BMJ Open

MUC5B is a Favorable Prognostic Marker for EGFR Mutant Non-Small Cell Lung Cancer

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2015-008366
Article Type:	Research
Date Submitted by the Author:	30-Mar-2015
Complete List of Authors:	Wakata, Kouki; Nagasaki University, Surgical Oncology Tsuchiya, Tomoshi; Nagasaki University, Surgical Oncology Tomoshige, Koichi; Nagasaki University, Surgical Oncology Takagi, Katsunori; Nagasaki University, Surgical Oncology Yamasaki, Naoya; Nagasaki University, Surgical Oncology Matsumoto, Keitaro; Nagasaki University, Surgical Oncology Miyazaki, Takuro; Nagasaki University, Surgical Oncology Nanashima, Atsushi; Nagasaki University, Surgical Oncology Whitsett, Jeffrey; CCHMC, Pulmonary Biology Maeda, Yutaka; CCHMC, Pulmonary Biology Nagayasu, Takeshi; Nagasaki University, Surgical Oncology
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Pathology, Respiratory medicine
Keywords:	Pathology < BASIC SCIENCES, Respiratory tract tumours < THORACIC MEDICINE, Surgical pathology < PATHOLOGY

MUC5B is a Favorable Prognostic Marker for *EGFR* Mutant Non-Small Cell Lung Cancer

Kouki Wakata[1], Tomoshi Tsuchiya[1], Koichi Tomoshige[1], Katsunori Takagi[1], Naoya Yamasaki[1], Keitaro Matsumoto[1], Takuro Miyazaki[1], Atsushi Nanashima[1], Jeffrey A. Whitsett[2], Yutaka Maeda[2], Takeshi Nagayasu[1]

Correspondence and requests for reprints should be addressed to (Yutaka Maeda and Takeshi Nagayasu are co-corresponding authors):

Correspondence to

Yutaka Maeda, D.V.M, Ph.D.

Perinatal Institute, Division of Neonatology, Perinatal and Pulmonary Biology

Cincinnati Children's Hospital Medical Center and the University of Cincinnati College of Medicine

Cincinnati, Ohio 45229, USA

Tel: +1-513-636-3322

Fax: +1-513-636-7868

E-mail: yutaka.maeda@cchmc.org

Takeshi Nagayasu, M.D., Ph.D.

Division of Surgical Oncology

Nagasaki University Graduate School of Biomedical Sciences

1-7-1 Sakamoto Nagasaki-city, Nagasaki 852-8501, Japan

Tel: +81-95-819-7304

Fax: +81-95-819-7306

E-mail: nagayasu@nagasaki-u.ac.jp

Key words: MUC5B; MUC5AC; epidermal growth factor receptor; NSCLC; mutation

Word count: 1967

Key messages

What is the key question?

Is there a specific prognostic biomarker for patients with surgically resected NSCLC that harbors an *EGFR* mutation?

What is the bottom line?

We identified MUC5B, a mucous protein, as a favorable prognostic marker specific for patients with *EGFR* mutant NSCLC.

Why read on?

MUC5B may serve as a biomarker that could be used to direct the therapeutic strategy for treating *EGFR* mutant NSCLC.

Abstract

Background: Genetic screening of pharmacologically treatable driver mutations, including epidermal growth factor receptor (*EGFR*) mutations, has become common in diagnosis for non-small cell lung cancer (NSCLC). Prognostic markers associated with each driver mutation detected in resected tumors will be useful in designing subsequent therapies.

Objectives: To determine the use of the mucin proteins MUC5B and MUC5AC as prognosis markers for NSCLCs carrying *EGFR* mutations.

Methods: Expression of MUC5B and MUC5AC were evaluated by immunohistochemical (IHC) analysis in 159 patients with NSCLC who underwent surgical resection (*EGFR* mutant type: n=78, *EGFR* wild type: n=81). The association of MUC5B or MUC5AC expression with clinicopathological characteristics and postoperative survival rate was analyzed according to *EGFR* mutation status.

Results: Patients whose tumors expressed MUC5B had significantly longer overall survival (OS) and relapse-free survival (RFS) compared to the MUC5B negative patients with *EGFR* mutant NSCLC (p=0.0098, p=0.0188, respectively). In patients with *EGFR* wild type NSCLC, there was no association with MUC5B expression. MUC5AC expression was not different between *EGFR* mutant and wild type NSCLC.

Conclusions: Present findings indicate that MUC5B, but not MUC5AC, is a novel prognostic biomarker for patients with NSCLC carrying *EGFR* mutations but not for patients with NSCLC carrying wild type *EGFR*.

Strengths and limitations of this study

1. A prognostic marker for each driver mutation in NSCLC has not yet been determined.

2. MUC5B is a favorable postoperative prognostic marker for *EGFR* mutant NSCLC.

3. MUC5AC is not correlated with postoperative prognosis regardless of *EGFR* mutation status.

4. The function of MUC5B in *EGFR* mutant NSCLC remains unknown.

Introduction

Lung cancer is the primary cause of cancer-related death in the United States and worldwide.[1] Non-small cell lung cancer (NSCLC) accounts for approximately 80-85% of all lung cancers.[1] Currently, targeted therapies for non-resectable NSCLC have progressed rapidly, based on the discovery of pharmacologically treatable driver mutations in epidermal growth factor receptor (*EGFR*) and fusions of anaplastic lymphoma kinase (ALK).[2,3] These moleculary targeted therapies have revealed distinct and/or overlapping tumorigenic pathways associated with each driver mutation, especially regarding the mechanisms of tumor recurrence.[4] Genetic screening of driver mutations, including *EGFR* mutations and *ALK* fusions, is now common for metastatic NSCLC but not for surigically resected primary NSCLC.[5] However, many patients suffer recurrence despite surgery.[6] Considering the current decreased costs of genetic screening and increasing understanding of molecular mechanisms associated with such "druggable" driver mutations, identification of such mutations and associated prognosis factors from resected primary NSCLC will help establish feasible schemes for therapies prior to or to limit tumor recurrence.

Recently, we reported that decreased expression of *Nkx2-1* (also known as TTF-1) in a mouse model of *EGFR* mutant NSCLC reduced the number and size of lung tumors [7] and extended the survival of the mice (see online supplementary figure S1). Unexpectedly, the decreased *Nkx2-1* induced the expression of a mucin protein MUC5B but not MUC5AC in *EGFR* mutant lung tumors in the mice,[7] suggesting that MUC5B may serve as a favorable prognostic marker associated with *EGFR* mutant NSCLC in humans. In the present study, we demonstrate that patients whose primarily resected *EGFR* mutant lung tumors express MUC5B survived significantly longer than those whose primarily resected *EGFR* mutant lung tumors did not express MUC5B. MUC5B was not associated with survival in *EGFR* wild type NSCLC. Our

finding provides a novel approach to assess prognosis for patients whose primarily resected lung tumors carry *EGFR* mutations.

Methods

Study Population

Among the patients who underwent surgical resection at Nagasaki University Hospital and related facilities between June 1996 and March 2013, patients who were tested for the presence or absence of *EGFR* mutations were selected for this study. We further selected the patients whose clinicopathological characteristics were retrieved from the patients' charts and whose prognosis was followed at our institution and related facilities. We enrolled 159 patients (*EGFR* mutant type: n=78, *EGFR* wild type: n=81) for this study (table 1). All investigations were approved by our institution and related facilities review boards, and informed consent was obtained from all participants prior to the study.

Clinicopathological Evaluation

Histological classification of NSCLC was designated as to three types – well, moderately and poorly differentiated – based on the predominant features according to the World Health Organization classification.[8] The patients remained for a median follow-up period of 1680 days, ranging from 55 to 4503 days. For all patients, periodic inspection with chest x-ray, CT scan and tumor marker assays was performed every several months to confirm the presence or absence of recurrence, even if patients experienced no complaints or no symptoms.

Antibody Information

For immunohistochemical staining, primary antibodies were used at the following concentrations: rabbit polyclonal anti-MUC5B (1:200; sc-20119, Santa Cruz Biotechnology), rabbit polyclonal anti-MUC5AC (1:50; sc-20118, Santa Cruz Biotechnology).

Sample Preparation, Selection and Immunohistochemistry

The 5-μm thick formalin-fixed, paraffin-embedded (FFPE) lung sections were deparaffinized in dimethylbenzene and dehydrated through a graded alcohol series. For antigen retrieval, the FFPE lung sections were incubated in 10 mM citric acid (pH 6.0) at 121°C for 15 min and then washed in phosphate-buffered saline (PBS). Next, the lung sections were immersed in 3% H_2O_2 solution for 30 min to block the endogenous peroxidase followed by incubation with each primary antibody at 4 °C overnight. After washing in PBS, the lung sections were incubated with the peroxidase-conjugated secondary antibodies (Simple Stain MAX-PO kit, Nichirei, Tokyo, Japan) for 30 min at room temperature. For IHC staining, the lung sections were visualized with a diaminobenzidine (DAB: brown) kit (Histofine, Nichirei) and counterstained with hematoxylin. The lung sections visualized with DAB were dehydrated with alcohol and dimethylbenzene and mounted in a conventional fashion.

Normal bronchial tissue specimens that moderately expressed MUC5B were prepared as positive controls in all cases. Normal gastric mucosa tissue specimens that moderately expressed MUC5AC were prepared as positive controls in all cases. Negative controls were also prepared in all cases. MUC5B and MUC5AC staining was evaluated by immunohistochemistry by two independent trained observers (K.W. and T.T.) using the following criteria: score 0, no staining; score 1, weak staining; score 2, distinct staining; score 3, very strong staining. Score 0 and 1 were further categorized as negative, and score 2 and 3 as positive.

Statistical Analysis

For univariate analysis, categorical data were analyzed by the chi-square test or Fisher's exact test or the Cochran-Armitage test. Continuous data were expressed as a mean using the Mann-Whitney U test or the Kruskal-Wallis test. The overall survival (OS) and relapse-free survival

(RFS) were calculated according to the Kaplan-Meier method, and differences between groups were tested for significance using the log-rank test. Subjects who neither died nor had recurrence were censored at the time of their last follow-up. The prognostic relevance of a single factor was determined by multivariate Cox regression analysis. A p-value of 0.05 or less was considered significant. SPSS version 17 software (SPSS Japan, Tokyo, Japan) was used for the analysis.

Results

Expression of MUC5B and MUC5AC in human NSCLC

We previously reported that a mouse model of *EGFR* mutant NSCLC in *Nkx2-1* heterozygous background ($EGFR^{L858R}$; $Nkx2\text{-}1^{+/-}$) expressed MUC5B (but not MUC5AC) in lung tumors while those in *Nkx2-1* wild type background ($EGFR^{L858R}$; $Nkx2\text{-}1^{+/+}$) did not express MUC5B.[7] The $EGFR^{L858R}$; $Nkx2\text{-}1^{+/-}$ mice survived significantly longer than the $EGFR^{L858R}$; $Nkx2\text{-}1^{+/+}$ mice (p= 0.0134; see online supplementary figure S1), suggesting that MUC5B is associated with a favorable prognosis in *EGFR* mutant NSCLC. Since MUC5B is an abundant cytoplasmic and secreted protein, we assessed whether MUC5B could be used as a prognostic marker for patients with NSCLC carrying *EGFR* mutations in primary resected human lung tumors. Primary resected NSCLC tumors were tested immunohistochemically for the presence of MUC5B. MUC5B staining was detected in the cytoplasm of NSCLC cells in 27 of the 78 samples with *EGFR* mutations and 29 of the 81 samples with wild type *EGFR* (figure 1A,B). The NSCLC samples were also tested using MUC5AC antibody, detecting expression of MUC5AC in the cytoplasm of NSCLC cells in 20 of the 73 samples with *EGFR* mutations and 24 of the 79 samples with wild type *EGFR* (figure 1C,D). These results indicate that both MUC5B and MUC5AC are expressed in a portion of human NSCLC.

Prognostic Association of MUC5B or MUC5AC with *EGFR* mutant or *EGFR* wild type NSCLC

Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying *EGFR* mutations or wild type *EGFR* were assessed. In a cohort of patients whose resected NSCLC tumors carrying *EGFR* mutations, patients whose tumors expressed MUC5B (MUC5B positive patients) survived significantly longer than patients whose tumors do not express MUC5B

(MUC5B negative patients) in both OS (p= 0.0098; figure 2A) and RFS (p= 0.0188; figure 2B). In a cohort of patients whose resected NSCLC tumors with wild type *EGFR*, there was no significant difference between the MUC5B positive patients and MUC5B negative patients in OS and RFS (figure 2C,D). Expression of MUC5AC in NSCLC was not associated with OS and RFS regardless of *EGFR* mutation status (figure 3). The expression of MUC5B in NSCLC tumors carrying *EGFR* mutations was not correlated with clinicopathological parameters, including age, gender, smoking status, histological types, tumor size, degree of differentiation, stage, tumor status, nodal status, lymphatic invasion and venous invasion (table 2). Univariate analysis showed significant differences ($p < 0.05$) in OS and RFS in expression of MUC5B, tumor size, degree of differentiation, stage, lymphatic invasion and venous invasion (table 3). Multivariate Cox regression analysis using the variables that were $p < 0.05$ in univariate analysis showed that the expression of MUC5B was independently associated with better OS and RFS ($p < 0.05$; table 4). These results indicate that MUC5B is a favorable prognostic marker for postoperative patients whose resected NSCLC tumors carry *EGFR* mutation but not those with wild type *EGFR*.

Discussion

In the present study, we demonstrate that expression of MUC5B in primary *EGFR* mutant NSCLC is associated with longer survival in patients with NSCLC. MUC5B, but not MUC5AC, is a favorable prognostic biomarker for NSCLC in humans carrying *EGFR* mutations.

MUC5B has been assessed as a prognostic biomarker for multiple cancers in several studies, using RT-PCR, microarray analysis and immunohistochemistry (see online supplementary table S1 and S2).[9–13] mRNA data assessing MUC5B as a prognostic biomarker was obtained from PrognoScan, a database for meta-analysis of the prognostic value of genes using microarray data deposited to the public domain (see online supplementary table S2).[14] The prognostic impact of MUC5B expression differed among cancer types. In lung cancer, six microarray studies analyzed by PrognoScan did not indicate MUC5B as either a good or a poor prognostic biomarker.[15–20] Immunohistochemical analysis indicated MUC5B as a poor prognosis biomarker (see online supplementary table S1),[9] a finding contradicting our present study. Previous mRNA microarray and immunostaining were based on all NSCLCs independent of driver mutation-based classification, which differs from our analysis that was based on classification by *EGFR* mutations. The utility of MUC5B as a prognostic factor differed in the two breast cancer studies, depending on the molecular basis of the tumors. In all breast cancers, PrognoScan indicated that MUC5B was associated with poor prognosis;[21,22] however, in ER (Estrogen Receptor)-positive breast cancers, MUC5B was associated with favorable prognosis,[23] indicating the potential importance of tumor classification on a molecular basis. In the present study, we assessed MUC5B as a biomarker for NSCLC based on *EGFR* mutation status rather than on all NSCLCs, identifying MUC5B as a favorable prognosis biomarker for *EGFR* mutant NSCLC.

Regulation of *MUC5B* in *EGFR* mutant NSCLC is not well understood. *MUC5AC* and *MUC5B* genes are closely located at a locus on human chromosome 11. Both are evolutionally

conserved gel-forming mucins secreted from airway epithelial cells in the lung. In normal lung, MUC5B is constitutively expressed at higher levels than MUC5AC.[24] In asthma and other inflammatory lung diseases, MUC5AC is highly induced in airway goblet cells.[24] In idiopathic pulmonary fibrosis (IPF), MUC5B but not MUC5AC is highly expressed in the airway goblet cells.[25,26] The SNP rs35705950 located at the *MUC5B* promoter is associated with induction of *MUC5B* mRNA in IPF;[26] however, we detected the SNP rs35705950 in only one of 27 cases in the *EGFR* mutant NSCLC expressing MUC5B (data not shown), indicating that the SNP is not associated with increased MUC5B in *EGFR* mutant NSCLC. Since MUC5B was induced in *EGFR* mutant lung tumors in *Nkx2-1* heterozygous mice ($EGFR^{L858R}$; $Nkx2\text{-}1^{+/-}$), *Muc5b* is suppressed by NKX2-1 in *EGFR* mutant NSCLC in mice.[7] Regulatory mechanisms controlling the *MUC5B* gene are not understood in *EGFR* mutant NSCLC in human. The function of MUC5B in cancer has been analyzed using a truncated MUC5B in MCF7 breast cancer cells, truncated MUC5B promoting tumorigenesis of MCF7 cells.[13] However, the use of truncated MUC5B may obscure the intrinsic role of full-length MUC5B since there is a possibility that the truncated MUC5B may function in a dominant-negative fashion. In lung, MUC5B is required for mucociliary clearance and innate immunity against bacterial infection.[27] The potential functions of MUC5B in lung cancer, including *EGFR* mutant lung cancer, are not known. The present study suggests that MUC5B or processes regulating *MUC5B* may influence the growth and metastasis of *EGFR* mutant NSCLC. MUC5B may serve as a surrogate biomarker influenced by a pathway involved in metastasis and recurrence associated with *EGFR* mutant NSCLC.

In conclusion, our data revealed the clinicopathological significance of MUC5B as a favorable prognostic factor in resected *EGFR* mutant NSCLC. Further studies are necessary to elucidate the gene regulatory mechanism and the function of MUC5B in *EGFR* mutant NSCLC.

Author affiliations:

[1]Division of Surgical Oncology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

[2]Perinatal Institute, Division of Neonatology, Perinatal and Pulmonary Biology, Cincinnati Children's Hospital Medical Center and the University of Cincinnati College of Medicine, Cincinnati, Ohio, USA

Acknowledgments: We thank the patients, their families, all of the investigators who participated in the study, central laboratories that did the *EGFR* mutation testing, and Dr. Mary Durbin for discussions.

Contributors: Conception and design, K.W., T.T., J.A.W., Y.M., T.N.; provision of study material, patients, and data acquisition, K.W., T.T., K.T., K.T., N.Y., K.M., T.M., A.N., J.A.W., Y.M.; data analysis and interpretation, K.W., T.T., K.T., K.M., J.A.W., Y.M., T.N.; drafting manuscript and intellectual content, K.W., T.T., J.A.W., Y.M., T.N.

Funding: This work was supported by Ministry of Education, Science, and Culture of Japan, National Institute of Health, American Lung Association, the University of Cincinnati Postdoctoral Fellow Research Program, and Cincinnati Children's Hospital Medical Center.

Competing interests: None

Patient consent: Obtained

Ethics approval: Ethics Committee from every participant hospital.

Provenance and peer review: Not commissioned; externally peer reviewed.

Data Sharing Statement: No additional data available.

References

1 Siegel R, Naishadham D JA. Cancer statistics for hispanics/latinos, 2012. *CA Cancer J Clin* 2012;**62**:283–98.

2 Lynch T, Bell D. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non–small-cell lung cancer to gefitinib. *N Engl J Med* 2004;**350**:2129–39.

3 Paez JG, Jänne P a, Lee JC, *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004;**304**:1497–500.

4 Camidge DR, Pao W, Sequist L V. Acquired resistance to TKIs in solid tumours: learning from lung cancer. *Nat Rev Clin Oncol* 2014;**104**:[Epub ahead of print].

5 Shames DS, Wistuba II. The evolving genomic classification of lung cancer. *J Pathol* 2014;**232**:121–33.

6 Chang WJ, Sun J-M, Lee JY, *et al.* A retrospective comparison of adjuvant chemotherapeutic regimens for non-small cell lung cancer (NSCLC): Paclitaxel plus carboplatin versus vinorelbine plus cisplatin. *Lung Cancer* 2014;**84**:51–5.

7 Maeda Y, Tsuchiya T, Hao H, *et al.* KrasG12D and Nkx2-1 haploinsufficiency induce mucinous adenocarcinoma of the lung. *J Clin Invest* 2012;**122**:4388－400.

8 Travis WD, Brambilla E, Noguchi M, *et al.* International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: international multidisciplinary classification of lung adenocarcinoma: executive summary. *Proc Am Thorac Soc* 2011;**8**:381–5.

9 Yu CJ, Yang PC, Shun CT, *et al.* Overexpression of MUC5 genes is associated with early post-operative metastasis in non-small-cell lung cancer. *Int J Cancer* 1996;**69**:457–65.

10 Pinto-de-Sousa J, Reis C a, David L, *et al.* MUC5B expression in gastric carcinoma: relationship with clinico-pathological parameters and with expression of mucins MUC1, MUC2, MUC5AC and MUC6. *Virchows Arch* 2004;**444**:224–30.

11 Varangot M, Barrios E, Sóñora C, *et al.* Clinical evaluation of a panel of mRNA markers in the detection of disseminated tumor cells in patients with operable breast cancer. *Oncol reports* 2005;**14**:537–45.

12 Partheen K, Levan K, Osterberg L, *et al.* Expression analysis of stage III serous ovarian adenocarcinoma distinguishes a sub-group of survivors. *Eur J Cancer* 2006;**42**:2846–54.

13 Valque H, Gouyer V, Gottrand F, *et al.* MUC5B leads to aggressive behavior of breast cancer MCF7 cells. *PLoS One* 2012;**7**:e46699.

14 Mizuno H, Kitada K, Nakai K, *et al.* PrognoScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med Genomics* 2009;**2**:18.

15 Shedden K, Taylor JMG, Enkemann S a, *et al.* Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008;**14**:822–7.

16 Okayama H, Kohno T, Ishii Y, *et al.* Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res* 2012;**72**:100–11.

17 Bild AH, Yao G, Chang JT, *et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;**439**:353–7.

18 Zhu C-Q, Ding K, Strumpf D, *et al.* Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. *J Clin Oncol* 2010;**28**:4417–24.

19 Lee E-S, Son D-S, Kim S-H, *et al.* Prediction of recurrence-free survival in postoperative non-small cell lung cancer patients by using an integrated model of clinical information and gene expression. *Clin Cancer Res* 2008;**14**:7397–404.

20 Raponi M, Zhang Y, Yu J, *et al.* Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res* 2006;**66**:7466–72.

21 Pawitan Y, Bjöhle J, Amler L, *et al.* Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005;**7**:R953–64.

22 Miller LD, Smeds J, George J, *et al.* An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A* 2005;**102**:13550–5.

23 Chanrion M, Negre V, Fontaine H, *et al.* A gene expression signature that can predict the recurrence of tamoxifen-treated primary breast cancer. *Clin Cancer Res* 2008;**14**:1744–52.

24 Evans C, Kim K, Tuvim M, *et al.* Mucus hypersecretion in asthma: causes and effects. *Curr Opin Pulm Med* 2009;**15**:4–11. doi:10.1097/MCP.0b013e32831da8d3.Mucus

25 Plantier L, Crestani B, Wert SE, *et al.* Ectopic respiratory epithelial cell differentiation in bronchiolised distal airspaces in idiopathic pulmonary fibrosis. *Thorax* 2011;**66**:651–7.

26 Seibold M, Wise A, Speer M, *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;**364**:1503–12.

27 Roy MG, Livraghi-Butrico A, Fletcher A a, *et al.* Muc5b is required for airway defence. *Nature* 2014;**505**:412–6.

Legends

Table 1: Baseline characteristics of the 159 patients with NSCLC

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

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

†Other mean NSCLC neuroendocrine.

Table 2: Association with clinicopathological data and the expression of MUC5B of patients with *EGFR* mutant NSCLC

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

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

Table 3: Univariate analysis for OS and RFS in *EGFR* mutant NSCLC patients

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 4: Multivariate analysis for OS and RFS in *EGFR* mutant NSCLC patients

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

Figure 1: Immunohistochemical staining for MUC5B and MUC5AC expression in NSCLC

Representative images of immune-positive staining for MUC5B in NSCLC (A) and negative staining (B), and positive staining for MUC5AC (C) and negative staining (D).

Figure 2: Survival curves for patients based on the expression of MUC5B in *EGFR* mutant or wild type NSCLC

Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying *EGFR* mutations or *EGFR* wild type. OS (A) and RFS (B) in the patients with *EGFR* mutant type NSCLC, and OS (C) and RFS (D) in the patients with EGFR wild type NSCLC.

Figure 3: Survival curves for patients based on the expression of MUC5AC in *EGFR* mutant or wild type NSCLC

Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying *EGFR* mutations or *EGFR* wild type. OS (A) and RFS (B) in the patients with *EGFR* mutant type NSCLC, and OS (C) and RFS (D) in the patients with *EGFR* wild type NSCLC.

Table 1. Baseline characteristics of the 159 patients with NSCLC

Number of patients	
Median age (range, years)	67.5(32-90)
Gender	
Male	103(65%)
Female	56(35%)
Smoking status	
Non-smoker	54(34%)
Smoker	105(66%)
Histological type	
Adenocarcinoma: Bronchoalveolar	33(21%)
Adenocarcinoma	91(57%)
Squamous cell carcinoma	32(20%)
Adenosquamous carcinoma	2(1%)
Other※	1(1%)
Median tumor size (range, mm)	32.4(8-120)
Degree of differentiation	
Well	55(35%)
Moderately	69(43%)
Poorly	28(18%)
MD	7(4%)
Stage	
IA/IB	81(51%)
IIA/IIB	40(25%)
IIIA/IIIB	38(24%)
Tumor status	
T1-2	136(86%)
T3-4	23(14%)
Nodal status	
N0	103(65%)
N1-3	56(35%)
Lymphatic invasion	

Negative	56(35%)
Positive	102(64%)
MD	1(1%)
Venous invasion	
Negative	76(48%)
Positive	82(52%)
MD	1(1%)
EGFR	
Wild type	81(51%)
Mutant type	78(49%)

Table 2. Association with clinicopathological data and the expression of MUC5B of patients with EGFR mutant NSCLC

Parameters	Total (n=78)	MUC5B Negative (-) (n=51)	MUC5B Positive (+) (n=27)	P-value
Median age (range, years)	66.9(41-85)	66.7(42-83)	67.3(41-85)	.674
Gender				
Male	35(45%)	22(43%)	13(48%)	
Female	43(55%)	29(57%)	14(52%)	.6721
Smoking status				
Non-smoker	42(54%)	28(55%)	14(52%)	
Smoker	36(46%)	23(45%)	13(48%)	.7971
Histological type				
Adenocarcinoma: Bronchoalveolar	27(35%)	16(31%)	11(41%)	
Adenocarcinoma	43(55%)	29(57%)	14(52%)	
Squamous cell carcinoma	8(10%)	6(12%)	2(7%)	.6522
Median tumor size (range, mm)	25.7(8-60)	26.1(8-60)	24.9(8-50)	.8771
Degree of differentiation				
Well	35(45%)	22(43%)	13(48%)	
Moderately	28(36%)	19(37%)	9(33%)	
Poorly	11(14%)	6(12%)	5(19%)	.7348
MD	4(5%)	4(8%)	0(0%)	
Stage				
IA/IB	46(59%)	29(57%)	17(63%)	
IIA/IIB	14(18%)	8(16%)	6(22%)	
IIIA/IIIB	18(23%)	14(27%)	4(15%)	.4162
Tumor status				
T1-2	68(87%)	42(82%)	26(96%)	
T3-4	10(13%)	9(18%)	1(4%)	.0797
Nodal status				
N0	53(68%)	35(69%)	18(67%)	
N1-3	25(32%)	16(31%)	9(33%)	.8599

Lymphatic invasion				
Negative	35(45%)	26(51%)	9(33%)	
Positive	42(54%)	24(47%)	18(67%)	.1165
MD	1(1%)	1(2%)	0(0%)	
Venous invasion				
Negative	45(58%)	30(59%)	15(56%)	
Positive	32(41%)	20(39%)	12(44%)	.7057
MD	1(1%)	1(2%)	0(0%)	

Table 3. Univariate analysis for OS and RFS in EGFR mutant NSCLC patients

	OS	RFS
Age	.7311	.1207
Gender (male vs female)	.3166	.813
Smoking (no vs yes)	.9754	.8508
MUC5B (positive vs negative)	.0098	.0187
Tumor size	.0058	.0001
Histological type (adeno vs sq)※	.0245	.6369
Differentiation (well vs moderately vs poorly)	.0172	.0137
Stage I vs Stage II/III	.0095	.0308
Lymphatic invasion (negative vs positive)	.0075	.0011
Venous invasion (negative vs positive)	.0045	.0021

Table 4. Multivariate analysis for OS and RFS in EGFR mutant NSCLC patients

～OS～

Independent variables	HR	95% CI	P-value
MUC5B (positive vs negative)	15.378	1.617 to 146.293	.0174
Tumor size	2.38	0.188 to 30.162	.5033
Histological type (adeno vs sq)※	1.706	0.376 to 7.737	.4885
Differentiation (well vs moderately vs poorly)	2.441	0.592 to 10.069	.2171
Stage I vs Stage II/III	2.112	0.529 to 8.433	.29
Lymphatic invasion (negative vs positive)	6.531	0.499 to 85.521	.1528
Venous invasion (negative vs positive)	0.39	0.066 to 2.305	.2986

～RFS～

Independent variables	HR	95% CI	P-value
MUC5B (positive vs negative)	2.783	1.0989 to 7.0483	.0309
Tumor size	2.8609	0.7700 to 10.6293	.1165
Histological type (adeno vs sq)※	0.6208	0.1633 to 2.3602	.4841
Differentiation (well vs moderately vs poorly)	0.7168	0.3095 to 1.6603	.4373
Stage I vs Stage II/III	1.2296	0.4435 to 3.4087	.6912
Lymphatic invasion (negative vs positive)	7.2608	1.6535 to 31.8834	.0086
Venous invasion (negative vs positive)	0.8025	0.2628 to 2.4500	.6992

Figure 1

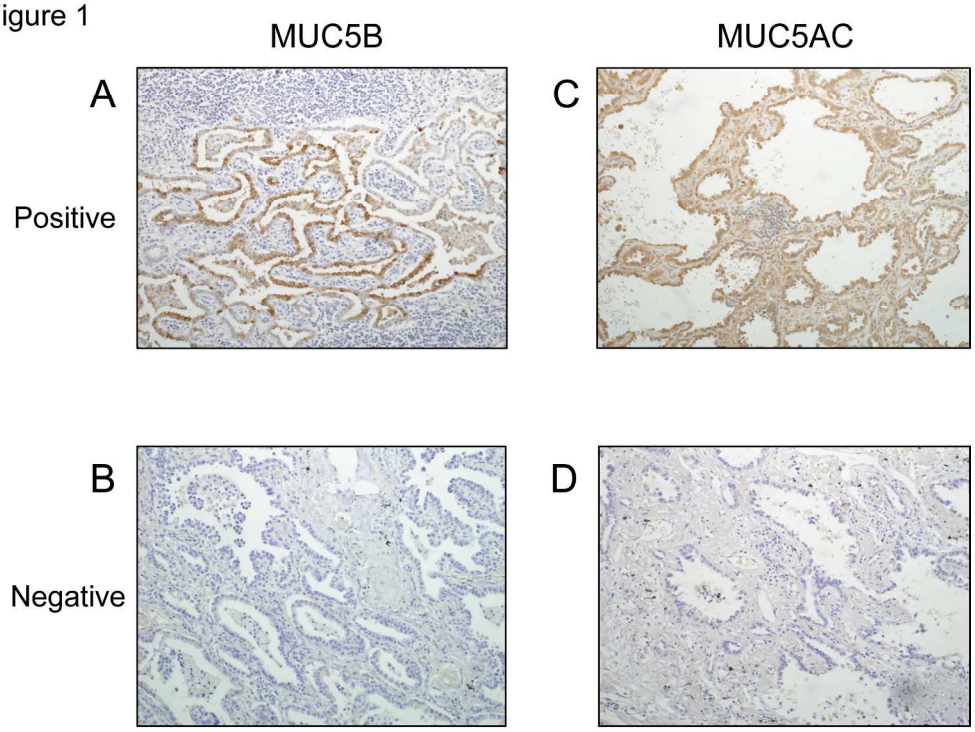


Figure 1: Immunohistochemical staining for MUC5B and MUC5AC expression in NSCLC
Representative images of immune-positive staining for MUC5B in NSCLC (A) and negative staining (B), and positive staining for MUC5AC (C) and negative staining (D).

216x160mm (300 x 300 DPI)

Figure 2

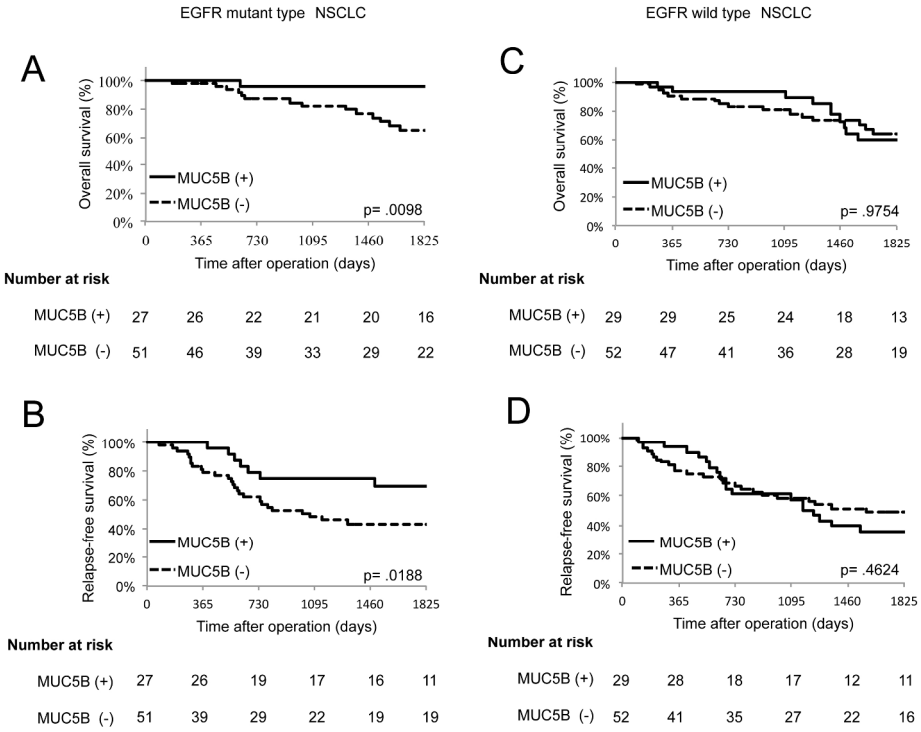


Figure 2: Survival curves for patients based on the expression of MUC5B in EGFR mutant or wild type NSCLC
Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying EGFR mutations or EGFR wild type. OS (A) and RFS (B) in the patients with EGFR mutant type NSCLC, and OS (C) and RFS (D) in the patients with EGFR wild type NSCLC.

234x184mm (300 x 300 DPI)

Figure 3

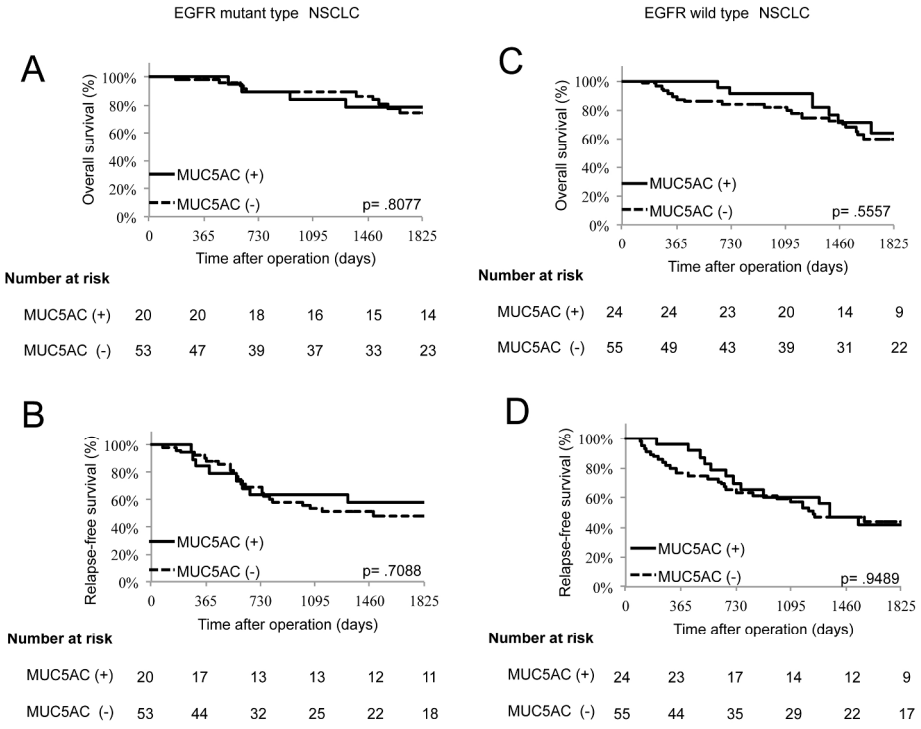


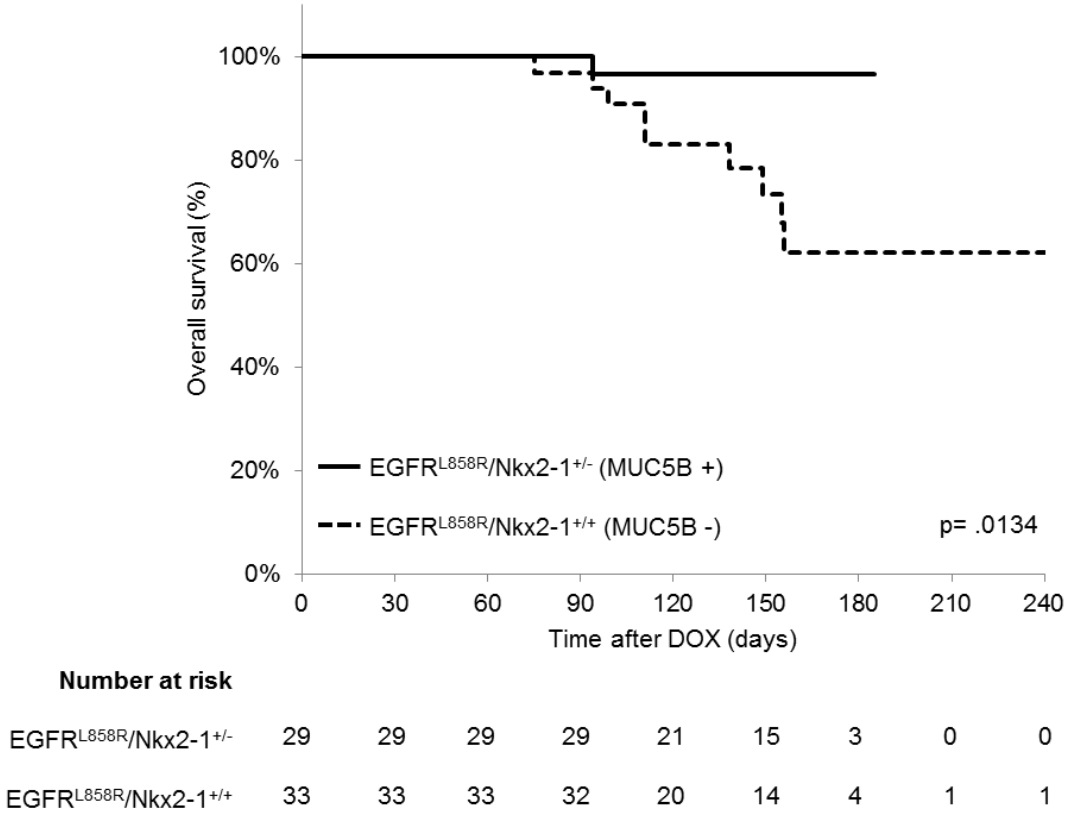
Figure 3: Survival curves for patients based on the expression of MUC5AC in EGFR mutant or wild type NSCLC

Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying EGFR mutations or EGFR wild type. OS (A) and RFS (B) in the patients with EGFR mutant type NSCLC, and OS (C) and RFS (D) in the patients with EGFR wild type NSCLC.

234x183mm (300 x 300 DPI)

Supplementary Methods

Transgenic mice were generated as previously described (7). 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*.



Number at risk

EGFR^{L858R}/Nkx2-1$^{+/-}$	29	29	29	29	21	15	3	0	0
EGFR^{L858R}/Nkx2-1$^{+/+}$	33	33	33	32	20	14	4	1	1

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= .0134). DOX (doxycycline) administration induces mutant *EGFR* (EGFR^{L858R}) in lung epithelium.

	Auther	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 repalse (p= .0015) and lower DFS (p= .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= .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< .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= .08)

Table S1: Association of MUC5B with prognosis in multiple cancers

NSCLC, non-small cell lung carcinoma; RT-PCR, reverse transcription-polymerase chain reaction; OS, overall survival; RFS, relapse-free survival; HR, hazard ratio.

Supplemental Table 2

<PrognoScan for MUC5B>

TYPE	ID_NAME	ID_DESCRIPTION	STPRC	DATASET	CANCER TYPE	SUBTYPE	ENDPOINT	COHORT	NTRIBUTC	ARRAY TYPE	PROBE ID	N	CUTPOINT	MINIMUM P-VALUE	CORRECTED P-VALUE	n(HR-high / HR-low	COX P-VALUE	ln(HR)	HR [95% CI-low CI-upp]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE5287	Bladder cancer		Overall Survival	Aarhus (1995-2004)	Als	HG-U133A	222268_x_at	30	0.166667	0.00250255	0.0543902	-1.67361	0.912781	-0.075511	0.93 [0.24 - 3.58]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE5287	Bladder cancer		Overall Survival	Aarhus (1995-2004)	Als	HG-U133A	213432_at	30	0.266667	0.154764	-	-0.652913	0.928131	-0.0949102	0.91 [0.12 - 7.15]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12417-GPL96	Blood cancer	AML	Overall Survival	AMLCG (1999-2003)	Metzeler	HG-U133A	222268_x_at	163	0.269939	0.147765	-	-0.310628	0.815868	-0.123671	0.88 [0.31 - 2.50]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12417-GPL96	Blood cancer	AML	Overall Survival	AMLCG (1999-2003)	Metzeler	HG-U133A	213432_at	163	0.846626	0.0134761	0.20158	-0.772452	0.0807591	-1.05141	0.35 [0.11 - 1.14]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12417-GPL570	Blood cancer	AML	Overall Survival	AMLCG (2004)	Metzeler	HG-U133_Plus_2	222268_x_at	79	0.240506	0.0041369	0.0814729	-0.878166	0.0559411	-1.33476	0.26 [0.07 - 1.03]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12417-GPL570	Blood cancer	AML	Overall Survival	AMLCG (2004)	Metzeler	HG-U133_Plus_2	213432_at	79	0.303797	0.0396709	0.425428	0.721236	0.808768	0.178555	1.20 [0.28 - 5.08]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE5122	Blood cancer	AML	Overall Survival	San Diego	Raponi	HG-U133A	222268_x_at	58	0.844828	0.073097	-	-0.699604	0.761854	-0.0347262	0.97 [0.77 - 1.21]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE5122	Blood cancer	AML	Overall Survival	San Diego	Raponi	HG-U133A	213432_at	58	0.37931	0.00195769	0.0445074	0.898877	0.0837673	0.34647	1.41 [0.95 - 2.09]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE8970	Blood cancer	AML	Overall Survival	San Diego	Raponi	HG-U133A	213432_at	34	0.470588	0.0778624	-	-0.665935	0.215135	-0.561057	0.57 [0.23 - 1.39]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE8970	Blood cancer	AML	Overall Survival	San Diego	Raponi	HG-U133A	222268_x_at	34	0.117647	0.0195136	0.263321	-1.27067	0.778421	-0.0611595	0.94 [0.61 - 1.44]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4475	Blood cancer	B-cell lymphoma	Overall Survival	Berlin (2003-2005)	Hummel	HG-U133A	222268_x_at	158	0.563291	0.000206207	0.00660067	0.84874	0.0261186	1.71926	5.58 [1.23 - 25.38]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4475	Blood cancer	B-cell lymphoma	Overall Survival	Berlin (2003-2005)	Hummel	HG-U133A	213432_at	158	0.803797	0.0445561	0.457845	-0.700569	0.728414	-0.318065	0.73 [0.12 - 4.38]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-346	Blood cancer	DLBCL	Overall Survival	GELA (1998-2000)	Jais	HG-U133A	213432_at	53	0.188679	0.155633	-	0.746811	0.869466	-0.0522197	0.95 [0.51 - 1.77]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-346	Blood cancer	DLBCL	Overall Survival	GELA (1998-2000)	Jais	HG-U133A	222268_x_at	53	0.886792	0.0314389	0.365517	0.950504	0.591159	0.129599	1.14 [0.71 - 1.83]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-346	Blood cancer	DLBCL	Event Free Survival	GELA (1998-2000)	Jais	HG-U133A	222268_x_at	53	0.886792	0.0179368	0.248003	1.034	0.713404	0.0829959	1.09 [0.70 - 1.69]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-346	Blood cancer	DLBCL	Event Free Survival	GELA (1998-2000)	Jais	HG-U133A	213432_at	53	0.735849	0.328071	-	-0.393293	0.478951	-0.215387	0.81 [0.44 - 1.46]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE16131-GPL96	Blood cancer	Follicular lymphoma	Overall Survival	NCI (1974-2001)	Dave	HG-U133A	222268_x_at	180	0.861111	0.120407	-	-0.568635	0.913684	-0.00989761	0.99 [0.83 - 1.18]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE16131-GPL96	Blood cancer	Follicular lymphoma	Overall Survival	NCI (1974-2001)	Dave	HG-U133A	213432_at	180	0.638889	0.176436	-	0.288403	0.534471	0.14136	1.15 [0.74 - 1.80]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2658	Blood cancer	Multiple myeloma	Disease Specific Survival	Arkansas	Zhan	HG-U133_Plus_2	222268_x_at	559	0.118068	0.00228932	0.0505861	-0.761384	0.00318225	-0.471227	0.62 [0.46 - 0.85]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2658	Blood cancer	Multiple myeloma	Disease Specific Survival	Arkansas	Zhan	HG-U133_Plus_2	213432_at	559	0.277281	0.00649211	0.116058	-0.561447	0.0214074	-0.476102	0.62 [0.41 - 0.93]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4271-GPL96	Brain cancer	Astrocytoma	Overall Survival	MDA	Phillips	HG-U133A	213432_at	77	0.181818	0.00822113	0.139178	-0.833979	0.0428764	-0.806423	0.45 [0.20 - 0.97]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4271-GPL96	Brain cancer	Astrocytoma	Overall Survival	MDA	Phillips	HG-U133A	222268_x_at	77	0.792208	0.248133	-	-0.384295	0.612941	-0.0829592	0.92 [0.67 - 1.27]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7696	Brain cancer	Glioblastoma	Overall Survival	Lausanne	Murat	HG-U133_Plus_2	222268_x_at	70	0.842857	0.0534195	-	-0.765317	0.817762	-0.128974	0.88 [0.29 - 2.63]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7696	Brain cancer	Glioblastoma	Overall Survival	Lausanne	Murat	HG-U133_Plus_2	213432_at	70	0.285714	0.0234037	0.299092	-0.637708	0.605879	-0.366261	0.69 [0.17 - 2.79]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	MGH-glioma	Brain cancer	Glioma	Overall Survival	CBTTB, MGH, BWH, CH	Nutt	HG-U95A	41365_at	50	0.86	0.133028	-	-0.735177	0.374952	-1.3273	0.27 [0.01 - 4.98]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4412-GPL96	Brain cancer	Glioma	Overall Survival	UCLA (1996-2003)	Freije	HG-U133A	213432_at	74	0.743243	0.0309414	0.361654	0.667304	0.18728	0.603886	1.83 [0.75 - 4.49]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4412-GPL96	Brain cancer	Glioma	Overall Survival	UCLA (1996-2003)	Freije	HG-U133A	222268_x_at	74	0.797297	0.0572263	-	-0.808815	0.891354	-0.0253437	0.97 [0.68 - 1.40]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE16581	Brain cancer	Meningioma	Overall Survival	UCLA	Lee	HG-U133_Plus_2	222268_x_at	67	0.58209	0.0825011	-	-1.66179	0.409086	-2.17636	0.11 [0.00 - 19.91]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE16581	Brain cancer	Meningioma	Overall Survival	UCLA	Lee	HG-U133_Plus_2	213432_at	67	0.283582	0.0831341	-	15.4754	0.267456	1.52689	4.60 [0.31 - 68.42]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE19615	Breast cancer		Distant Metastasis Free Survival	DF/HCC	Li	HG-U133_Plus_2	222268_x_at	115	0.608696	0.0340859	0.385577	1.12341	0.226588	0.596698	1.82 [0.69 - 4.78]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE19615	Breast cancer		Distant Metastasis Free Survival	DF/HCC	Li	HG-U133_Plus_2	213432_at	115	0.826087	0.0747692	-	-15.3292	0.618168	-0.0828371	0.92 [0.66 - 1.27]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE3143	Breast cancer		Overall Survival	Duke	Bild	HG-U95A	41365_at	158	0.348101	0.00462066	0.0889345	-0.784772	0.0703306	-0.246794	0.78 [0.60 - 1.02]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7849	Breast cancer		Disease Free Survival	Duke (1990-2001)	Anders	HG-U95A	41365_at	76	0.894737	0.0209488	0.276837	1.26611	0.10966	0.30777	1.36 [0.93 - 1.98]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12276	Breast cancer		Relapse Free Survival	EMC	Bos	HG-U133_Plus_2	222268_x_at	204	0.102941	0.0375123	0.410401	-0.485774	0.24141	-0.0649973	0.94 [0.84 - 1.04]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12276	Breast cancer		Relapse Free Survival	EMC	Bos	HG-U133_Plus_2	213432_at	204	0.372549	0.072348	-	0.264676	0.394158	0.0300675	1.03 [0.96 - 1.10]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE6532-GPL570	Breast cancer		Distant Metastasis Free Survival	GUYT	Loi	HG-U133_Plus_2	213432_at	87	0.321839	0.0501144	-	-0.738472	0.366257	-0.17488	0.84 [0.57 - 1.23]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE6532-GPL570	Breast cancer		Relapse Free Survival	GUYT	Loi	HG-U133_Plus_2	222268_x_at	87	0.229885	0.00632575	0.113744	2.26527	0.134784	0.816368	2.26 [0.78 - 6.59]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE6532-GPL570	Breast cancer		Relapse Free Survival	GUYT	Loi	HG-U133_Plus_2	213432_at	87	0.321839	0.0501144	-	-0.738472	0.366257	-0.17488	0.84 [0.57 - 1.23]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE6532-GPL570	Breast cancer		Distant Metastasis Free Survival	GUYT	Loi	HG-U133_Plus_2	222268_x_at	87	0.229885	0.00632575	0.113744	2.26527	0.134784	0.816368	2.26 [0.78 - 6.59]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE9195	Breast cancer		Distant Metastasis Free Survival	GUYT2	Loi	HG-U133_Plus_2	213432_at	77	0.168831	0.0131742	0.19825	-1.46961	0.140125	-1.55818	0.21 [0.03 - 1.67]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE9195	Breast cancer		Relapse Free Survival	GUYT2	Loi	HG-U133_Plus_2	222268_x_at	77	0.181818	0.194541	-	-0.761545	0.846	0.215856	1.24 [0.14 - 10.96]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE9195	Breast cancer		Relapse Free Survival	GUYT2	Loi	HG-U133_Plus_2	213432_at	77	0.168831	0.0794856	-	-1.01232	0.202305	-0.936011	0.39 [0.09 - 1.65]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE9195	Breast cancer		Distant Metastasis Free Survival	GUYT2	Loi	HG-U133_Plus_2	222268_x_at	77	0.181818	0.0455006	0.463875	-1.21651	0.397719	-0.979289	0.38 [0.04 - 3.63]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12093	Breast cancer		Distant Metastasis Free Survival	IO, NCI, TUM, CCF (1992-2000)	Zhang	HG-U133A	222268_x_at	136	0.125	0.2694	-	1.08027	0.968682	0.00975977	1.01 [0.62 - 1.64]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12093	Breast cancer		Distant Metastasis Free Survival	IO, NCI, TUM, CCF (1992-2000)	Zhang	HG-U133A	213432_at	136	0.205882	0.125785	-	-0.744092	0.777402	-0.0366527	0.96 [0.75 - 1.24]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE11121	Breast cancer		Distant Metastasis Free Survival	Mainz (1988-1998)	Schmidt	HG-U133A	213432_at	200	0.63	0.0373341	0.409139	0.605812	0.833264	0.0175713	1.02 [0.86 - 1.20]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE11121	Breast cancer		Distant Metastasis Free Survival	Mainz (1988-1998)	Schmidt	HG-U133A	222268_x_at	200	0.71	0.0309603	0.361801	0.632659	0.684141	0.073379	1.08 [0.76 - 1.53]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE1378	Breast cancer		Relapse Free Survival	MGH (1987-2000)	Ma	Arcturus 22k	1549	60	0.116667	0.0835798	-	1.59027	0.85216	-0.135054	0.87 [0.21 - 3.62]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE1379	Breast cancer		Relapse Free Survival	MGH (1987-2000)	Ma	Arcturus 22k	1549	60	0.683333	0.0404414	0.430683	-0.972391	0.806123	-0.259746	0.77 [0.10 - 6.14]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE9893	Breast cancer		Overall Survival	Montpellier, Bordeaux, Turin (1989-2001)	Chanrion	Human 21K V12.0	9179	155	0.367742	3.97E-11	3.88E-09	-1.82272	0.000151417	-1.26278	0.28 [0.15 - 0.54]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2034	Breast cancer		Distant Metastasis Free Survival	Rotterdam (1980-1995)	Wang	HG-U133A	213432_at	286	0.188811	0.212447	-	-0.28869	0.785528	0.0164024	1.02 [0.90 - 1.14]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2034	Breast cancer		Distant Metastasis Free Survival	Rotterdam (1980-1995)	Wang	HG-U133A	222268_x_at	286	0.370629	0.0104142	0.166462	-0.492547	0.0879186	-0.190965	0.83 [0.66 - 1.03]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE1456-GPL96	Breast cancer		Disease Specific Survival	Stockholm (1994-1996)	Pawitan	HG-U133A	222268_x_at	159	0.779874	0.000169232	0.00555612	1.30435	0.0330242	0.570926	1.77 [1.05 - 2.99]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE1456-GPL96	Breast cancer		Relapse Free Survival	Stockholm (1994-1996)	Pawitan	HG-U133A	213432_at	159	0.716981	0.00415096	0.0816928	0.884271	0.0727154	0.141571	1.15 [0.99 - 1.34]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE1456-GPL96	Breast cancer		Overall Survival	Stockholm (1994-1996)	Pawitan	HG-U133A	222268_x_at	159	0.773585	0.000528658	0.0148607	1.0594	0.123804	0.348984	1.42 [0.91 - 2.21]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE1456-GPL96	Breast cancer		Disease Specific Survival	Stockholm (1994-1996)	Pawitan	HG-U133A	213432_at	159	0.72327	0.000960689	0.0246417	1.16167	0.0292251	0.197455	1.22 [1.02 - 1.45]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE1456-GPL96	Breast cancer		Relapse Free Survival	Stockholm (1994-1996)	Pawitan	HG-U133A	222268_x_at	159	0.779874	0.00121343	0.0299613	1.00412	0.166617	0.31606	1.37 [0.88 - 2.15]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE1456-GPL96	Breast cancer		Overall Survival	Stockholm (1994-1996)	Pawitan	HG-U133A	213432_at	159	0.72327	0.00363923	0.0735602	0.894998	0.0419133	0.160041	1.17 [1.01 - 1.37]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-158	Breast cancer		Relapse Free Survival	UCSF, CPMC (1989-1997)	Chin	HG-U133A	222268_x_at	117	0.222222	0.0919756	-	0.731198	0.533343	-0.244955	0.78 [0.36 - 1.69]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-158	Breast cancer		Disease Specific Survival	UCSF, CPMC (1989-1997)	Chin	HG-U133A	222268_x_at	117	0.230769	0.0603684	-	1.09309	0.636689	-0.218405	0.80 [0.32 - 1.99]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-158	Breast cancer		Relapse Free Survival	UCSF, CPMC (1989-1997)	Chin	HG-U133A	213432_at	117	0.153846	0.0395986	0.424932	-0.769592	0.352153	-0.105736	0.90 [0.72 - 1.12]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-158	Breast cancer		Disease Specific Survival	UCSF, CPMC (1989-1997)	Chin	HG-U133A	213432_at	117	0.111111	0.130354	-	-0.732362	0.653212	-0.057769	0.94 [0.73 - 1.21]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-158	Breast cancer		Overall Survival	UCSF, CPMC (1989-1997)	Chin	HG-U133A	222268_x_at	117	0.222222	0.0919756	-	0.731198	0.533343	-0.244955	0.78 [0.36 - 1.69]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-158	Breast cancer		Overall Survival	UCSF, CPMC (1989-1997)	Chin	HG-U133A	213432_at	117	0.153846	0.0395986	0.424932	-0.769592	0.352153	-0.105736	0.90 [0.72 - 1.12]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-158	Breast cancer		Distant Metastasis Free Survival	UCSF, CPMC (1989-1997)	Chin	HG-U133A	222268_x_at	117	0.230769	0.0290694	0.346845	1.47141	0.976801	-0.0131362	0.99 [0.41 - 2.39]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	E-TABM-158	Breast cancer		Distant Metastasis Free Survival	UCSF, CPMC (1989-1997)	Chin	HG-U133A	213432_at	117	0.478632	0.206757	-	-0.498966	0.66397	-0.0569218	0.94 [0.73 - 1.22]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE3494-GPL96	Breast cancer		Disease Specific Survival	Uppsala (1987-1989)	Miller	HG-U133A	222268_x_at	236	0.377119	0.0175716	0.244381	0.719954	0.131954	0.342612	1.41 [0.90 - 2.20]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE3494-GPL96	Breast cancer		Disease Specific Survival	Uppsala (1987-1989)	Miller	HG-U133A	213432_at	236	0.661017	0.0036049	0.0730048	0.768302	0.03738	0.141609	1.15 [1.01 - 1.32]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4922-GPL96	Breast cancer		Disease Free Survival	Uppsala (1987-1989)	Ivshina	HG-U133A	222268_x_at	249	0.116466	0.0725866	-	0.742208	0.236414	0.202393	1.22 [0.88 - 1.71]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4922-GPL96	Breast cancer		Disease Free Survival	Uppsala (1987-1989)	Ivshina	HG-U133A	213432_at	249	0.136546	0.0552101	-	0.738082	0.690211	0.0242737	1.02 [0.91 - 1.15]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2990	Breast cancer		Relapse Free Survival	Uppsala, Oxford	Sotiriou	HG-U133A	222268_x_at	62	0.225806	0.140862	-	1.40305	0.720892	-0.0900179	0.91 [0.56 - 1.50]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2990	Breast cancer		Relapse Free Survival	Uppsala, Oxford	Sotiriou	HG-U133A	213432_at	125	0.88	0.258202	-	-0.582238	0.599283	-0.0769519	0.93 [0.69 - 1.23]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2990	Breast cancer		Distant Metastasis Free Survival	Uppsala, Oxford	Sotiriou	HG-U133A	222268_x_at	125	0.744	0.104351	-	-0.854011	0.738575	-0.261003	0.77 [0.17 - 3.57]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2990	Breast cancer		Relapse Free Survival	Uppsala, Oxford	Sotiriou	HG-U133A	222268_x_at	125	0.744	0.0158547	0.226946	-0.957955	0.142829	-0.93021	0.39 [0.11 - 1.37]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2990	Breast cancer		Distant Metastasis Free Survival	Uppsala, Oxford	Sotiriou	HG-U133A	222268_x_at	54	0.203704	0.122281	-	15.3283	0.713137	-0.112872	0.89 [0.49 - 1.63]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2990	Breast cancer		Distant Metastasis Free Survival	Uppsala, Oxford	Sotiriou	HG-U133A	213432_at	125	0.536	0.0160191	0.228646	0.942968	0.381146	0.139788	1.15 [0.84 - 1.57]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2990	Breast cancer		Distant Metastasis Free Survival	Uppsala, Oxford	Sotiriou	HG-U133A	213432_at	54	0.296296	0.248382	-	1.14791	0.921565	-0.0215771	0.98 [0.64 - 1.50]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2990	Breast cancer		Relapse Free Survival	Uppsala, Oxford	Sotiriou	HG-U133A	213432_at	62	0.709677	0.229894	-	0.567588	0.647373	0.0774408	1.08 [0.78 - 1.51]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7390	Breast cancer		Distant Metastasis Free Survival	psala, Oxford, Stockholm, IGR, GUYT, CRH (1980-1998)	Desmedt	HG-U133A	222268_x_at	198	0.888889	0.0460209	0.467165	0.653262	0.384742	0.0819481	1.09 [0.90 - 1.31]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7390	Breast cancer		Relapse Free Survival	psala, Oxford, Stockholm, IGR, GUYT, CRH (1980-1998)	Desmedt	HG-U133A	213432_at	198	0.787879	0.079211	-	-0.505994	0.510913	-0.0374262	0.96 [0.86 - 1.08]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7390	Breast cancer		Overall Survival	psala, Oxford, Stockholm, IGR, GUYT, CRH (1980-1998)	Desmedt	HG-U133A	222268_x_at	198	0.222222	0.10297	-	0.613718	0.71488	0.0356447	1.04 [0.86 - 1.25]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7390	Breast cancer		Overall Survival	psala, Oxford, Stockholm, IGR, GUYT, CRH (1980-1998)	Desmedt	HG-U133A	213432_at	198	0.353535	0.0528911	-	0.589154	0.815359	0.0156911	1.02 [0.89 - 1.16]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7390	Breast cancer		Distant Metastasis Free Survival	psala, Oxford, Stockholm, IGR, GUYT, CRH (1980-1998)	Desmedt	HG-U133A	213432_at	198	0.353535	0.175603	-	0.377008	0.930912	-0.00585968	0.99 [0.87 - 1.13]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE7390	Breast cancer		Relapse Free Survival	psala, Oxford, Stockholm, IGR, GUYT, CRH (1980-1998)	Desmedt	HG-U133A	222268_x_at	198	0.217172	0.0207192	0.274701	0.688533	0.144728	0.115246	1.12 [0.96 - 1.31]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12945	Colorectal cancer		Overall Survival	Berlin	Staub	HG-U133A	213432_at	62	0.322581	0.0313502	0.36483	-1.19153	0.673616	-0.14782	0.86 [0.43 - 1.72]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12945	Colorectal cancer		Disease Free Survival	Berlin	Staub	HG-U133A	222268_x_at	51	0.72549	0.0784358	-	1.47188	0.0991694	3.55409	34.96 [0.51 - 2389.29]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12945	Colorectal cancer		Disease Free Survival	Berlin	Staub	HG-U133A	213432_at	51	0.137255	0.0637611	-	-1.5413	0.311886	-0.825235	0.44 [0.09 - 2.17]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE12945	Colorectal cancer		Overall Survival	Berlin	Staub	HG-U133A	222268_x_at	62	0.870968	0.0913438	-	1.07633	0.545657	0.876698	2.40 [0.14 - 41.28]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17536	Colorectal cancer		Disease Specific Survival	MCC	Smith	HG-U133_Plus_2	213432_at	177	0.146893	0.195059	-	0.599591	0.544381	0.0604966	1.06 [0.87 - 1.29]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17536	Colorectal cancer		Disease Free Survival	MCC	Smith	HG-U133_Plus_2	222268_x_at	145	0.8	0.100362	-	-0.846887	0.739787	-0.179574	0.84 [0.29 - 2.41]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17536	Colorectal cancer		Overall Survival	MCC	Smith	HG-U133_Plus_2	213432_at	177	0.327684	0.184582	-	-0.322154	0.964992	-0.00382845	1.00 [0.84 - 1.18]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17536	Colorectal cancer		Disease Free Survival	MCC	Smith	HG-U133_Plus_2	213432_at	145	0.117241	0.162614	-	-0.582244	0.793316	0.0320597	1.03 [0.81 - 1.31]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17536	Colorectal cancer		Disease Specific Survival	MCC	Smith	HG-U133_Plus_2	222268_x_at	177	0.112994	0.00385583	0.0770355	-0.984904	0.0396355	-0.893192	0.41 [0.17 - 0.96]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17536	Colorectal cancer		Overall Survival	MCC	Smith	HG-U133_Plus_2	222268_x_at	177	0.350282	0.0177476	0.24613	-0.570472	0.028257	-0.830626	0.44 [0.21 - 0.92]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE14333	Colorectal cancer		Disease Free Survival	Melbourne	Jorissen	HG-U133_Plus_2	222268_x_at	226	0.712389	0.0446578	0.458498	0.577054	0.540653	0.0643078	1.07 [0.87 - 1.31]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE14333	Colorectal cancer		Disease Free Survival	Melbourne	Jorissen	HG-U133_Plus_2	213432_at	226	0.168142	0.005301	0.0990904	-0.85839	0.306848	-0.0541525	0.95 [0.85 - 1.05]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17537	Colorectal cancer		Overall Survival	VMC	Smith	HG-U133_Plus_2	213432_at	55	0.272727	0.0174365	0.243033	-1.02643	0.28077	-0.207913	0.81 [0.56 - 1.19]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17537	Colorectal cancer		Disease Specific Survival	VMC	Smith	HG-U133_Plus_2	222268_x_at	49	0.122449	0.0247828	0.311154	-1.39468	0.354336	-0.954909	0.38 [0.05 - 2.90]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17537	Colorectal cancer		Disease Free Survival	VMC	Smith	HG-U133_Plus_2	213432_at	55	0.272727	0.000133522	0.00451543	-1.7658	0.00722948	-0.592375	0.55 [0.36 - 0.85]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17537	Colorectal cancer		Disease Specific Survival	VMC	Smith	HG-U133_Plus_2	213432_at	49	0.265306	0.00768824	0.132221	-1.43255	0.274932	-0.263244	0.77 [0.48 - 1.23]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17537	Colorectal cancer		Overall Survival	VMC	Smith	HG-U133_Plus_2	222268_x_at	55	0.145455	0.0707413	-	-0.904392	0.30946	-0.775874	0.46 [0.10 - 2.06]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE17537	Colorectal cancer		Disease Free Survival	VMC	Smith	HG-U133_Plus_2	222268_x_at	55	0.818182	0.0129034	0.19524	-15.5065	0.0201724	-2.08509	0.12 [0.02 - 0.72]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE22138	Eye cancer	Uveal melanoma	Distant Metastasis Free Survival	BRCIC	Laurent	HG-U133_Plus_2	213432_at	63	0.190476	6.32E-05	0.00233604	-1.37077	0.282916	-10.0641	0.00 [0.00 - 4048.33]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE22138	Eye cancer	Uveal melanoma	Distant Metastasis Free Survival	BRCIC	Laurent	HG-U133_Plus_2	222268_x_at	63	0.190476	0.00095618	0.0245445	-1.16718	0.0310448	-18.2148	0.00 [0.00 - 0.19]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2837	Head and neck cancer	Squamous cell carcinoma	Relapse Free Survival	VUMC, VAMC, UTMDACC (1992-2005)	Chung	U133_X3P	222268_3p_x_at	28	0.357143	0.04462	0.458256	-1.1553	0.250504	-5.7954	0.00 [0.00 - 59.71]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2837	Head and neck cancer	Squamous cell carcinoma	Relapse Free Survival	VUMC, VAMC, UTMDACC (1992-2005)	Chung	U133_X3P	Hs.321110.0.S1_3p_x_at	28	0.607143	0.14703	-	-0.918039	0.285754	-11.2103	0.00 [0.00 - 11764.43]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2837	Head and neck cancer	Squamous cell carcinoma	Relapse Free Survival	VUMC, VAMC, UTMDACC (1992-2005)	Chung	U133_X3P	Hs.102482.1.A1_3p_at	28	0.142857	0.0685508	-	-1.19105	0.284826	-14.681	0.00 [0.00 - 203362.22]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE2837	Head and neck cancer	Squamous cell carcinoma	Relapse Free Survival	VUMC, VAMC, UTMDACC (1992-2005)	Chung	U133_X3P	Hs.321110.0.S1_3p_at	28	0.214286	0.0163753	0.232305	-1.24263	0.148495	-40.5154	0.00 [0.00 - 1876209.36]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	jacob-00182-CANDF	Lung cancer	Adenocarcinoma	Overall Survival	CAN/DF	Shedden	HG-U133A	213432_at	82	0.878049	0.240369	-	0.522425	0.876809	0.012293	1.01 [0.87 - 1.18]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	jacob-00182-CANDF	Lung cancer	Adenocarcinoma	Overall Survival	CAN/DF	Shedden	HG-U133A	222268_x_at	82	0.658537	0.0741784	-	0.612601	0.757625	0.0669511	1.07 [0.70 - 1.64]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	jacob-00182-HLM	Lung cancer	Adenocarcinoma	Overall Survival	HLM	Shedden	HG-U133A	213432_at	79	0.810127	0.149077	-	0.469214	0.719569	0.0222022	1.02 [0.91 - 1.15]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	jacob-00182-HLM	Lung cancer	Adenocarcinoma	Overall Survival	HLM	Shedden	HG-U133A	222268_x_at	79	0.822785	0.0235559	0.300438	0.759418	0.139414	0.409003	1.51 [0.88 - 2.59]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	jacob-00182-MSK	Lung cancer	Adenocarcinoma	Overall Survival	MSK	Shedden	HG-U133A	222268_x_at	104	0.307692	0.0158806	0.227214	1.02915	0.346073	0.185491	1.20 [0.82 - 1.77]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE31210	Lung cancer	Adenocarcinoma	Overall Survival	NCCRI	Okayama	HG-U133_Plus_2	222268_x_at	204	0.784314	0.0115268	0.179583	0.912977	0.176318	0.325094	1.38 [0.86 - 2.22]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE31210	Lung cancer	Adenocarcinoma	Relapse Free Survival	NCCRI	Okayama	HG-U133_Plus_2	222268_x_at	204	0.112745	0.137888	-	-0.535869	0.967005	0.00874848	1.01 [0.67 - 1.53]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE31210	Lung cancer	Adenocarcinoma	Overall Survival	NCCRI	Okayama	HG-U133_Plus_2	213432_at	204	0.676471	0.0271541	0.331219	0.786874	0.0638283	0.145775	1.16 [0.99 - 1.35]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE31210	Lung cancer	Adenocarcinoma	Relapse Free Survival	NCCRI	Okayama	HG-U133_Plus_2	213432_at	204	0.828431	0.0295614	0.35078	0.664076	0.190598	0.0756894	1.08 [0.96 - 1.21]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	jacob-00182-UM	Lung cancer	Adenocarcinoma	Overall Survival	UM	Shedden	HG-U133A	222268_x_at	178	0.898876	0.0102658	0.164678	0.704777	0.463982	0.103295	1.11 [0.84 - 1.46]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE3141	Lung cancer	NSCLC	Overall Survival	Duke	Bild	HG-U133_Plus_2	222268_x_at	111	0.756757	0.157667	-	0.405406	0.580544	0.059817	1.06 [0.86 - 1.31]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE3141	Lung cancer	NSCLC	Overall Survival	Duke	Bild	HG-U133_Plus_2	213432_at	111	0.576577	0.249077	-	-0.312461	0.665212	0.0270191	1.03 [0.91 - 1.16]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE14814	Lung cancer	NSCLC	Overall Survival	JRB.10	Zhu	HG-U133A	213432_at	90	0.844444	0.0455218	0.464009	0.721806	0.307879	0.157437	1.17 [0.86 - 1.58]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE14814	Lung cancer	NSCLC	Disease Specific Survival	JRB.10	Zhu	HG-U133A	213432_at	90	0.844444	0.155772	-	0.607385	0.418538	0.143987	1.15 [0.81 - 1.64]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE14814	Lung cancer	NSCLC	Overall Survival	JRB.10	Zhu	HG-U133A	222268_x_at	90	0.577778	0.12857	-	0.490688	0.42491	0.368906	1.45 [0.58 - 3.58]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE14814	Lung cancer	NSCLC	Disease Specific Survival	JRB.10	Zhu	HG-U133A	222268_x_at	90	0.788889	0.0736108	-	0.710765	0.405953	0.434562	1.54 [0.55 - 4.30]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE8894	Lung cancer	NSCLC	Relapse Free Survival	Seoul (1995-2005)	Lee	HG-U133_Plus_2	222268_x_at	138	0.130435	0.122962	-	-0.487517	0.456656	-0.236856	0.79 [0.42 - 1.47]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE8894	Lung cancer	NSCLC	Relapse Free Survival	Seoul (1995-2005)	Lee	HG-U133_Plus_2	213432_at	138	0.84058	0.0279049	0.337403	0.622948	0.344771	0.0446189	1.05 [0.95 - 1.15]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4573	Lung cancer	Squamous cell carcinoma	Overall Survival	Michigan (1991-2002)	Raponi	HG-U133A	222268_x_at	129	0.899225	0.0108652	0.171834	0.906163	0.415948	0.150742	1.16 [0.81 - 1.67]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE4573	Lung cancer	Squamous cell carcinoma	Overall Survival	Michigan (1991-2002)	Raponi	HG-U133A	213432_at	129	0.658915	0.00268797	0.0576385	0.747073	0.0778376	0.157681	1.17 [0.98 - 1.40]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE9891	Ovarian cancer		Overall Survival	AOCS, RBH, WH, NKI-AVL (1992-2006)	Tothill	HG-U133_Plus_2	213432_at	278	0.154676	0.000701965	0.0189115	-0.717531	0.00178554	-0.266836	0.77 [0.65 - 0.91]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE9891	Ovarian cancer		Overall Survival	AOCS, RBH, WH, NKI-AVL (1992-2006)	Tothill	HG-U133_Plus_2	222268_x_at	278	0.776978	0.00152948	0.0363149	-0.90355	0.00320118	-1.00539	0.37 [0.19 - 0.71]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	DUKE-OC	Ovarian cancer		Overall Survival	Duke	Bild	HG-U133A	222268_x_at	133	0.548872	0.00948687	0.155171	0.6075	0.21746	-0.249638	0.78 [0.52 - 1.16]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	DUKE-OC	Ovarian cancer		Overall Survival	Duke	Bild	HG-U133A	213432_at	133	0.887218	0.00974187	0.15831	-1.65835	0.180505	-0.0768546	0.93 [0.83 - 1.04]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE8841	Ovarian cancer		Overall Survival	Milan (1992-2003)	Marchini	G4100A	8877	81	0.111111	0.0193573	0.261826	-1.32277	0.100403	-0.353955	0.70 [0.46 - 1.07]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE8841	Ovarian cancer		Overall Survival	Milan (1992-2003)	Marchini	G4100A	12783	81	0	1	-	0	1	0	1.00 [1.00 - 1.00]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE8841	Ovarian cancer		Overall Survival	Milan (1992-2003)	Marchini	G4100A	6993	81	0.111111	0.113891	-	-0.91062	0.487354	-0.392573	0.68 [0.22 - 2.04]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE26712	Ovarian cancer		Overall Survival	MSKCC (1990-2003)	Bonome	HG-U133_Plus_2	213432_at	185	0.778378	0.101571	-	-0.369136	0.439381	-0.114418	0.89 [0.67 - 1.19]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE26712	Ovarian cancer		Disease Free Survival	MSKCC (1990-2003)	Bonome	HG-U133_Plus_2	222268_x_at	185	0.767568	0.0742086	-	0.338238	0.468884	0.137584	1.15 [0.79 - 1.67]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE26712	Ovarian cancer		Disease Free Survival	MSKCC (1990-2003)	Bonome	HG-U133_Plus_2	213432_at	185	0.12973	0.0395386	0.424519	0.497956	0.548251	-0.0789755	0.92 [0.71 - 1.20]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE26712	Ovarian cancer		Overall Survival	MSKCC (1990-2003)	Bonome	HG-U133_Plus_2	222268_x_at	185	0.783784	0.0595283	-	0.388301	0.540023	0.127411	1.14 [0.76 - 1.71]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE14764	Ovarian cancer		Overall Survival	TOC	Denkert	HG-U133A	222268_x_at	80	0.575	0.265518	-	0.487312	0.948225	0.0132949	1.01 [0.68 - 1.51]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE14764	Ovarian cancer		Overall Survival	TOC	Denkert	HG-U133A	213432_at	80	0.6375	0.0921577	-	-0.836403	0.97688	0.0033982	1.00 [0.80 - 1.26]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE19234	Skin cancer	Melanoma	Overall Survival	NYU	Bogunovic	HG-U133_Plus_2	222268_x_at	38	0.315789	0.00538415	0.100307	-1.27738	0.0383015	-0.402311	0.67 [0.46 - 0.98]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE19234	Skin cancer	Melanoma	Overall Survival	NYU	Bogunovic	HG-U133_Plus_2	213432_at	38	0.736842	0.00776952	0.133291	1.22306	0.0138139	1.27093	3.56 [1.30 - 9.80]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE30929	Soft tissue cancer	Liposarcoma	Distant Recurrence Free Survival	MSKCC (1993-2008)	Gobble	HG-U133A	222268_x_at	140	0.278571	0.0496542	0.489545	-0.596332	0.372695	-0.604911	0.55 [0.14 - 2.06]
GENE	MUC5B	mucin 5B, oligomeric mucus/gel-forming	None	GSE30929	Soft tissue cancer	Liposarcoma	Distant Recurrence Free Survival	MSKCC (1993-2008)	Gobble	HG-U133A	213432_at	140	0.85	0.000687088	0.0185714	-15.3632	0.154222	-1.40217	0.25 [0.04 - 1.69]

Table S2: PrognoScan analysis of MUC5B in multiple cancers

AML, acute myelocytic leukemia; DLBCL, diffuse large B-cell lymphoma; NSCLC, non-small cell lung carcinoma.

BMJ Open

A favorable prognostic marker for EGFR mutant non-small cell lung cancer: immunohistochemical analysis of MUC5B

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2015-008366.R1
Article Type:	Research
Date Submitted by the Author:	08-May-2015
Complete List of Authors:	Wakata, Kouki; Nagasaki University, Surgical Oncology Tsuchiya, Tomoshi; Nagasaki University, Surgical Oncology Tomoshige, Koichi; Nagasaki University, Surgical Oncology Takagi, Katsunori; Nagasaki University, Surgical Oncology Yamasaki, Naoya; Nagasaki University, Surgical Oncology Matsumoto, Keitaro; Nagasaki University, Surgical Oncology Miyazaki, Takuro; Nagasaki University, Surgical Oncology Nanashima, Atsushi; Nagasaki University, Surgical Oncology Whitsett, Jeffrey; CCHMC, Pulmonary Biology Maeda, Yutaka; CCHMC, Pulmonary Biology Nagayasu, Takeshi; Nagasaki University, Surgical Oncology
Primary Subject Heading:	Oncology
Secondary Subject Heading:	Pathology, Respiratory medicine
Keywords:	Pathology < BASIC SCIENCES, Respiratory tract tumours < THORACIC MEDICINE, Surgical pathology < PATHOLOGY

A favorable prognostic marker for EGFR mutant non-small cell lung cancer: immunohistochemical analysis of MUC5B

Kouki Wakata[1], Tomoshi Tsuchiya[1], Koichi Tomoshige[1], Katsunori Takagi[1], Naoya Yamasaki[1], Keitaro Matsumoto[1], Takuro Miyazaki[1], Atsushi Nanashima[1], Jeffrey A. Whitsett[2], Yutaka Maeda[2], Takeshi Nagayasu[1]

Correspondence and requests for reprints should be addressed to (Yutaka Maeda and Takeshi Nagayasu are co-corresponding authors):

Correspondence to

Yutaka Maeda, D.V.M, Ph.D.

Perinatal Institute, Division of Neonatology, Perinatal and Pulmonary Biology

Cincinnati Children's Hospital Medical Center and the University of Cincinnati College of Medicine

Cincinnati, Ohio 45229, USA

Tel: +1-513-636-3322

Fax: +1-513-636-7868

E-mail: yutaka.maeda@cchmc.org

Takeshi Nagayasu, M.D., Ph.D.

Division of Surgical Oncology

Nagasaki University Graduate School of Biomedical Sciences

1-7-1 Sakamoto Nagasaki-city, Nagasaki 852-8501, Japan

Tel: +81-95-819-7304

Fax: +81-95-819-7306

E-mail: nagayasu@nagasaki-u.ac.jp

Key words: MUC5B; MUC5AC; epidermal growth factor receptor; NSCLC; mutation

Word count: 2547

Key messages

What is the key question?

Is there a specific prognostic biomarker for patients with surgically resected NSCLC that harbors an *EGFR* mutation?

What is the bottom line?

We identified MUC5B, a mucous protein, as a favorable prognostic marker specific for patients with *EGFR* mutant NSCLC.

Why read on?

MUC5B may serve as a biomarker that could be used to direct the therapeutic strategy for treating *EGFR* mutant NSCLC.

Abstract

Objectives: To determine the use of the mucin proteins MUC5B and MUC5AC as prognosis markers for non-small cell lung cancer (NSCLC) carrying *EGFR* mutations.

Setting: Patients who underwent surgical resection at Nagasaki University Hospital and related facilities in Japan between June 1996 and March 2013.

Participant: One hundred and fifty-nine Japanese patients (male: n=103; female: n=56) with NSCLC who underwent surgical resection (*EGFR* mutant type: n=78, *EGFR* wild type: n=81).

Results: Patients whose tumors expressed MUC5B had significantly longer overall survival (OS) and relapse-free survival (RFS) compared to the MUC5B negative patients with *EGFR* mutant NSCLC (p=0.0098 and p=0.0187, respectively). In patients with *EGFR* wild type NSCLC, there was no association with MUC5B expression. MUC5AC expression was not different between *EGFR* mutant and wild type NSCLC.

Conclusions: Present findings indicate that MUC5B, but not MUC5AC, is a novel prognostic biomarker for patients with NSCLC carrying *EGFR* mutations but not for patients with NSCLC carrying wild type *EGFR*.

Strengths and limitations of this study

1. A prognostic marker for each driver mutation in NSCLC has not yet been determined.

2. MUC5B is a favorable postoperative prognostic marker for *EGFR* mutant NSCLC.

3. MUC5AC is not correlated with postoperative prognosis regardless of *EGFR* mutation status.

4. The function of MUC5B in *EGFR* mutant NSCLC remains unknown.

Introduction

Lung cancer is the primary cause of cancer-related death in the United States and worldwide [1]. Non-small cell lung cancer (NSCLC) accounts for approximately 80-85% of all lung cancers [1]. Currently, targeted therapies for non-resectable NSCLC have progressed rapidly, based on the discovery of pharmacologically treatable driver mutations in epidermal growth factor receptor (*EGFR*) and fusions of anaplastic lymphoma kinase (ALK) [2,3]. These moleculary targeted therapies have revealed distinct and/or overlapping tumorigenic pathways associated with each driver mutation, especially regarding the mechanisms of tumor recurrence [4]. Genetic screening of driver mutations, including *EGFR* mutations and *ALK* fusions, is now common for metastatic NSCLC but not for surigically resected primary NSCLC [5]. In the ALCHEMIST lung cancer trials (http://www.cancer.gov/researchandfunding/areas/clinical-trials/nctn/alchemist), patients whose primary lung tumors carry *EGFR* mutations (*EGFR*-mutant patients) are being tested for adjuvant therapy of erlotinib targeting EGFR mutations. However, a favorable or poor prognostic biomarker associated with *EGFR* mutations is not known. Such biomarkers will be useful to determine *EGFR*-mutant patients who would benefit most from the adjuvant therapy of erlotinib and to avoid such unnecessary therapy after surgery in patients who would not benefit.

Recently, we reported that decreased expression of *Nkx2-1* (also known as TTF-1) in a mouse model of *EGFR* mutant NSCLC reduced the number and size of lung tumors [6] and extended the survival of the mice (see online supplementary figure S1). Unexpectedly, the decreased *Nkx2-1* induced the expression of a mucin protein MUC5B but not MUC5AC in *EGFR*-mutant lung tumors in the mice [6], suggesting that MUC5B may serve as a favorable prognostic marker associated with *EGFR* mutant NSCLC in humans. In the present study, we assessed whether the expression of MUC5B in the primarily resected *EGFR* mutant or wild type lung tumors is linked

to survival of the patients after surgery. Our study provides a novel approach to assess prognosis for patients whose primarily resected lung tumors carry *EGFR* mutations.

Methods

Study Population

Among the patients who underwent surgical resection at Nagasaki University Hospital and related facilities between June 1996 and March 2013, patients who were tested for the presence or absence of *EGFR* mutations were selected for this study. The *EGFR* mutations were confirmed internally or externally (LSI Medience Corporation, Japan). We further selected the patients whose clinicopathological characteristics were retrieved from the patients' charts and whose prognosis was followed at our institution and related facilities. We enrolled 159 patients (*EGFR* mutant type: n=78, *EGFR* wild type: n=81) for this study (table 1). All investigations were approved by our institution and related facilities review boards, and informed consent was obtained from all participants prior to the study.

Clinicopathological Evaluation

Histological classification of NSCLC was designated as to three types – well, moderately and poorly differentiated – based on the predominant features according to the World Health Organization classification.[7] The patients remained for a median follow-up period of 1680 days, ranging from 55 to 4503 days. For all patients, periodic inspection with chest x-ray, CT scan and tumor marker assays was performed at least every six months to confirm the presence or absence of recurrence, even if patients experienced no complaints or no symptoms.

Antibody Information

For immunohistochemical staining, primary antibodies were used at the following concentrations: rabbit polyclonal anti-MUC5B (1:200; sc-20119, Santa Cruz Biotechnology), rabbit polyclonal anti-MUC5AC (1:50; sc-20118, Santa Cruz Biotechnology).

Sample Preparation, Selection and Immunohistochemistry

The 5-μm thick formalin-fixed, paraffin-embedded (FFPE) lung sections were deparaffinized in dimethylbenzene and dehydrated through a graded alcohol series. For antigen retrieval, the FFPE lung sections were incubated in 10 mM citric acid (pH 6.0) at 121°C for 15 min and then washed in phosphate-buffered saline (PBS). Next, the lung sections were immersed in 3% H_2O_2 solution for 30 min to block the endogenous peroxidase followed by incubation with each primary antibody at 4 °C overnight. After washing in PBS, the lung sections were incubated with the peroxidase-conjugated secondary antibodies (Simple Stain MAX-PO kit, Nichirei, Tokyo, Japan) for 30 min at room temperature. For IHC staining, the lung sections were visualized with a diaminobenzidine (DAB: brown) kit (Histofine, Nichirei) and counterstained with hematoxylin. The lung sections visualized with DAB were dehydrated with alcohol and dimethylbenzene and mounted in a conventional fashion.

Normal bronchial tissue specimens that moderately expressed MUC5B were prepared as positive controls in all cases. Normal gastric mucosa tissue specimens that moderately expressed MUC5AC were prepared as positive controls in all cases. Negative controls were also prepared in all cases. MUC5B and MUC5AC staining was evaluated by immunohistochemistry by two independent trained observers (K.W. and T.T.). The pathological criteria was determined by reference to guideline for human epidermal growth factor receptor 2 (Her2/neu) testing in breast cancer (score 0, no staining observed or incomplete faint/barely perceptible cytoplasmic staining of < 10% of tumor cells; score 1, incomplete faint/barely perceptible cytoplasmic staining of > 10 % of tumor cells; score 2, incomplete weak/moderate cytoplasmic staining of > 10 % of tumor cells; score 3, complete and intense cytoplasmic staining of > 30 % of tumor cells) [8]. Score 0 and 1 were further categorized as negative, and score 2 and 3 as positive.

Statistical Analysis

For univariate analysis, categorical data were analyzed by the chi-square test or Fisher's exact test or the Cochran-Armitage test. Continuous data were expressed as a mean using the Mann-Whitney U test or the Kruskal-Wallis test. The overall survival (OS) and relapse-free survival (RFS) were calculated according to the Kaplan-Meier method, and differences between groups were tested for significance using the log-rank test. Subjects who neither died nor had recurrence were censored at the time of their last follow-up. The prognostic relevance of a single factor was determined by multivariate Cox regression analysis. A p-value of 0.05 or less was considered significant. SPSS version 17 software (SPSS Japan, Tokyo, Japan) was used for the analysis.

Results

Expression of MUC5B and MUC5AC in human NSCLC

Since MUC5B is an abundant cytoplasmic and secreted protein, we assessed whether MUC5B could be used as a prognostic marker for patients with NSCLC carrying *EGFR* mutations in primary resected human lung tumors. Primary resected NSCLC tumors were tested immunohistochemically for the presence of MUC5B. MUC5B staining was detected in the cytoplasm of NSCLC cells in 27 of the 78 samples with *EGFR* mutations and 29 of the 81 samples with wild type *EGFR* (figure 1A,B). The NSCLC samples were also tested using MUC5AC antibody, detecting expression of MUC5AC in the cytoplasm of NSCLC cells in 20 of the 73 samples with *EGFR* mutations and 24 of the 79 samples with wild type *EGFR* (figure 1C,D). These results indicate that both MUC5B and MUC5AC are expressed in a portion of human NSCLC.

Prognostic Association of MUC5B or MUC5AC with *EGFR* mutant or *EGFR* wild type NSCLC

Expression of MUC5B in NSCLC tumors carrying *EGFR* mutations was not correlated with clinicopathological parameters, including age, gender, smoking status, histological type, tumor size, degree of differentiation, stage, tumor status, nodal status, lymphatic invasion, venous invasion or adjuvant chemotherapy (table 2). Expression of MUC5B in NSCLC tumors with wild type *EGFR* was not correlated with all of the clinicopathological parameters except histological type (see online supplementary table S1).

Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying *EGFR* mutations or wild type *EGFR* were assessed. In a cohort of patients whose resected NSCLC tumors carrying *EGFR* mutations, univariate analysis showed significant differences ($p<0.05$) in

OS in expression of MUC5B, tumor size, histological type, degree of differentiation, stage, lymphatic invasion and venous invasion, and in RFS in expression of MUC5B, tumor size, degree of differentiation, stage, lymphatic invasion and venous invasion (table 3). Patients whose tumors expressed MUC5B (MUC5B positive patients) survived significantly longer than patients whose tumors do not express MUC5B (MUC5B negative patients) in both OS (5-year OS; 95.8% versus 65.1%, p=0.0098; figure 2A) and RFS (5-year RFS; 69.9% versus 44.0%, p=0.0187; figure 2B). Multivariate Cox regression analysis using the variables that were p<0.05 in univariate analysis showed that the expression of MUC5B was independently associated with better OS and RFS (p<0.05; table 4). In a cohort of patients whose resected NSCLC tumors with wild type *EGFR*, univariate analysis showed significant differences (p<0.05) in OS in smoking status, stage, venous invasion and adjuvant chemotherapy, and in RFS in lymphatic invasion and venous invasion (see online supplementary table S2). There was no significant difference between the MUC5B positive patients and MUC5B negative patients in OS and RFS (5-year OS; 59.5% versus 63.6%, 5-year RFS; 36.0% versus 48.5%, respectively, figure 2C,D). Expression of MUC5AC in NSCLC was not associated with OS and RFS regardless of *EGFR* mutation status (figure 3). These results indicate that MUC5B is a favorable prognostic marker for postoperative patients whose resected NSCLC tumors carry *EGFR* mutation but not those with wild type *EGFR*.

Discussion

In the present study, we demonstrate that expression of MUC5B in primary *EGFR* mutant NSCLC is associated with longer survival in patients with NSCLC. MUC5B, but not MUC5AC, is a favorable prognostic biomarker for NSCLC in humans carrying *EGFR* mutations. Our study also indicates that adjuvant chemotherapy is not effective for EGFR-mutant patients, suggesting that mutant EGFR-targeting drugs, including gefitinib or erlotinib, should be used as an adjuvant therapy mainly for MUC5B-negative EGFR-mutant patients who have a poorer prognosis than MUC5B-positive EGFR-mutant patients. Our results using MUC5B as a prognosis biomarker for EGFR-mutant patients should be integrated into the ALCHEMIST lung cancer trials to determine patients who would benefit most from the adjuvant therapy.

MUC5B has been assessed as a prognostic biomarker for multiple cancers in several studies, using RT-PCR, microarray analysis and immunohistochemistry (see online supplementary table S3).[9–14] mRNA data assessing MUC5B as a prognostic biomarker is available at PrognoScan, a database for meta-analysis of the prognostic value of genes using microarray data deposited to the public domain [15]. The prognostic impact of MUC5B expression differed among cancer types. In lung cancer, six microarray studies analyzed by PrognoScan did not indicate MUC5B as either a good or a poor prognostic biomarker [16–21]. Immunohistochemical analysis indicated MUC5B as a poor prognosis biomarker (see online supplementary table S3) [9,14], a finding contradicting our present study. Previous mRNA microarray and immunostaining were based on all NSCLCs independent of driver mutation-based classification, which differs from our analysis that was based on classification by *EGFR* mutations. The utility of MUC5B as a prognostic factor differed in the two breast cancer studies, depending on the molecular basis of the tumors. In all breast cancers, PrognoScan indicated that MUC5B was associated with poor prognosis [22,23]; however, in ER (Estrogen Receptor)-positive breast cancers, MUC5B was

associated with favorable prognosis [24], indicating the potential importance of tumor classification on a molecular basis. In the present study, we assessed MUC5B as a biomarker for NSCLC based on *EGFR* mutation status rather than on all NSCLCs, identifying MUC5B as a favorable prognosis biomarker for *EGFR* mutant NSCLC.

Regulation of *MUC5B* in *EGFR* mutant NSCLC is not well understood. *MUC5AC* and *MUC5B* genes are closely located at a locus on human chromosome 11. Both are evolutionally conserved gel-forming mucins secreted from airway epithelial cells in the lung. In normal lung, MUC5B is constitutively expressed at higher levels than MUC5AC.[25] In asthma and other inflammatory lung diseases, MUC5AC is highly induced in airway goblet cells.[25] In idiopathic pulmonary fibrosis (IPF), MUC5B but not MUC5AC is highly expressed in the airway goblet cells.[26,27] The SNP rs35705950 located at the *MUC5B* promoter is associated with induction of *MUC5B* mRNA in IPF;[27] however, we detected the SNP rs35705950 in only one of 27 cases in the *EGFR* mutant NSCLC expressing MUC5B (data not shown), indicating that the SNP is not associated with increased MUC5B in *EGFR* mutant NSCLC. Since MUC5B was induced in *EGFR* mutant lung tumors in *Nkx2-1* heterozygous mice ($EGFR^{L858R}$; $Nkx2\text{-}1^{+/-}$), *Muc5b* is suppressed by NKX2-1 in *EGFR* mutant NSCLC in mice.[6] Regulatory mechanisms controlling the *MUC5B* gene are not understood in *EGFR* mutant NSCLC in human. The function of MUC5B in cancer has been analyzed using a truncated MUC5B in MCF7 breast cancer cells, truncated MUC5B promoting tumorigenesis of MCF7 cells.[13] However, the use of truncated MUC5B may obscure the intrinsic role of full-length MUC5B since there is a possibility that the truncated MUC5B may function in a dominant-negative fashion. In lung, MUC5B is required for mucociliary clearance and innate immunity against bacterial infection.[28] The potential functions of MUC5B in lung cancer, including *EGFR* mutant lung cancer, are not known. The present study suggests that MUC5B or processes regulating *MUC5B* may influence the growth

and metastasis of *EGFR* mutant NSCLC. MUC5B may serve as a surrogate biomarker influenced by a pathway involved in metastasis and recurrence associated with *EGFR* mutant NSCLC.

In conclusion, our data revealed the clinicopathological significance of MUC5B as a favorable prognostic factor in resected *EGFR* mutant NSCLC. Further studies are necessary to elucidate the gene regulatory mechanism and the function of MUC5B in *EGFR* mutant NSCLC.

Author affiliations:

[1]Division of Surgical Oncology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

[2]Perinatal Institute, Division of Neonatology, Perinatal and Pulmonary Biology, Cincinnati Children's Hospital Medical Center and the University of Cincinnati College of Medicine, Cincinnati, Ohio, USA

Acknowledgments: We thank the patients, their families, all of the investigators who participated in the study, central laboratories that did the *EGFR* mutation testing, and Dr. Mary Durbin for discussions.

Contributors: Conception and design, K.W., T.T., J.A.W., Y.M., T.N.; provision of study material, patients, and data acquisition, K.W., T.T., K.T., K.T., N.Y., K.M., T.M., A.N., J.A.W., Y.M.; data analysis and interpretation, K.W., T.T., K.T., K.M., J.A.W., Y.M., T.N.; drafting manuscript and intellectual content, K.W., T.T., J.A.W., Y.M., T.N.

Funding: This work was supported by Ministry of Education, Science, and Culture of Japan, National Institute of Health, American Lung Association, the University of Cincinnati Postdoctoral Fellow Research Program, and Cincinnati Children’s Hospital Medical Center.

Competing interests: None

Patient consent: Obtained

Ethics approval: Ethics Committee from every participant hospital.

Provenance and peer review: Not commissioned; externally peer reviewed.

Data Sharing Statement: No additional data available.

References

1 Siegel R, Naishadham D JA. Cancer statistics for hispanics/latinos, 2012. *CA Cancer J Clin* 2012;**62**:283–98.

2 Lynch T, Bell D. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non–small-cell lung cancer to gefitinib. *N Engl J Med* 2004;**350**:2129–39.

3 Paez JG, Jänne P a, Lee JC, *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* 2004;**304**:1497–500.

4 Camidge DR, Pao W, Sequist L V. Acquired resistance to TKIs in solid tumours: learning from lung cancer. *Nat Rev Clin Oncol* 2014;**104**:[Epub ahead of print].

5 Shames DS, Wistuba II. The evolving genomic classification of lung cancer. *J Pathol* 2014;**232**:121–33.

6 Maeda Y, Tsuchiya T, Hao H, *et al.* KrasG12D and Nkx2-1 haploinsufficiency induce mucinous adenocarcinoma of the lung. *J Clin Invest* 2012;**122**:4388－400.

7 Travis WD, Brambilla E, Noguchi M, *et al.* International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: international multidisciplinary classification of lung adenocarcinoma: executive summary. *Proc Am Thorac Soc* 2011;**8**:381–5.

8 Wolff AC, Hammond MEH, Hicks DG, *et al.* Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* 2013;**31**:3997–4013.

9 Yu CJ, Yang PC, Shun CT, *et al.* Overexpression of MUC5 genes is associated with early post-operative metastasis in non-small-cell lung cancer. *Int J Cancer* 1996;**69**:457–65.

10 Pinto-de-Sousa J, Reis C a, David L, *et al.* MUC5B expression in gastric carcinoma: relationship with clinico-pathological parameters and with expression of mucins MUC1, MUC2, MUC5AC and MUC6. *Virchows Arch* 2004;**444**:224–30.

11 Varangot M, Barrios E, Sóñora C, *et al.* Clinical evaluation of a panel of mRNA markers in the detection of disseminated tumor cells in patients with operable breast cancer. *Oncol reports* 2005;**14**:537–45.

12 Partheen K, Levan K, Osterberg L, *et al.* Expression analysis of stage III serous ovarian adenocarcinoma distinguishes a sub-group of survivors. *Eur J Cancer* 2006;**42**:2846–54.

13 Valque H, Gouyer V, Gottrand F, *et al.* MUC5B leads to aggressive behavior of breast cancer MCF7 cells. *PLoS One* 2012;**7**:e46699.

14 Nagashio R, Ueda J, Ryuge S, *et al.* Diagnostic and Prognostic Significances of MUC5B and TTF-1 Expressions in Resected Non-Small Cell Lung Cancer. *Sci Rep* 2015;**5**:8649.

15 Mizuno H, Kitada K, Nakai K, *et al.* PrognoScan: a new database for meta-analysis of the prognostic value of genes. *BMC Med Genomics* 2009;**2**:18.

16 Shedden K, Taylor JMG, Enkemann S a, *et al.* Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008;**14**:822–7.

17 Okayama H, Kohno T, Ishii Y, *et al.* Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res* 2012;**72**:100–11.

18 Bild AH, Yao G, Chang JT, *et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;**439**:353–7.

19 Zhu C-Q, Ding K, Strumpf D, *et al.* Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. *J Clin Oncol* 2010;**28**:4417–24.

20 Lee E-S, Son D-S, Kim S-H, *et al.* Prediction of recurrence-free survival in postoperative non-small cell lung cancer patients by using an integrated model of clinical information and gene expression. *Clin Cancer Res* 2008;**14**:7397–404.

21 Raponi M, Zhang Y, Yu J, *et al.* Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res* 2006;**66**:7466–72.

22 Pawitan Y, Bjöhle J, Amler L, *et al.* Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005;**7**:R953–64.

23 Miller LD, Smeds J, George J, *et al.* An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A* 2005;**102**:13550–5.

24 Chanrion M, Negre V, Fontaine H, *et al.* A gene expression signature that can predict the recurrence of tamoxifen-treated primary breast cancer. *Clin Cancer Res* 2008;**14**:1744–52.

25 Evans C, Kim K, Tuvim M, *et al.* Mucus hypersecretion in asthma: causes and effects. *Curr Opin Pulm Med* 2009;**15**:4–11. doi:10.1097/MCP.0b013e32831da8d3.Mucus

26 Plantier L, Crestani B, Wert SE, *et al.* Ectopic respiratory epithelial cell differentiation in bronchiolised distal airspaces in idiopathic pulmonary fibrosis. *Thorax* 2011;**66**:651–7.

27 Seibold M, Wise A, Speer M, *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;**364**:1503–12.

28 Roy MG, Livraghi-Butrico A, Fletcher A a, *et al.* Muc5b is required for airway defence. *Nature* 2014;**505**:412–6.

Figure 1. Immunohistochemical staining for MUC5B and MUC5AC expression in NSCLC

Representative images of immune-positive staining for MUC5B in NSCLC (A) and negative staining (B), and positive staining for MUC5AC (C) and negative staining (D).

Figure 2. Survival curves for patients based on the expression of MUC5B in *EGFR* mutant or wild type NSCLC

Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying *EGFR* mutations or *EGFR* wild type. OS (A) and RFS (B) in the patients with *EGFR* mutant type NSCLC, and OS (C) and RFS (D) in the patients with EGFR wild type NSCLC.

Figure 3. Survival curves for patients based on the expression of MUC5AC in *EGFR* mutant or wild type NSCLC

Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying *EGFR* mutations or *EGFR* wild type. OS (A) and RFS (B) in the patients with *EGFR* mutant type NSCLC, and OS (C) and RFS (D) in the patients with *EGFR* wild type NSCLC.

Table 1. Baseline Characteristic of the 159 Patients with NSCLC

Number of Patients	
Median age (range, years)	67.5(32-90)
Gender	
Male	103(65%)
Female	56(35%)
Smoking status	
Non-smoker	54(34%)
Smoker	105(66%)
Histological type	
Adenocarcinoma: Bronchoalveolar	33(21%)
Adenocarcinoma	91(57%)
Squamous cell carcinoma	32(20%)
Adenosquamous carcinoma	2(1%)
Other†	1(1%)
Median tumor size (range, mm)	32.4(8-120)
Degree of differentiation	
Well	55(35%)
Moderately	69(43%)
Poorly	28(18%)
MD	7(4%)
Stage	
IA/IB	81(51%)
IIA/IIB	40(25%)
IIIA/IIIB	38(24%)
Tumor status	
T1-2	136(86%)
T3-4	23(14%)
Nodal status	
N0	103(65%)
N1-3	56(35%)
Lymphatic invasion	
Negative	56(35%)
Positive	102(64%)

MD	1(1%)
Venous invasion	
Negative	76(48%)
Positive	82(52%)
MD	1(1%)
Adjuvant chemotherapy	
Yes	85(53%)
No	74(47%)
EGFR	
Wild type	81(51%)
Mutant type	78(49%)

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

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

†Other mean NSCLC neuroendocrine.

Table 2. Association with Clinicopathological Data and the Expression of MUC5B of Patients with EGFR Mutant NSCLC

Parameters	Total (n=78)	MUC5B Negative (-) (n=51)	MUC5B Positive (+) (n=27)	P-value
Median age (range, years)	66.9(41-85)	66.7(42-83)	67.3(41-85)	0.674
Gender				
Male	35(45%)	22(28%)	13(17%)	
Female	43(55%)	29(37%)	14(18%)	0.6721
Smoking status				
Non-smoker	42(54%)	28(36%)	14(18%)	
Smoker	36(46%)	23(29%)	13(17%)	0.7971
Histological type				
Adenocarcinoma: Bronchoalveolar	27(35%)	16(21%)	11(14%)	
Adenocarcinoma	43(55%)	29(37%)	14(18%)	
Squamous cell carcinoma	8(10%)	6(8%)	2(3%)	0.6522
Median tumor size (range, mm)	25.7(8-60)	26.1(8-60)	24.9(8-50)	0.8771
Degree of differentiation				
Well	35(45%)	22(28%)	13(17%)	
Moderately	28(36%)	19(24%)	9(12%)	
Poorly	11(14%)	6(8%)	5(6%)	0.7348
MD	4(5%)	4(5%)	0(0%)	
Stage				
IA/IB	46(59%)	29(37%)	17(22%)	
IIA/IIB	14(18%)	8(10%)	6(8%)	
IIIA/IIIB	18(23%)	14(18%)	4(5%)	0.4162
Tumor status				
T1-2	68(87%)	42(54%)	26(33%)	
T3-4	10(13%)	9(12%)	1(1%)	0.0797
Nodal status				
N0	53(68%)	35(45%)	18(23%)	
N1-3	25(32%)	16(21%)	9(12%)	0.8599
Lymphatic invasion				
Negative	35(45%)	26(33%)	9(12%)	

Positive	42(54%)	24(31%)	18(23%)	0.1165
MD	1(1%)	1(1%)	0(0%)	
Venous invasion				
Negative	45(58%)	30(38%)	15(19%)	
Positive	32(41%)	20(26%)	12(15%)	0.7057
MD	1(1%)	1(1%)	0(0%)	
Adjuvant chemotherapy				
Yes	33(42%)	20(26%)	13(17%)	
No	45(58%)	31(40%)	14(18%)	0.4475

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

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

Table 3. Univariate Analysis for OS and RFS in EGFR Mutant NSCLC Patients

		OS		RFS	
Parameters	**N**	**%Survival**	**P-value**	**%Survival**	**P-value**
Age					
<70	39	75.5		63.6	
≧70	39	76.3	0.7311	41.5	0.1207
Gender					
Male	35	69		54.4	
Female	43	81.5	0.3166	51.8	0.813
Smoking status					
Non-smoker	43	75.9		49.1	
Smoker	36	73.3	0.9754	55.8	0.8508
MUC5B expression					
positive	27	95.8		69.9	
negative	51	65.1	0.0098	44	0.0187
Tumor size					
<20mm	25	95		85.7	
≧20mm	44	64.4	0.0058	31.2	0.0001
Histological type					
Adenocarcinoma	69	78.8		54	
Squamous cell carcinoma	8	46.9	0.0245	50	0.6369
Degree of differentiation					
Well	35	90		72.6	
Moderately	28	62.3		34.5	
Poorly	11	50.6	0.0172	38.1	0.0137
Stage					
I	45	87.8		62.8	
II/III	33	59.3	0.0095	38.8	0.0308
Lymphatic invasion					
Negative	35	92.3		74.2	
Positive	42	62.7	0.0075	35.8	0.0011
Venous invasion					
Negative	45	89		66.5	
Positive	32	55.7	0.0045	25.9	0.0021
Adjuvant chemotherapy					
Yes	33	62		37.5	
No	45	82.8	0.2569	63.4	0.0183

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 4. Multivariate Analysis for OS and RFS in EGFR Mutant NSCLC Patients

～OS～

Independent Variables	HR	95% CI	P-value
MUC5B (positive vs negative)	0.053	0.0064 to 0.4402	0.0065
Differentiation (well vs moderately vs poorly)	0.3762	0.1357 to 1.0428	0.0602
Stage I vs Stage II/III	0.5199	0.1687 to 1.6019	0.2545
Lymphatic invasion (positive vs negative)	2.9524	0.6572 to 13.2631	0.1578

～RFS～

Independent Variables	HR	95% CI	P-value
MUC5B (positive vs negative)	0.2913	0.1233 to 0.6886	0.005
Stage I vs Stage II/III	0.5937	0.2735 to 1.2888	0.1874
Lymphatic invasion (positive vs negative)	3.7624	1.5294 to 9.2556	0.0039

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

Figure 1

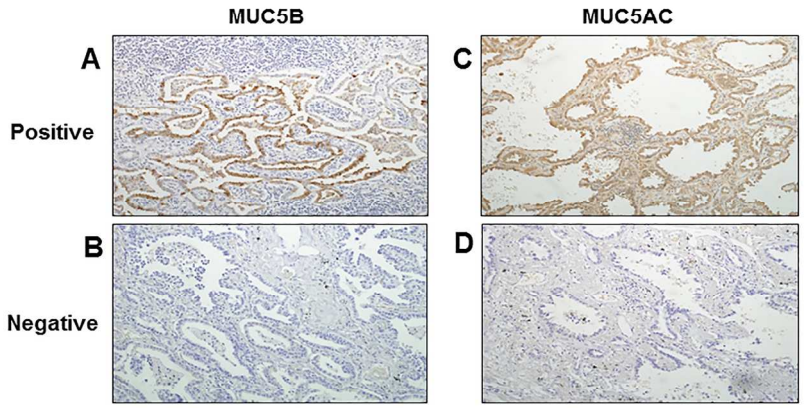


Figure 1. Immunohistochemical staining for MUC5B and MUC5AC expression in NSCLC
Representative images of immune-positive staining for MUC5B in NSCLC (A) and negative staining (B), and positive staining for MUC5AC (C) and negative staining (D).

254x142mm (300 x 300 DPI)

Figure 2

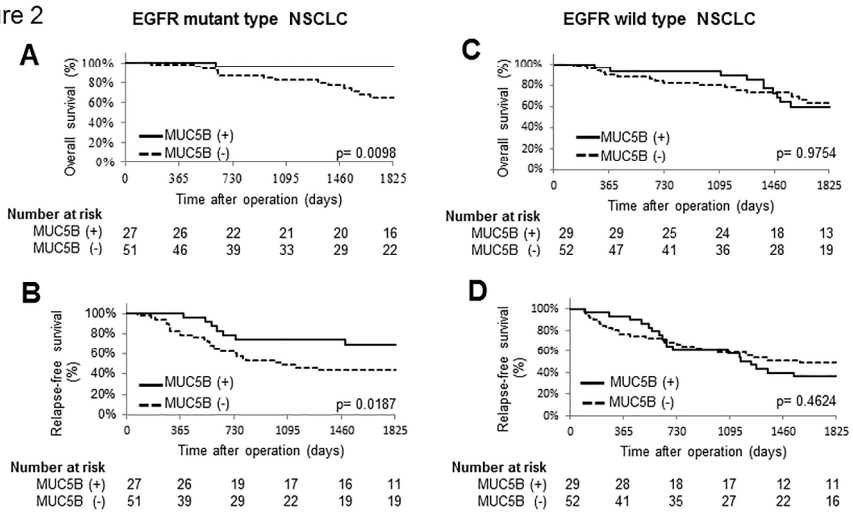


Figure 2. Survival curves for patients based on the expression of MUC5B in EGFR mutant or wild type NSCLC
Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying EGFR mutations or EGFR wild type. OS (A) and RFS (B) in the patients with EGFR mutant type NSCLC, and OS (C) and RFS (D) in the patients with EGFR wild type NSCLC.

254x142mm (300 x 300 DPI)

Figure 3

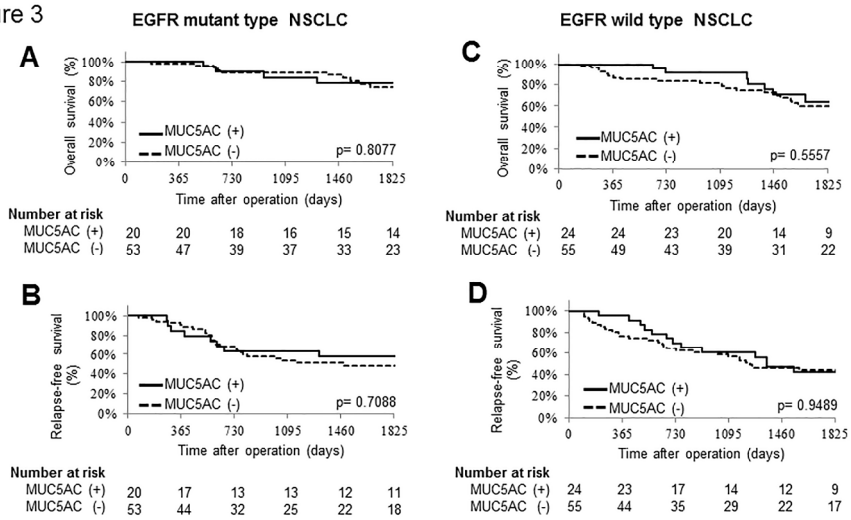


Figure 3. Survival curves for patients based on the expression of MUC5AC in EGFR mutant or wild type NSCLC

Overall and relapse-free survivals (OS and RFS) for patients with NSCLC carrying EGFR mutations or EGFR wild type. OS (A) and RFS (B) in the patients with EGFR mutant type NSCLC, and OS (C) and RFS (D) in the patients with EGFR wild type NSCLC.

254x142mm (300 x 300 DPI)

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.

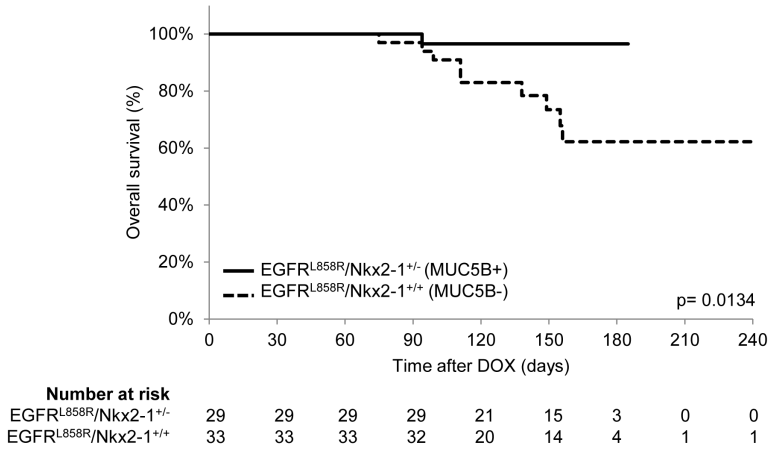


Figure S1. Kaplan-Meier analysis of overall survival in *EGFR*-mutant NSCLC mice

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	Total (n=81)	MUC5B Negative (-) (n=52)	MUC5B Positive (+) (n=29)	P-value
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.