Integrative Analysis of genomic sequencing data reveals higher prevalence of LRP1B mutations in lung adenocarcinoma patients with COPD

Dakai Xiao[1,2,3#], Fuqiang Li[4,5#], Hui Pan[1,2], Han Liang[4,5], Kui Wu*[4,5,6], Jianxing He*[1,2,3]

1. Department of Thoracic Surgery, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China
2. Research Center for Translational Medicine, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China
3. Guangzhou Institute of Respiratory Disease &State Key Laboratory for Respiratory Disease, Guangzhou 510120, China
4. Cancer Institute, BGI-Research, BGI-Shenzhen, Shenzhen 518083, China
5. China National GeneBank-Shenzhen, BGI-Shenzhen, 518120, China
6. Department of Biology, University of Copenhagen, DK-2200, Copenhagen N, Denmark

[#]These authors contributed equally to this work

Corresponding authors:

Jianxing He, M.D, Ph.D

Department of Thoracic Surgery

The First Affiliated Hospital of Guangzhou Medical University

E-mail: Drjianxing.he@gmail.com

Kui Wu, Ph.D

BGI-Shenzhen

E-mail: wukui@genomics.cn

Supplementary Figure Legend

Fig. S1:

Somatic mutations and their association with COPD in LUAD patients stratified by smoking status in GMU cohort. Mutated genes and mutant frequencies in the primary tumor with and without COPD from smokers (left panel) and non-smokers (right panel) were shown. Genders, smoking status, tumor stage, survival and mutation types were also listed at the bottom according to the samples

Fig.S2

Recurrent Somatic mutations and their association with COPD in LUAD patients. Mutated genes and the mutant frequencies in the primary tumors with or without COPD from TCGA cohort were shown. Genders, smoking status, tumor stage, survival and mutation types were listed at the bottom according to the samples.

Fig.S1

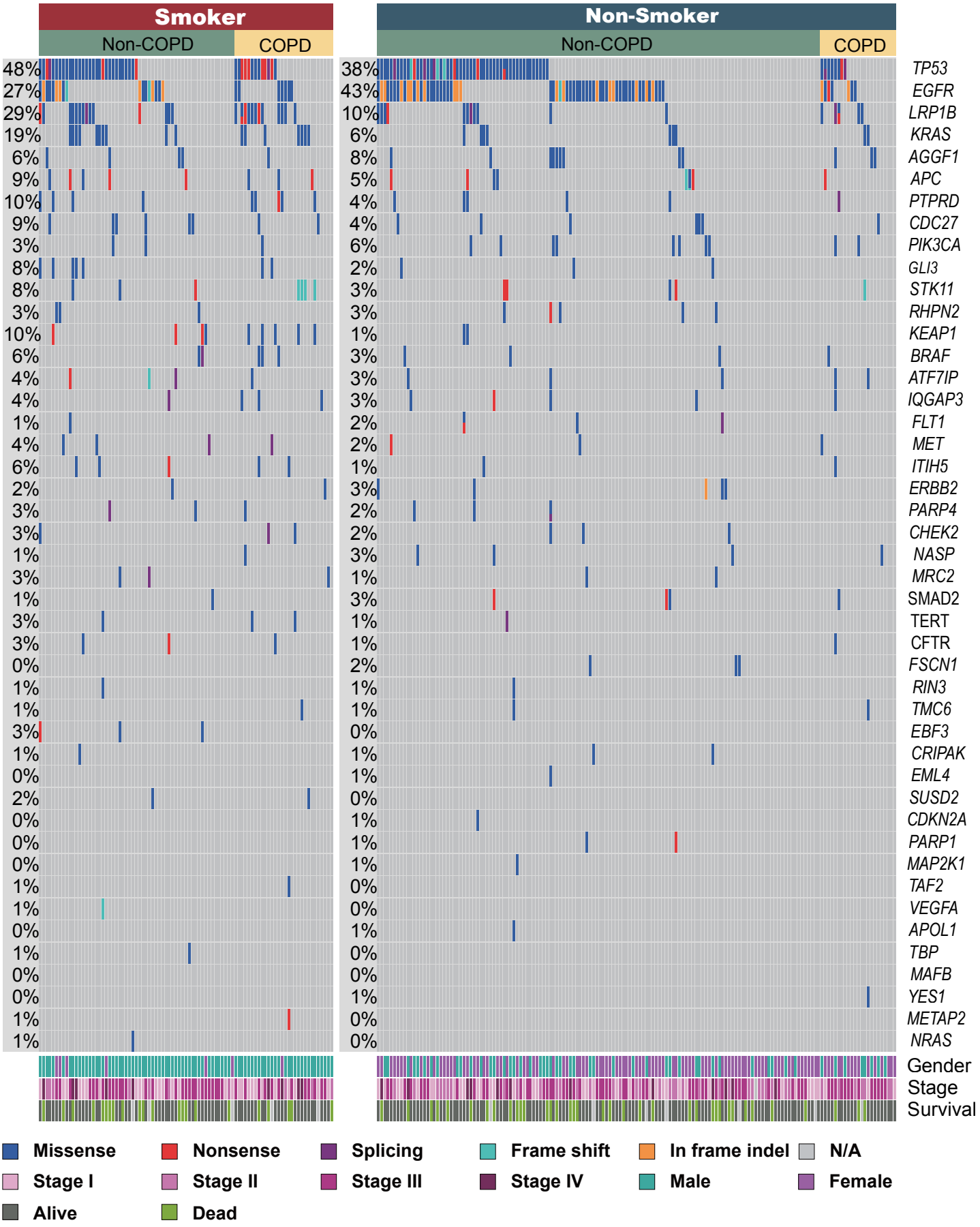

Smoker
Non-Smoker
Non-COPD
COPD
Non-COPD
COPD
48% TP53
27% EGFR
29% LRP1B
19% KRAS
6% AGGF1
9% APC
10% PTPRD
9% CDC27
3% PIK3CA
8% GLI3
8% STK11
3% RHPN2
10% KEAP1
6% BRAF
4% ATF7IP
4% IQGAP3
1% FLT1
4% MET
6% ITIH5
2% ERBB2
3% PARP4
3% CHEK2
1% NASP
3% MRC2
1% SMAD2
3% TERT
3% CFTR
0% FSCN1
1% RIN3
1% TMC6
3% EBF3
1% CRIPAK
0% EML4
2% SUSD2
0% CDKN2A
0% PARP1
0% MAP2K1
1% TAF2
1% VEGFA
0% APOL1
1% TBP
0% MAFB
0% YES1
1% METAP2
1% NRAS
38% TP53
43% EGFR
10% LRP1B
6% KRAS
8% AGGF1
5% APC
4% PTPRD
4% CDC27
6% PIK3CA
2% GLI3
3% STK11
3% RHPN2
1% KEAP1
3% BRAF
3% ATF7IP
3% IQGAP3
2% FLT1
2% MET
1% ITIH5
3% ERBB2
2% PARP4
2% CHEK2
3% NASP
1% MRC2
3% SMAD2
1% TERT
1% CFTR
2% FSCN1
1% RIN3
1% TMC6
0% EBF3
1% CRIPAK
1% EML4
0% SUSD2
1% CDKN2A
1% PARP1
1% MAP2K1
0% TAF2
0% VEGFA
1% APOL1
0% TBP
0% MAFB
1% YES1
0% METAP2
0% NRAS
Gender
Stage
Survival
Missense
Nonsense
Splicing
Frame shift
In frame indel
N/A
Stage I
Stage II
Stage III
Stage IV
Male
Female
Alive
Dead

Fig.S2

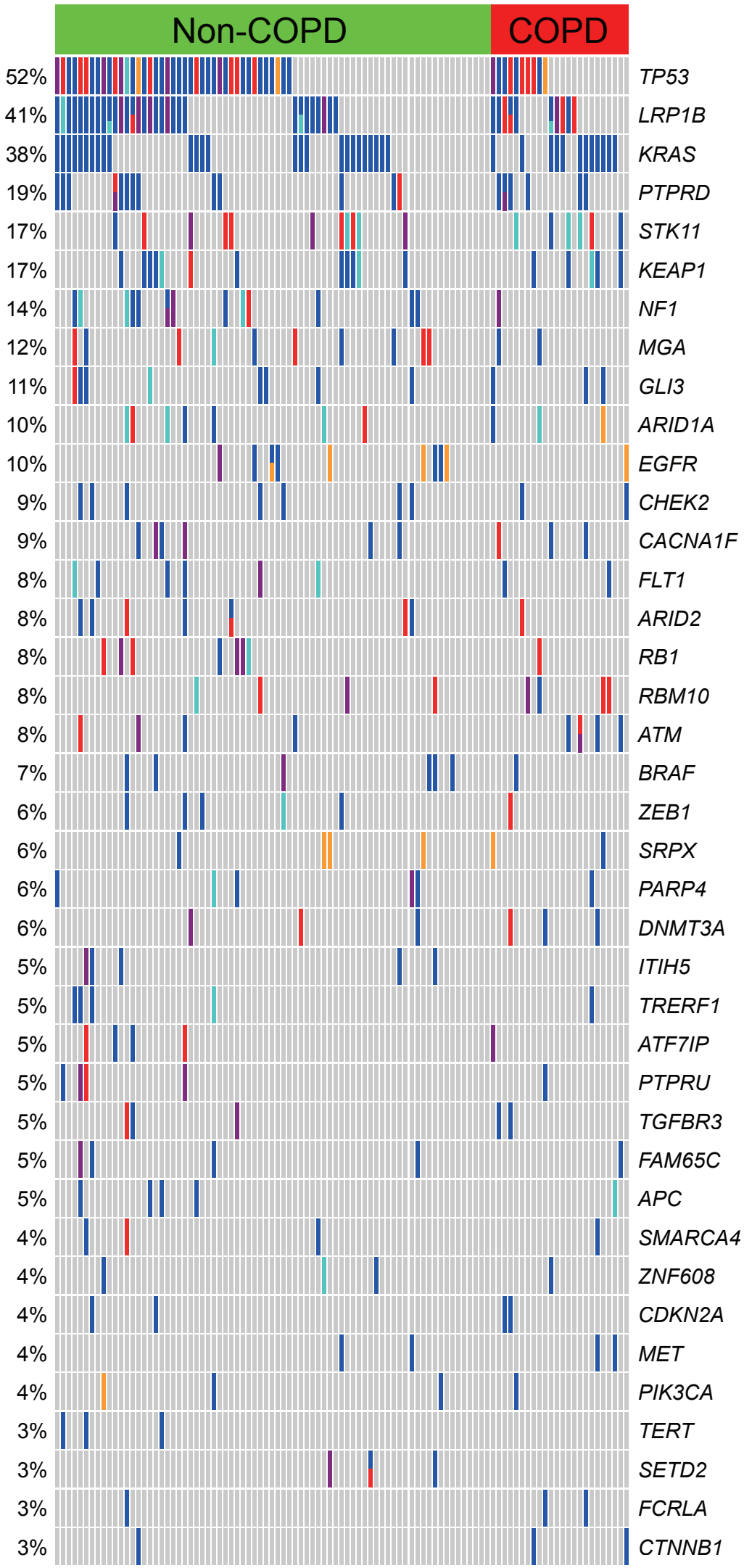
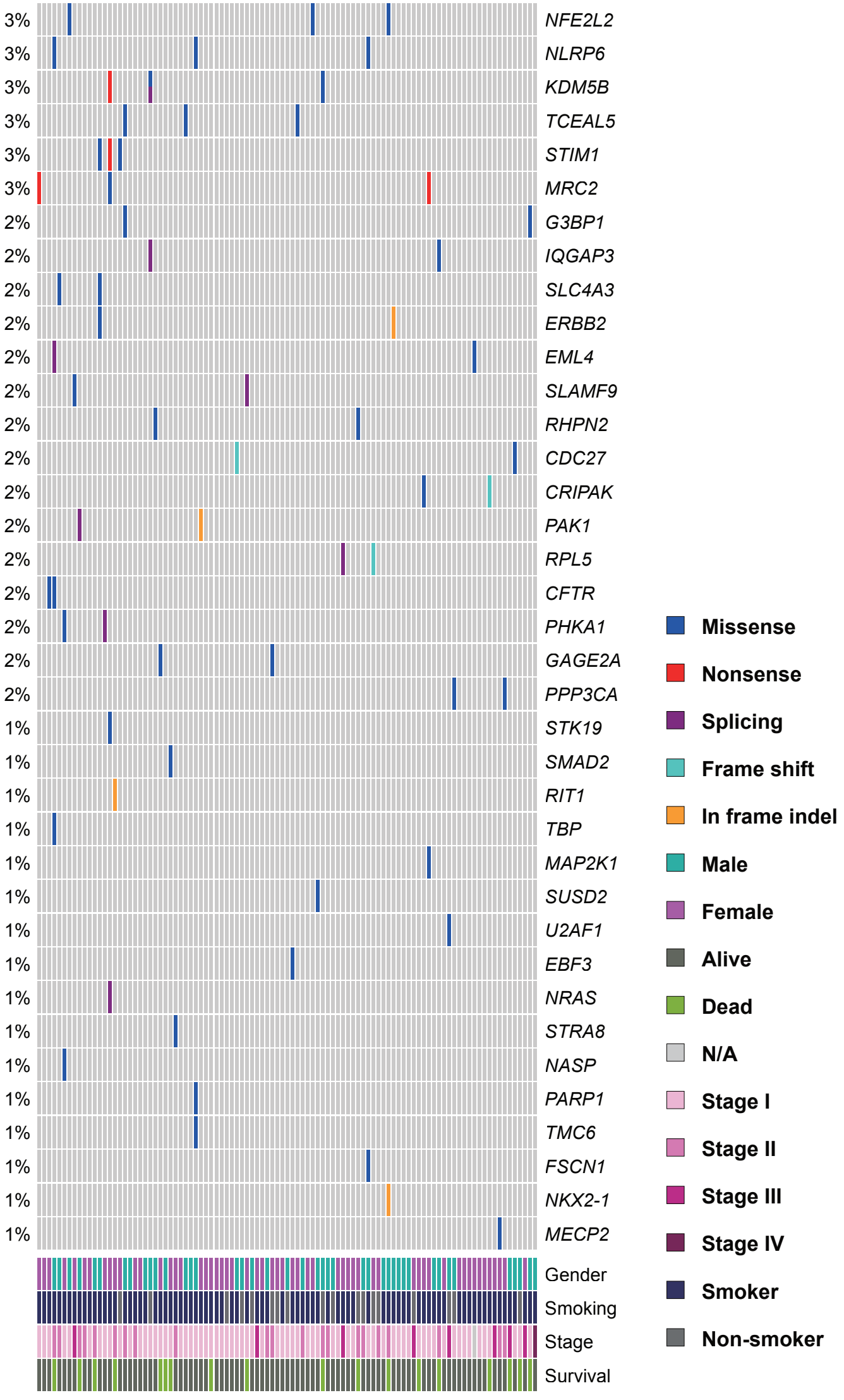

Non-COPD
COPD
52% TP53
41% LRP1B
38% KRAS
19% PTPRD
17% STK11
17% KEAP1
14% NF1
12% MGA
11% GLI3
10% ARID1A
10% EGFR
9% CHEK2
9% CACNA1F
8% FLT1
8% ARID2
8% RB1
8% RBM10
8% ATM
7% BRAF
6% ZEB1
6% SRPX
6% PARP4
6% DNMT3A
5% ITIH5
5% TRERF1
5% ATF7IP
5% PTPRU
5% TGFBR3
5% FAM65C
5% APC
4% SMARCA4
4% ZNF608
4% CDKN2A
4% MET
4% PIK3CA
3% TERT
3% SETD2
3% FCRLA
3% CTNNB1

Fig. S2
(Continued)



Supplementary Tables

Table S1

Univariate analysis with negative binomial regression comparing the counts of nonsense mutation and mutation in splicing site between COPD and non-COPD groups

	No.of patients	No.of Nonsense mutation median(range)	*p* value	No. of mutation in splicing site median(range)	*p* value
COPD			**0.0358**		**0.0371**
Yes	18	8(0-62)		2.5(0-28)	
No	67	4(0-65)		2(0-31)	

Table S2
Demographic and clinical information of LUAD patients from TCGA cohort

Variable	Overall (n=99)	COPD (n=24)	Non-COPD (n=75)	*P*-value
Age,mean±SD	66.66±9.49	68.83	65.96	0.134
Gender				0.75
Male	42	9	33	
Female	57	15	42	
Smoking				1
No	14	3	11	
Yes	85	21	64	
FEV1/FVC	80.24±21.62	53.13	88.92	<0.0001
Stage				0.057
Ⅰ	68	15	53	
Ⅱ	21	3	18	
Ⅲ	8	4	4	
Ⅳ	1	1	0	
NA	1			

Table S3

Univariate analysis with negative binomial regression comparing the counts of missense mutation between COPD and non-COPD groups in TCGA cohort

Variable	**No. of patients (n=99)**	**No.missense mutations median(range)**	Univariate analysis	Multivariate analysis
			p value	p value
Age，year			0.000649	0.0012
≤65	42	242.5(2-3738)		
>65	57	135(10-1073)		
Gender			0.254	0.4516
Male	42	195(10-3738)		
Female	57	203 (2-1819)		
Smoking			0.0169	0.0121
Yes	85	215(10-3738)		
No	9	191(45-240)		
NA	5	-		
Stage				
Ⅰ	68	206(10-3738)	(reference）	(reference）
Ⅱ	21	247(15-1819)	0.328	0.6198
Ⅲ	8	124.5(2-334)	0.036	0.4821
Ⅳ	1	51	0.074	0.0922
NA	1	-		
COPD			0.133	0.7166
Yes	24	195(51-903)		
No	75	203(2-3738)		

Table S4

Univariable analysis of factors associated with overall survival of LUAD patients

Variable	HR,95%CI	*P*-value
Age(year),≤65 vs >65	0.70(0.43-1.15)	0.158
Gender,male vs female	1.80(1.17-2.78)	**0.0078**
Smoking,yes vs no	0.97(0.61-1.56)	0.911
COPD,yes vs no	1.02(0.59-1.78)	0.946
Stage(Ⅰ vs Ⅱ vs Ⅲ vs Ⅳ)	-	**0.0014**
WBC,elevated vs normal	1.25(0.78-2.00)	0.356
Neutrophil, elevated vs normal	4.57(0.75-27.78)	0.099
Eosnophil, elevated vs normal	1.55(0.89-2.69)	0.122
Basophil, elevated vs normal	0.36(0.03-4.64)	0.434
Monocyte, elevated vs normal	0.60(0.22-1.69)	0.335
Lymphocyte, elevated vs normal	0.76(0.14-4.16)	0.75
LMR, High vs Low	0.57(0.36-0.91)	**0.018**
CEA, elevated vs normal	1.01(0.65-1.56)	0.975
CA125, elevated vs normal	2.16(1.07-4.34)	**0.031**
CA153, elevated vs normal	3.45(1.72-6.90)	0.00047