

Supplementary Methods

Transgenic mice were generated as previously described.[6] Mice were maintained according to protocols approved by the Institutional Animal Care and Use Committee at the Cincinnati Children's Hospital Medical Center. Mice were housed in a pathogen-free barrier facility in humidity and temperature-controlled rooms on a 12:12 h light/dark cycle, allowed food and water *ad libitum*.

Figure S1.

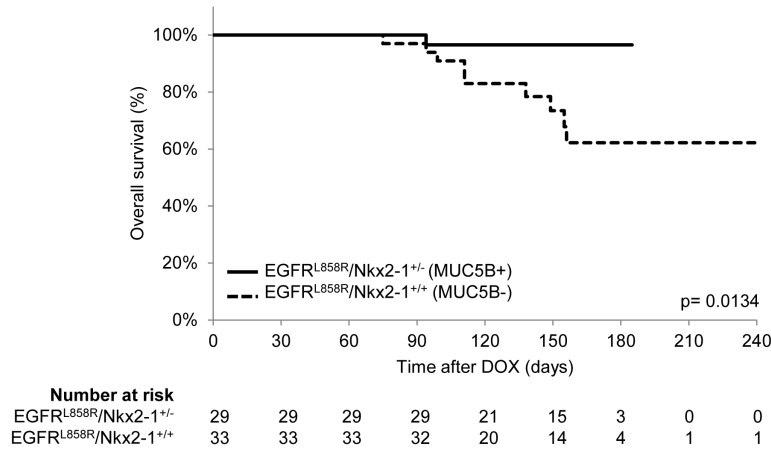


Figure S1. Kaplan-Meier analysis of overall survival in *EGFR*-mutant NSCLC mice

EGFR^{L858R}/Nkx2-1^{+/-} mice whose lung tumors express MUC5B significantly survived longer than EGFR^{L858R}/Nkx2-1^{+/+} mice whose lung tumors lack MUC5B (p=0.0134). DOX (doxycycline) administration induces mutant *EGFR* (EGFR^{L858R}) in lung epithelium.

Table S1. Association with Clinicopathological Data and the Expression of MUC5B of Patients with EGFR Wild Type NSCLC

Parameters	MUC5B			P-value
	Total (n=81)	Negative (-) (n=52)	Positive (+) (n=29)	
Median age (range, years)	67.7(32-90)	67.1(32-90)	68.8(41-79)	0.2954
Gender				
Male	68(84%)	45(56%)	23(28%)	
Female	13(16%)	7(9%)	6(7%)	0.5934
Smoking status				
Non-smoker	12(15%)	8(10%)	4(5%)	
Smoker	69(85%)	44(54%)	25(31%)	0.8467
Histological type				
Adenocarcinoma: Bronchoalveolar	8(10%)	5(6%)	3(4%)	
Adenocarcinoma	49(60%)	24(30%)	25(31%)	
Squamous cell carcinoma	24(30%)	23(28%)	1(1%)	0.0013
Median tumor size (range, mm)	37.8(10-120)	36.2(10-120)	40.7(14-80)	0.1661
Degree of differentiation				
Well	20(25%)	13(16%)	7(9%)	
Moderately	42(52%)	24(30%)	18(22%)	
Poorly	17(21%)	14(17%)	3(4%)	0.186
MD	2(2%)	1(1%)	1(1%)	
Stage				
IA/IB	35(43%)	26(32%)	9(11%)	
IIA/IIB	26(32%)	12(15%)	14(17%)	
IIIA/IIIB	20(25%)	14(17%)	6(7%)	0.0631
Tumor status				
T1-2	68(84%)	42(52%)	26(32%)	
T3-4	13(16%)	10(12%)	3(4%)	0.3599
Nodal status				
N0	50(62%)	33(41%)	17(21%)	
N1-3	31(38%)	19(23%)	12(15%)	0.6674
Lymphatic invasion				
Negative	21(26%)	13(16%)	8(10%)	
Positive	60(74%)	39(48%)	21(26%)	0.1165
Venous invasion				

Negative	31(38%)	19(23%)	12(15%)	
Positive	50(62%)	33(41%)	17(21%)	0.6674
Adjuvant chemotherapy				
Yes	52(64%)	33(41%)	19(23%)	
No	29(36%)	19(23%)	10(12%)	0.7057

Definition of abbreviations: NSCLC = non-small cell lung cancer; MD = missing data.

* Data are median (range) or number (%) unless otherwise stated.

Table S2. Univariate Analysis for OS and RFS in EGFR Wild Type NSCLC Patients

	OS	RFS
Age	0.6803	0.84
Gender (male vs female)	0.1194	0.4276
Smoking (no vs yes)	0.041	0.0744
MUC5B (positive vs negative)	0.9754	0.4624
Tumor size	0.9915	0.7123
Histological type (adeno vs sq)*	0.4381	0.2378
Differentiation (well vs moderately vs poorly)	0.0597	0.0722
Stage I vs Stage II/III	0.0346	0.2792
Lymphatic invasion (negative vs positive)	0.1568	0.0284
Venous invasion (negative vs positive)	0.0341	0.0338
Adjuvant chemotherapy (yes or no)	0.026	0.1639

Definition of abbreviations: NSCLC = non-small cell lung cancer; OS = overall survival; RFS = relapse-free survival; adeno = adenocarcinoma; sq = squamous cell carcinoma.

*Data are p-values by Kaplan-Meier analysis.

Table S3. Association of MUC5B with Prognosis in Multiple Cancers

	Author	Year	Species	Organ	Sample size (n)	Objective	Analysis Method	Result (MUC5B Positive)
1	Yu CJ ⁹⁾	1996	human	lung	60	To elucidate the clinical significance of mucin gene overexpression in lung cancer	Slot-blot analysis and immunohistochemistry in surgical specimens of NSCLC	Associated with relapse (p= 0.0015) and lower DFS (p= 0.0037)
2	Pinto-de-Sousa J ¹⁰⁾	2004	human	gastric	50	To elucidate the clinical significance of mucin gene overexpression in gastric cancer	Immunohistochemistry in surgical specimens of gastric carcinomas	No significance (p= 0.59)
3	Varangot M ¹¹⁾	2005	human	breast	80	To evaluate the prognostic value of MUC5B mRNA expression in bone marrow aspirates	Multimarker RT-PCR assay in pre-operative bone marrow aspirates	Unexpected favorable clinical outcome.
4	Partheen K ¹²⁾	2006	human	ovarian	54	In order to find novel candidate biomarkers	Microarray (with hierarchical cluster analysis) and quantitative RT-PCR assay	A hierarchical sub-group that included 60% of the survivors shows higher mRNA expression (p< 0.001)
5	Valque H ¹³⁾	2012	mouse	breast	22	To understand better the implication of MUC5B in cancer pathogenesis	Histological and immunological analysis	Correlate with poor survival with no significance (p= 0.08)
6	Nagashio R ¹⁴⁾	2015	human	lung	247	To evaluate the relationships between MUC5B expression in tumor cells and the clinicopathological parameters of ACs	Immunohistochemistry in surgical specimens of NSCLC	Significantly associated with poorer survival (p= 0.017)

Definition of abbreviations: NSCLC = non-small cell lung cancer; DFS = disease-free survival; RT-PCR = reverse transcription-polymerase chain reaction; OS = overall survival; RFS = relapse-free survival; HR = hazard ratio; ACs = adenocarcinomas.