=2205)	Controls (n=1497)	p	Cases (n=1118)	Controls (n=665)	p	Cases (n=1087)	Controls (n=832)	p
Age (mean±SD)	64.9±10.7	58.2±12.1	<0.01	65.8±10.5	60.2±12.8	<0.01	64.8±11.0	56.8±11.4	<0.01
Gender, N (%)									
Female	1087 (49.3%)	822 (55.6%)	<0.01						
Male	1118 (50.7%)	665 (44.4%)							
Smoking, N (%)									
Never	204 (9.3%)	522 (34.9%)	<0.01	77 (6.8%)	206 (30.8%)	<0.01	128 (11.7%)	324 (38.6%)	<0.01
Ex-smoker	1174 (53.2%)	688 (46.9%)		649 (57.6%)	362 (53.8%)		535 (48.8%)	333 (39.7%)	
Current smoker	827 (37.5%)	287 (19.2%)		401 (35.6%)	105 (15.6%)		434 (39.6%)	183 (21.8%)	
Years since quit (median) ^{a, b}	12 (1-59)	18 (1-65)	<0.01	14 (1-59)	20 (1-65)	<0.01	12 (1-55)	17 (1-59)	<0.01
Pack-years ^b	50 (0.1-231)	25 (0.1-218)		<0.01	58 (0.2-231)		29 (0.1-210)	<0.01	
Tumor stage (%)									
I and II	48.6%			49.0%			48.0%		
III and IV	51.3%			51.0%			52.0%		
Cell type (%)									
Adenocarcinoma	57.0%			50.6%			63.4%		
Squamous cell carcinoma	21.9%			28.1%			15.6%		
Others	21.1%			21.3%			21.0%		
Kraslet7 genotypes									
<i>TT</i>	1805 (81.9%)	1248 (83.4%)	0.32	944 (84.4%)	549 (82.6%)	0.46	861 (79.2%)	699 (84.0%)	0.02
<i>TG</i>	378 (17.4%)	231 (15.4%)		161 (14.4%)	105 (15.8%)		217 (20.0%)	126 (15.1%)	
<i>GG</i>	22 (1.0%)	18 (1.2%)		13 (1.2%)	11 (1.7%)		9 (0.8%)	7 (0.8%)	
<i>TG+GG</i>	400 (18.1%)	249 (16.6)	0.25*	174 (15.6%)	116 (17.4%)	0.32*	226 (20.6%)	133 (15.8%)*	0.008

^aEx-smokers only; ^bMedian (range), tested by non-parametric Wilcoxon's rank sum test; ^cContinuous variables tested with the Student's *t*-test; categorical variables tested using the χ^2 test.

* Compared with *TT* genotype, Fisher's exact test.

Supplementary Table 3. Primer Sequences Used

	Primer	Sequence (5'-3')
For sequencing the <i>KRAS</i> 3'UTR	SMJ104	CTAGCTAGCATACAATTTGTA CTTTTTTCTTAAGGCATAC
	LJC1	GGCACACCACCACCCCAAATCTC
	LJC2	CCATCTTCAGTGCCAGTCTTGGG
	LJC3	GGGTCGTATACCAAAGGCCTTAG
	LJC4	GCCTGAACTAGTTCACAGACAAGGG
	LJC5	CTAGCTAGCTCAATGCAGAATTCATGCTATCCAG
For sequencing only LCS6 and RFLP analysis	LJC21	GGTGT CAGAGTCTCGCTCTT
	LJC3	GGGTCGTATACCAAAGGCCTTAG
	LJC27*	CCTGAGTAGCTGGGATTACA
	LJC28*	GGATACCATATACCCAGTGCCTT
For sequencing the <i>NRAS</i> 3'UTR	LJC13	CCACTTTCAAGCTGCACTGACAC
	LJC8	CTAGCTGGAGTTACTGGTGCAATGAGC
	LJC9	GATACCTATGAGGATTTGGAGGC
	LJC10	GCATGGTAGCCTTCAGACAGAAC
	LJC11	CTGCTTCTTGTAATTCATCTCTGC
	LJC12	CAACTTAAAATATCGGCCCTTCC
For making KRAS wild-type	SMJ104	CTAGCTAGCATACAATTTGTA CTTTTTTCTTAAGGCATAC
	LJC5	CTAGCTAGCTCAATGCAGAATTCATGCTATCCAG
For making KRAS mLCS6	LJC16	CGAACTCCTGACCTCAAGTGATgCACCCACCTT
	LJC17	ATCACTTGAGGTCAGGAGTTCGAGACCAGCCT

* = used for nested PCR

		Primary Lung Tumor/Adjacent Non-tumor Tissue	Primary Lung Tumor	Non-primary Lung Tumor with or without Adjacent Non- tumor Tissue	Non-cancerous Lung
LCS1	4th bp = A	35	16	5	3
	4th bp = A/C	6 [†]	1	1 [!]	0
	25th bp = G	14	5	2	3
	rs712(-) = T/G	27 [§]	12	4	0
	Total Patients	41	17	6	3
LCS9	20th bp = T	37	15	6	1
	20th bp = T/C	3 [‡]	2	0	2
	15th bp = C	39	17	6	3
	15th bp = C/T	1 [!]	0	0	0
	Total Patients	40	17	6	3
LCS4	14th bp = C	40	18	6	3
	14th bp = C/T	1 [¶]	0	0	0
	Total Patients	41	18	6	3

Supplementary Table 4. Alterations in *KRAS* 3'UTR LCSs. All LCSs in which mutations were identified in are listed here, as are the number of NSCLC patients with these mutations. 68% of the primary lung tumors and 67% of the non-primary lung tumors examined were heterozygous at the rs712 locus in LCS1. A mutation at the fourth nucleotide of LCS1 was found in patients of both sexes and in a variety of non-small cell lung cancers types. The change at the twentieth nucleotide of LCS9 and was found in both sexes and in cancerous and non-cancerous lungs. The mutation at the twelfth base of LCS9 was seen in the adjacent tissue sample of a female, adenosquamous carcinoma patient. A mutation at the fifteenth base pair of LCS9 was found in the adjacent tissue of a female, squamous cell carcinoma patient, where as the primary lung tumor sample was normal. Lastly, there was one case of a mutation in LCS4. Both the tumor and adjacent tissue samples were heterozygous at this site. [†]Only tumor sample (3 patients), only adjacent sample (3 patients). [§]Tumor and adjacent samples (24 patients), only tumor sample (1 patient), only adjacent sample (2 patients). [‡]Only tumor sample (2 patients), tumor and adjacent samples (1 patient). [!] Only adjacent sample. [¶]Tumor and adjacent samples.