

AD _____

Award Number: **W81XWH-04-1-0142**

TITLE: V.I.T.A.L. (Vanguard Investigations of Therapeutic Approaches to Lung Cancer)

PRINCIPAL INVESTIGATOR:

Waun Ki Hong, M.D., Principle Investigator
Reuben Lotan, Ph.D., Co-Principle Investigator
David Stewart, M.D., Co-Principle Investigator

CONTRACTING ORGANIZATION:

The University of Texas M.D. Anderson Cancer Center
Houston, TX 77030

REPORT DATE: January 2011

TYPE OF REPORT: Final Addendum

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

X Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.				
1. REPORT DATE (DD-MM-YYYY) 01-01-2011		2. REPORT TYPE Final Addendum		3. DATES COVERED (From - To) 14 Dec 2009 - 14 Dec 2010
4. TITLE AND SUBTITLE V.I.T.A.L. (Vanguard Investigations of Therapeutic Approaches Lung Cancer)			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER W81XWH-04-1-0142	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Waun Ki Hong, M.D. Edward Kim, M.D. Ignacio Wistuba, M.D. J. Jack Lee, Ph.D.			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The University of Texas M.D. Anderson Cancer Center Houston, TX 77030			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Commar Fort Detrick, Maryland 21702-5012			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited				
13. SUPPLEMENTARY NOTES				
14. ABSTRACT The VITAL Research Program will provide a better understanding of the cellular and molecular processes that drive lung tumorigenesis so that an accurate risk model for recurrence and/or the development of the secondary primary tumor can be developed, and the biologic agents most effective in reducing these events in the group of high-risk patients can be identified. We will be incorporating retrospective clinical trial specimens to develop our risk model and validating it with specimens collected from our Vanguard study. The research projects are proceeding well as proposed, producing valuable findings with cell lines, and will validate these results using the clinical samples obtained from the VITAL trials in the coming years.				
15. SUBJECT TERMS Lung cancer, risk model, cancer recurrence, clinical trials				
16. SECURITY CLASSIFICATION OF: U			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 226
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U		19a. NAME OF RESPONSIBLE PERSON USAMRMC
				19b. TELEPHONE NUMBER (include area code)

TABLE OF CONTENTS

INTRODUCTION	2
BODY.....	3
Project 1.....	3
Project 2.....	5
Project 3.....	14
Project 4.....	18
Project 5.....	24
Core B Biostatistics and Data Management	25
Core C Pathology and Specimen Procurement.....	27
KEY RESEARCH ACCOMPLISHMENTS	30
REPORTABLE OUTCOMES	32
CONCLUSIONS	34
APPENDICES.....	36
Appendix 1: (Core B) Trial Summary via Event Charts and Data Management Activities for VITAL	
Appendix 2: Statistical Report for ReVITALization Study	
Appendix 3: Publications and Abstracts	

INTRODUCTION

Smoking-related cancers such as lung and head and neck cancers are a major cause of cancer death in the United States. About 25% of lung cancer patients are diagnosed with stage I or II disease and undergo surgery with curative intent, but the 5-year survival for this group of patients is only 30%-70%. Patients with a strong history of smoking and prior early-stage cancer are found to be at high risk for cancer recurrence or development of second primary tumors (SPTs). An effective adjuvant therapy after surgery in this group of patients is not well established yet. The survival benefit of adjuvant chemotherapy was uncertain until recent findings reported by Winton and colleagues (Winton et al., 2005). They found that adjuvant chemotherapy (vinorelbine and cisplatin) increases the 5-year survival of surgically resected non-small cell lung cancer (NSCLC) patients, resolving the debate over the benefit of adjuvant chemotherapy. Thus, better-designed clinical trials and basic research are needed to establish the standard of care for these patients after surgery.

The program VITAL (Vanguard Trial of Investigational Therapeutics in Adjuvant Treatment of Lung Cancer) initiated in 2003 was developed to gain a better understanding of the molecular events underlying the progression of NSCLC in order to develop a risk model for cancer recurrence and development of smoking-related SPT in the high-risk population, and to identify effective preventive agents for this group of patients. Specifically, our objectives are:

- To identify biologically-based treatments for prevention of cancer recurrence and development of second primary tumors in high-risk patients;
- To understand molecular events in premalignant tissues that contribute to progression or malignancy;
- To develop a risk prediction model for disease recurrence and development of second primary tumors in high-risk patients by combining clinical treatment outcomes with molecular and imaging data.

Three clinical trials were proposed, in part, to acquire the necessary correlative samples to develop this risk model, which will significantly improve decision-making for patients and physicians in the management of this challenging disease. Histologic assessment was planned to determine whether malignant changes would occur during this time period. Despite substantial efforts, our patient accrual was significantly lower than expected due to a number of factors; thus, a ReVITALization plan was previously proposed (see revised Aims below) and approved by the DoD. Implementation of the alternative ReVITALization strategy over the past two years was based on the revised project aims that were developed to accomplish our goal of the development of the risk model. An overview of the changes is provided below with additional details in each relevant project.

A request for a final unfunded 6-month extension has been submitted for approval to the DoD.

ReVITALization Aims:

- 1. Circumvent low accruals using surgical specimens in our tissue bank.** These specimens (about 500 samples) of resected lung cancer will be utilized for biomarker assessment and will serve as the foundation for a biomarker-based risk assessment model.
- 2. Continue enrolling patients in our Vanguard trial to accrue 50-60 patients.** This cohort will provide sufficient biospecimens for the aims proposed in the other projects of the VITAL program. Additionally, the clinical data obtained from these patients will be

used to test the biomarker-based risk assessment model and the follow-up bronchoscopy specimens will provide important information for biomarker changes in the bronchial epithelium.

3. **Close the celecoxib and erlotinib trials to focus resources on specimen analyses to develop the biomarker risk model.**
4. **Perform two additional discovery projects related to increased risk that are only now possible due to continued progress in VITAL.**
 - a. Identify gene expression signatures in bronchial brush specimens using high-throughput genomics approach.
 - b. Identify genes expression signatures in epithelial cells detected by LIFE bronchoscopy that determine aggressiveness.

This report summarizes work conducted over the past year of the research period, and highlights key research accomplishments and reportable outcomes with the bibliography of all publications and meeting abstracts derived from VITAL during this timeframe.

PROGRESS REPORT

Project 1: Biologic Approaches for Adjuvant Treatment of Aerodigestive Tract Cancer

(PI and co-PIs: Drs. Waun Ki Hong, Edward S. Kim, Rodolfo C. Morice, David J. Stewart)

Aim 1 Assess the smoking-related disease-free survival in patients who are current or former smokers with a prior definitively-treated stage I/II lung or head and neck cancer.

The main objective for this project was to open the Vanguard study at MDACC as well as the 2 other participating sites. Enrollment was planned for a total of 300 patients with definitively treated stage I/II lung or head and neck cancer and at least a 20-pack-year smoking history. Patients undergo baseline testing including chest x-ray, CT scan, labs, bronchoscopy, and other specimen collections (i.e., sputum, saliva, serologies). Bronchoscopies and specimen collection are performed at baseline and at months 12, 24, and 36. White-light alone or white-light and autofluorescence modalities are used. Abnormal areas detected by bronchoscopy are biopsied. Histologic assessment is performed to determine whether malignant changes will occur during the time period. If severe dysplasia, carcinoma *in situ*, or carcinoma is discovered, patients follow the plans outlined in the clinical protocol. Once patients have completed 3 years of testing, they are followed until the study is completed. As per the revised ReVITALization aims, the study will be closed when a total of 50-60 patients have been accrued; all patients will be followed as outlined above.

Summary of Research Findings

We have now completed the patient enrollment in the VITAL/Vanguard trials, and the study was closed to new patients in March 2009. A total of 54 patients were enrolled to the Vanguard trial, and our number of evaluable patients (having completed both the baseline and 12-month bronchoscopy) is currently at 38, with 27 patients remaining on study. The median age is 61 years (range, 53-81 years), and 26 of the 54 patients (53%) were male. A total of 44 patients had NSCLC, with 36 patients having stage I (82%), 7 stage II (16%) and 1 IIIA (2%) disease. A

total of 38 pts (70%) have completed bronchoscopy at 12 months, 21 pts (39%) at 24 months, and 10 pts (19%) at 36 months. To date, a total of 9 pts (18%) have recurred and 3 pts (6%) have had SPTs recorded. Twelve additional bronchoscopies (2 at 24 months and 10 at 36 months) are scheduled to be completed by April 2011. Our plan is to continue to collect clinical data from these patients during the requested 6-month unfunded extension. Twenty-four- and 36-month bronchoscopies and longer follow-up for these last patients will be supported, after exhaustion of all VITAL funds, through other funding sources (e.g., institutional, philanthropy).

Dr. Ignacio Wistuba and his lab (Core C) have analyzed the bronchial epithelium and have completed the assessment of possible molecular markers. Planned biomarker analyses for Projects 2-5 have been supplemented using retrospectively collected specimens through the ReVITALization effort to maximize data acquisition while minimizing the time required to develop the proposed risk model (See Aim 3 below and Core B and C reports). The ReVITALization effort collected over 500 specimens and has completed the database with both pathological and clinical data.

Aim 2 Evaluate effects of biologic agents as adjuvant therapy on the modulation of histology and specific biomarkers in this high-risk population.

Current adjuvant chemotherapy offers some benefits in the high-risk patients, but is not a long-term preventive strategy. Our plan was to open several biologic adjuvant clinical trials with novel agents such as celecoxib, erlotinib, lonafarnib, and possibly others; however, poor accrual due to changes in the standard of care for lung cancer patients prohibited successful completion of the proposed trials and, thus, the trials were deferred as previously described. Our revised aim is now focused on the timely development of the risk model (see Aim 3 / Revised Task 3). The retrospective specimens will be used for the proposed analyses and correlated with the available clinical data.

Summary of Research Findings

As noted previously, the celecoxib trial was closed, and the proposed erlotinib trial was deferred. We will continue to focus our efforts on the productive analysis of acquired samples from the clinical trial and from our tissue bank, leading to the timely development of a lung cancer risk model.

Aim 3 Develop a lung cancer risk model to help predict the likelihood of development of relapse or new smoking related primary tumors

Patients with a history of smoking and a prior surgically resected stage I/II head and neck or lung cancer are at high risk for cancer recurrence or SPTs. There are no standard interventions that have been proven to help reduce the risk of cancer occurrence. A Gail risk model implemented in the initial management of breast cancer screening has proven useful and has helped with early detection and more stringent follow-up in the higher risk cohorts. Patients enrolled in the Vanguard trial will have aggressive post-operative follow-up with analysis including frequent serologies, bronchial specimens and CT scanning. Trends in these multiple biomarkers will be analyzed and used to develop a predictive model. Establishing a risk model will eventually help identify patients who may be at higher risk for lung cancer development and promote earlier interventions for prevention.

As noted previously, we have revised Aim 3 to develop a lung cancer risk model to help predict the likelihood of cancer recurrence and second primary tumor (SPT) development utilizing

clinical, pathologic and biomarker information obtained prospectively and retrospectively from the high-risk population of patients.

Summary of Research Findings

The clinical database and related data entry was completed via both patient record review and through the distribution of 300 IRB-approved letters to patients requesting contact to acquire and record additional long-term follow-up data. The ReVitalization project has also completed the pathological and clinical database of over 500 patients with prior resected lung cancer. In a close collaboration between VITAL project investigators and Cores B and C, preliminary statistical analyses have been performed on data from 542 patients whose specimens were studied for 21 markers (IGF1R, IGFBP3, InsulinR, pAKT, pIGF1R, pSRC, pmTOR, pAMPK, pEGFR, pS6, FEN1, MCM2, MCM6, SFN, TPX2, UBE2C, CASK, CD51, CXCR2, EpCAM, SPP1) to identify key markers for potential inclusion in the risk model development (see reports in Projects 2 and Cores B and C in this application for further detail).

Key Research Accomplishments

- Increased number of evaluable patients to 38, with 27 patients remaining on study.
- Continued to collect patient clinical data and tissues for distribution to support research projects in the VITAL grant.
- Completed tissue analysis of patients who have completed their 12-month bronchoscopy.
- Performed statistical analyses on data from 542 patients (including 10 from the Vanguard trial) whose tumor specimens were analyzed for 21 markers.
- Completed tissue identification and clinical data collection of over 500 archived tissue specimens from the pathology database for the ReVITALization plan.

Conclusion

Completion of the development of the pathological and clinical database for the ReVITALization Proposal was a key milestone for this program. The analysis of the tissue for important biomarkers has been completed, and results were correlated with the clinical outcomes. With expertise from the Pathology and Biostatistics Cores, we are currently in the development of a risk model for development of SPT and recurrence. This model has served as an example for other ongoing projects to effectively integrate pathology and clinical databases. A six-month no-cost extension has been requested to allow completion of the statistical risk modeling, the final component of proposed work for this grant.

Project 2: Identification of Biomarkers of Response to Chemoprevention Agents in Lung Epithelium

(PI and co-PIs: Ignacio Wistuba, M.D., Reuben Lotan, Ph.D., John Minna, M.D.)

Lung cancer continues to be the deadliest among all cancers in the United States with over 165,000 deaths annually for the last few years and an overall 5-year survival rate of less than 15% (Jemal et al., 2007). Early detection of premalignant lesions or tumors appears to be an efficient approach to reducing the morbidity and mortality from lung cancer because the survival of early stage lung cancer patients is much better than that of patients with advanced cancers. Therefore, new strategies for the early diagnosis, prevention and treatment of this dreadful disease are urgently needed (Wistuba and Gazdar 2006; Sato et al., 2007). The development of early detection tools for lung cancer requires improved molecular testing by identification and understanding of early events in the multi-step process of lung carcinogenesis, which involves

the accumulation of genetic and epigenetic alterations over the long course of exposure to carcinogens such as tobacco smoke (Mao, 2002; Wistuba and Gazdar 2006; Sato et al., 2007). To date, there are no validated biomarkers for early detection. Moreover, one or a few genes may not provide sufficient specificity given the multi-factorial process of lung carcinogenesis and heterogeneous nature of lung cancer. Thus, the effort to search for more specific and sensitive biomarkers of early lung cancer is warranted. The development of high-throughput gene expression analyses, e.g., DNA-chips or microarrays, provides opportunities to define biomarkers (signatures) of risk of cancer development. During the last few years, several studies reported molecular classification of human lung carcinomas on the basis of gene expression and described numerous putative biological markers of cancer (Meyerson et al., 2004). However, only limited number of studies has attempted to identify genes that are modulated at early stages of human lung carcinogenesis such as premalignant state because of the limited availability of premalignant lung tissues suitable for RNA extraction. We hypothesized that immortalized, transformed, and tumorigenic human bronchial epithelial cell (HBEC) line models will have similar abnormalities in gene expression profiles as premalignant and malignant tissues *in vivo*. Therefore, such cell models will be useful to identify markers of early disease.

We proposed to use genomic and proteomic analyses to identify changes in gene expression (including mRNA and miRNAs) and proteins which correlate/associate with cancer risk in the carcinogen damaged aerodigestive tract field and also use these signatures to monitor the response of this field to chemoprevention. We will develop and use a model HBEC system to study the effect of specific oncogenic changes and also the response of these manipulated HBECs to various carcinogenic and chemoprevention agents. Thus, we will determine modifications of these changes by chemopreventive agents in premalignant cells *in vitro* and to use probes for the modified genes and proteins to analyze tissue specimens from individuals participating in the chemoprevention clinical trials.

Following the relocation of the original Project PI, Dr. Li Mao, Dr. Ignacio Wistuba (Director, Biomarker Core) was appointed by the VITAL Principal Investigator, Dr. Waun Ki Hong, to assume ongoing leadership of this project.

Aim 1 Develop immortalized human bronchial epithelial cell cultures using a subset of patient tissue specimens collected in Project 1 and characterize the expression profiles of these cells using oligonucleotide based microarrays.

The main goal of this aim of this project is to establish these cultures from lung cancer patients and persons without lung cancer, including those patients entered onto the clinical trial described in Project 1, and to characterize their gene expression profiles.

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

Aim 2 Characterize effects of the chemo preventive agents used in Project 1 on cell proliferation and apoptosis in the immortalized human bronchial epithelial cell cultures developed in Specific Aim 1.

We will determine the potential role of different chemopreventive agents [e.g., celecoxib, N-[4-hydroxyphenyl]retinamide (4-HPR), Iressa (gefitinib), and SCH63663] alone or in combination with one another for their effects on cell proliferation and apoptosis in cell cultures established in

Aim 1. We will also determine the relative sensitivity among the various cell cultures to each of the agents by determining the 50% growth inhibitory concentration (IC₅₀).

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

Aim 3 Identify gene expression and protein “signatures” which reflect lung tumorigenesis and sensitivity or resistance to chemo preventive regimens proposed in Project 1, and to validate the signatures and to determine their biological importance in precancer cell models of lung cancer.

Summary of Research Findings

A. Perform high density array CGH analysis of lung cancers to identify regions commonly amplified in lung cancer.

This subaim was completed as previously reported.

B. Study of the expression of all 48 nuclear hormone receptors (NRs) reveals tumor-specific differences in microdissected tumor and normal lung epithelium.

This subaim was completed as previously reported.

Aim 4 Develop techniques to assess these molecular signatures in tissue specimens and serum obtained in Project 1, and assess the relevance of these molecular signatures as *in vivo* biomarkers using baseline and post-treatment specimens.

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

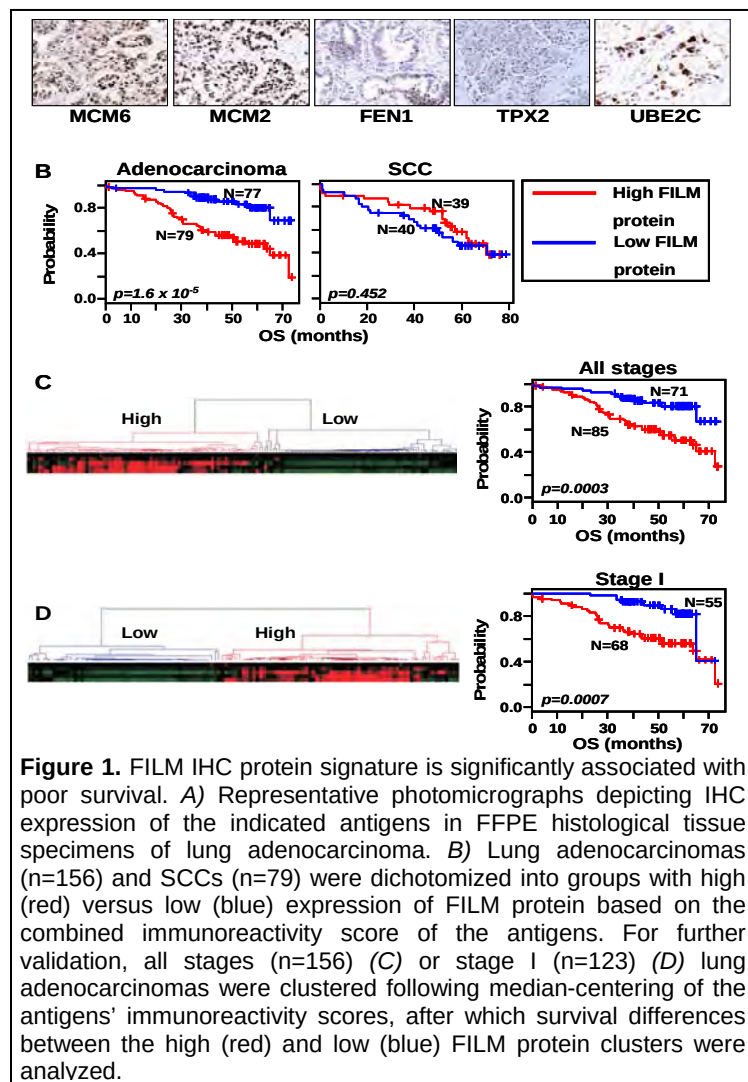
Aim 5. Identify gene expression signatures that characterize progression from immortalized to transformed to tumorigenic human bronchial epithelial cells based on already available high-throughput gene expression microarray data and validate these signatures using tissue microarrays (TMAs) containing normal bronchial epithelium, hyperplasia, squamous metaplasia, dysplasias, squamous cell carcinomas, atypical adenomatous hyperplasia, and adenocarcinomas (Years 4-5).

Summary of Research Findings

In the last report, we produced data depicting the derivation of a five-gene signature (FILM, five-gene *in vitro* lung carcinogenesis model) that was effective in predicting the survival of lung adenocarcinoma patients upon analysis of publicly available NSCLC microarray datasets. Importantly, the signature was effective in predicting the survival of stage I lung adenocarcinomas and was not prognostic in lung squamous cell carcinomas (SCCs). The signature is comprised of genes (UBE2C, MCM2, MCM6, TPX2, and FEN1) that we had previously shown, by gene-interaction network analysis, to be differentially expressed and functionally modulated between the components of an *in vitro* lung carcinogenesis model encompassing normal, immortalized, transformed, and tumorigenic lung epithelial cells. In the past year, we assessed the immunohistochemical (IHC) expression of the protein products of the aforementioned genes in NSCLC tissue microarrays (TMAs) of formalin-fixed and paraffin-embedded (FFPE) tissues (Dr. Ignacio Wistuba, Director of Thoracic Molecular Pathology Laboratory). We deemed this approach to be important, as the levels of functionally relevant proteins need not match the expression level of transcripts. The IHC analysis was important to

perform to complement the mRNA data described above, both from the point of validating the finding at the protein level, and for testing whether a corresponding FILM protein signature will be useful for prospective prognostic analysis of FFPE NSCLC specimens.

We analyzed the protein expression of the FILM signature by IHC analysis in FFPE histological

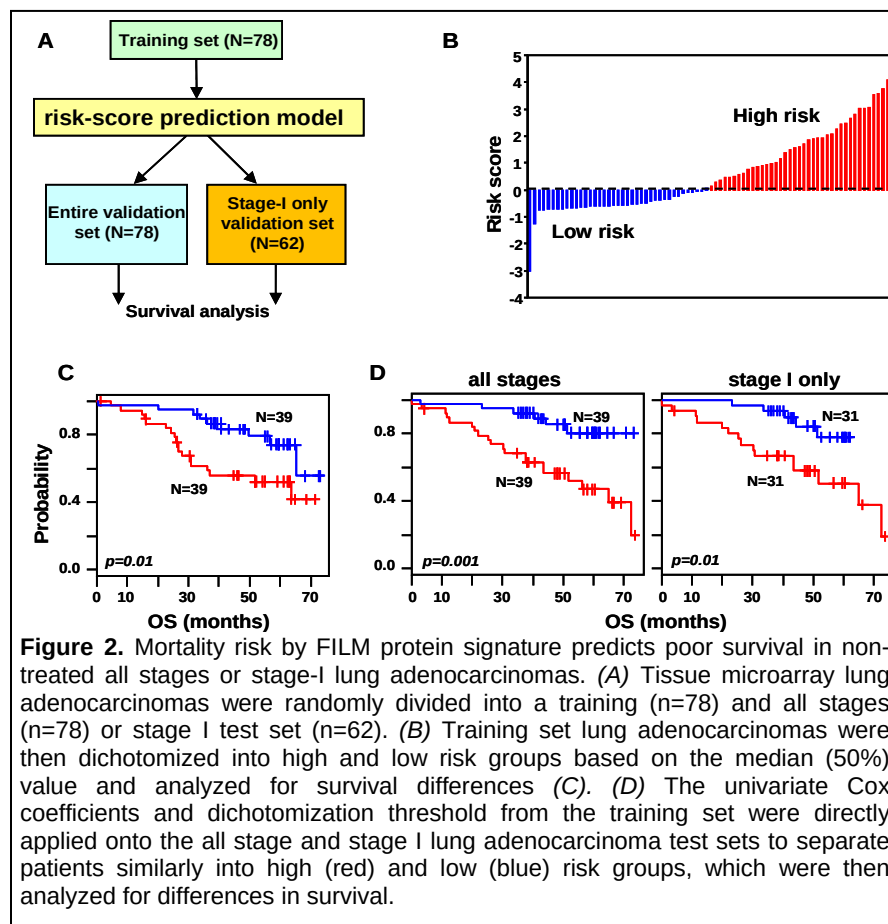


tissue specimens obtained from 156 lung adenocarcinoma and 79 SCC patients who did not receive any therapy before or after tumor resection (Figure 1). A combined total immunoreactivity score for FILM protein expression was computed, and patients were then dichotomized based on the median FILM protein score. In accordance with the findings obtained with the transcript signature, all-stages ($p < 0.001$, HR=3.6, 95% CI=1.9-6.8) or stage I-only ($p < 0.001$, HR=3.5, 95% CI=1.7-7.5) lung adenocarcinoma patients with higher expression of FILM protein exhibited significantly poorer survival compared to patients with lower expression. In contrast, the protein signature was not prognostic in lung SCC (Figure 1B). To confirm the prognostic capacity of the FILM protein signature, all stages (Figure 1C) or stage I (Figure 1D) lung adenocarcinoma patients were analyzed by hierarchical cluster analysis by average linkage based on the centered expression of the immunoreactivity scores of FILM protein expression. Two clusters were identified with dissimilar expression of FILM protein. All

stages or stage I-only lung adenocarcinoma patients in the cluster with higher FILM protein (high cluster) exhibited significantly poorer survival compared to patients in the low FILM protein cluster (both $p < 0.001$) (Figures 1C and 1D).

We then determined to further confirm the robustness of the FILM protein signature in predicting the survival of lung adenocarcinoma patients. Lung adenocarcinoma patients were randomized into a training (n=78) set and an all-stages (n=78) or stage-I only (n=62) test set (Figure 2A). Risk scores were computed based on the FILM proteins centered immunoreactivity scores and Cox coefficients in the training set (Figure 2A). Patients were then dichotomized based on the median (50%) risk score into high versus low risk groups (Figure 2B) and subsequently analyzed for survival differences. Lung adenocarcinoma patients in the training set and predicted to be at high mortality risk based on FILM protein exhibited significantly poorer survival than patients at low risk ($p = 0.01$) (Figure 2C). Cox regression coefficients and

dichotomization cut-off threshold generated from the training set were then directly applied to the test set (n=78) or only stage-I patients within the set (n=62) and analyzed similarly. In



accordance, all stages (p=0.001, HR=4.0, 95% CI=1.6-9.8) or stage I (p=0.01, HR=3.4, 95% CI=1.4-9.1) lung adenocarcinoma patients predicted to be at high mortality risk based on FILM protein exhibited significantly poor survival (Figure 2D and Table 1).

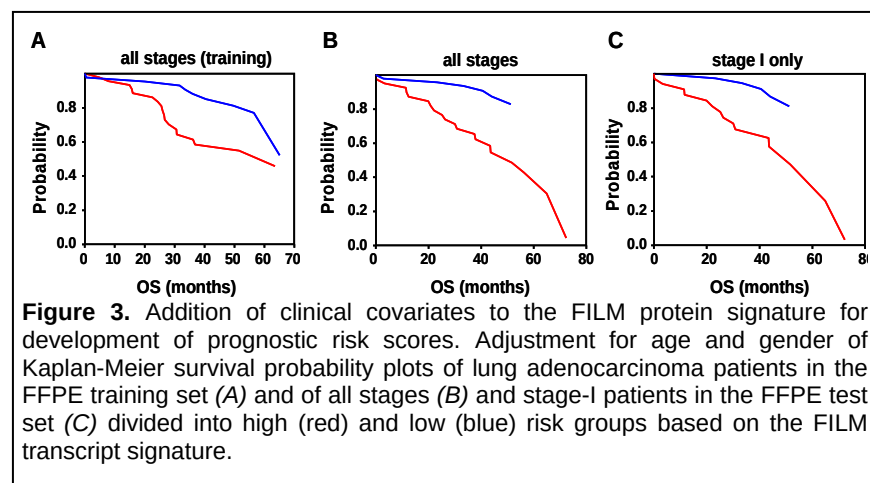
Moreover, multivariate Cox hazard regression analyses demonstrated that risk assessed by FILM protein was an independent predictor of survival in all stages (p=0.005, HR=5.78, 95% CI=1.7-20) almost similar in connotation to the most significant variable, age (p=0.002, HR=1.1, 95% CI=1.03-1.15) (Table 1). In addition, FILM protein risk was also an independent, and

better than IB stage, predictor of survival in stage I patients (p=0.01, HR=6, 95% CI=1.4-26.2) (Table 1). Inclusion of clinical covariates (age and gender) enhanced the capacity of the FILM protein signature to separate patients with poor survival from those with excellent survival (Figure 3). When comparing at 50 months follow-up and onwards, and following adjustment of survival probability plots for age and gender, no events or deaths were noted in low risk patients in contrast to survival analysis without the clinical covariates.

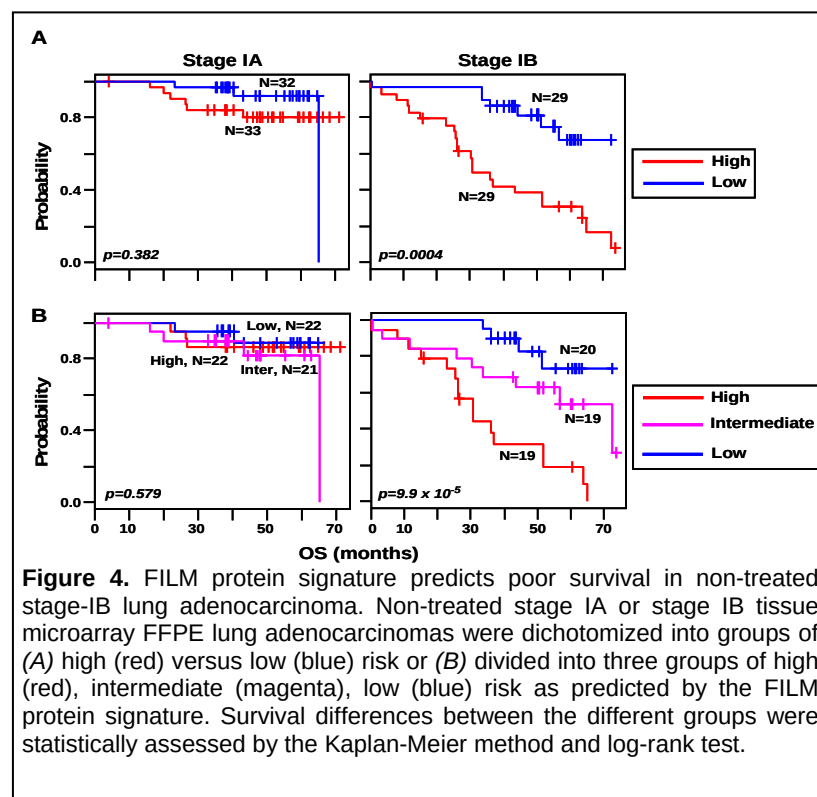
All stages				
	Univariate		Multivariate	
	H.R. (95% C.I.)	p-value	H.R. (95% C.I.)	p-value
High risk	3.998 (1.575-9.346)	0.001	5.776 (1.668-20.002)	0.005
Stage III vs I, II or IV	2.505 (0.731-8.583)	0.144	7.011 (1.467-33.506)	0.018
Poor differentiation	1.401 (0.580-3.383)	0.453	0.814 (0.176-3.762)	0.793
Solid subtype	1.007 (0.994-1.020)	0.275	1.003 (0.983-1.023)	0.777
BAC subtype	0.997 (0.983-1.012)	0.71	1.014 (0.995-1.034)	0.14
Never vs ever smoker	0.969 (0.386-2.432)	0.946	1.366 (0.419-4.453)	0.605
Male gender	1.816 (0.823-4.003)	0.139	1.226 (0.491-3.062)	0.662
Age	1.032 (0.987-1.079)	0.161	1.088 (1.031-1.149)	0.002

Stage I				
	Univariate		Multivariate	
	H.R. (95% C.I.)	p-value	H.R. (95% C.I.)	p-value
High risk	3.359 (1.392-9.110)	0.011	6.049 (1.399-26.154)	0.01
Stage IB vs IA	2.556 (0.917-7.124)	0.072	3.154 (0.977-10.191)	0.055
Poor differentiation	2.135 (0.773-5.900)	0.143	1.034 (0.187-5.722)	0.97
Solid subtype	1.010 (0.996-1.025)	0.161	0.997 (0.972-1.026)	0.921
BAC subtype	0.996 (0.988-1.012)	0.646	1.007 (0.988-1.026)	0.459
Never vs ever smoker	0.728 (0.210-2.517)	0.616	0.770 (0.148-4.024)	0.757
Male gender	2.210 (0.071-5.004)	0.095	1.940 (0.597-6.209)	0.209
Age	1.037 (0.984-1.094)	0.173	1.060 (0.990-1.135)	0.09

Table 1. Film protein risk is an independent predictor of poor survival in stage-I lung adenocarcinomas. HR, Hazard ratio; CI, confidence interval.



higher risk exhibited significantly poorer survival than patients at low risk ($p < 0.001$, HR=4.2, 95% CI=1.8-9.9) (Figure 4A). Similarly, when patients were categorized into groups of low, intermediate, and high risk, the FILM protein signature was able to further separate stage-IB patients with poor survival from those with excellent survival ($p < 0.001$, HR=7.26, 95% CI=2.40-21.98) (Figure 4B). These results demonstrate that the FILM protein signature, like its transcript version, is valuable for predicting the survival of lung adenocarcinoma patients and that the protein signature may specifically be valuable for identifying stage I or IB patients who may benefit from adjuvant therapy.



Since the FILM protein signature was effective in stage I lung adenocarcinoma prognosis, we tested its prognostic capacity in stage IA and stage IB patients separately. When patients were dichotomized based on mortality risk computed by FILM FFPE protein signature, stage IB but not stage IA patients with

It is worthwhile to note that we had previously started to assess the prognostic efficacy of the protein products of the FILM transcript signature independently by each protein marker alone (shown in the previous report). The prognostic efficacy of the protein signature was superior in comparison to that of each IHC protein marker alone. Therefore, further independent studies are warranted to externally validate the potential clinical use of this five-gene/protein signature, and in particular, the prognostic capacity of the FILM protein FFPE signature towards translation to the clinic for identifying non-treated stage-I and specifically IB lung adenocarcinoma

patients with poor survival that will benefit from adjuvant therapy.

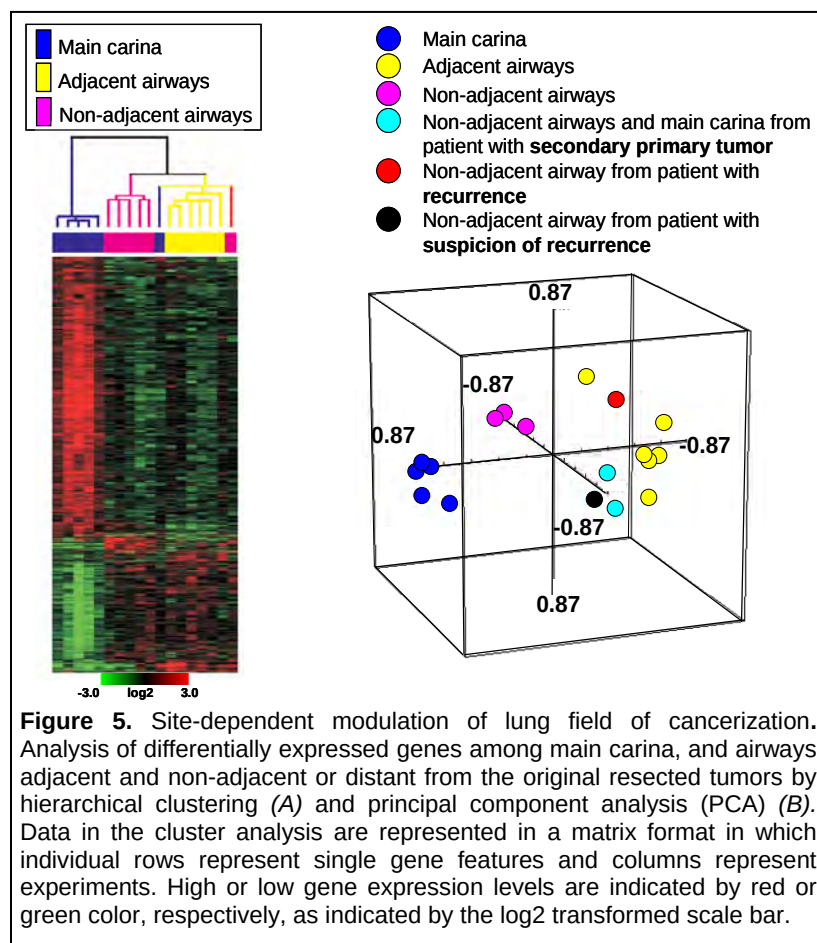
Aim 6. Identify gene expression signatures in bronchial brush specimens from the 50-60 patients enrolled in the Vanguard study using high-throughput genomics approach (Years 4-5).

Summary of Research Findings

In the previous year, we reported results from the analysis of the temporal patterns of gene expression in the cancerization field and the global differential gene expression among different regions within the NSCLC field of cancerization. We analyzed the transcriptomes of bronchial brushings obtained at baseline and at different time points (baseline, 12, and 24 months following surgery) from a limited set of early stage NSCLC patients who were definitively treated by resective surgery. We identified, as indicated in the previous report, genes that were differentially expressed among airways from different regions of the lung at various distances from the original resected tumor (site-dependent molecular modulation of the field of cancerization) as well as genes that were differentially modulating in lung airways at 24 and 12 months compared to airways at baseline following surgery (i.e. time-dependent molecular modulation of the field of cancerization). Moreover, we also showed in the previous report that gene-interaction networks mediated by the extracellular-regulated kinase 1 and 2 (ERK1/2) were significantly modulated (upregulated in both expression and predicted function) in airways (main carina) sampled at 24 months compared to those sampled at baseline.

We analyzed the site-dependent modulation of airway gene expression in the NSCLC field of cancerization by molecular profiling analysis of main carina, adjacent (ADJ), and airways non-

adjacent or distant (NON-ADJ) (n=6 each) from the original resected tumors sampled from the same six patients. Supervised clustering analysis of the differentially expressed genes (n=931) divided the different anatomically located airways into discrete clusters (Figure 5A). Interestingly, principal component analysis (PCA) of the differentially expressed genes demonstrated that while ADJ airway samples grouped separately and closely together, one MC and 3 NON-ADJ airway samples resided closely with ADJ samples, which were then found to originate from 3 pts with evidence of recurrence, SPT, or suspicion of recurrence (Figure 5B). Moreover, pathways analysis of the spatially modulated genes revealed that gene-networks mediated by nuclear factor-kappa B (NF- κ B) and ERK1/2-



mediated were most significantly elevated ($p < 0.001$) in function in ADJ airway samples versus main carinas (Figure 6). These findings highlight expression signatures and pathways (ERK1/2 and NF- κ B) in a “cancerization field” that may drive lung cancer pathogenesis and be associated with recurrence or SPT development in ES NSCLC pts and thus useful for derivation of biomarkers to guide personalized prevention strategies.

These promising preliminary findings highlight gene expression signatures that are modulated in a time- and location-dependent manner in the lung cancerization field. Therefore, we seek to expand these exciting findings, outside the scope of this grant proposal, by analyzing both mRNA and miRNA global differential expression patterns in more detail in NSCLC cancerization fields, by simultaneous mRNA and miRNA profiling analysis of airways successively distant from the tumors as well as the tumors and adjacent normal parenchyma. This powerful and integrative approach will allow us to more directly query molecular patterns in the airways and especially those that are common in both tumors and the airways from the same patients. Biologically, this will increase our understanding of the field of cancerization separately in lung adenocarcinomas and SCCs and thus further our comprehension of the molecular pathogenesis occurring in NSCLC. From a translational perspective, this work will allow us to generate molecular risk profiles that can not only prospectively predict recurrence, but also used to detect lung cancer in individuals at high risk of developing the disease (e.g., heavy smokers). Towards this, under the leadership of Dr. Ignacio Wistuba, we used the aforementioned preliminary analysis of the Vanguard field of cancerization molecular profiles and submitted a grant application for the previous fiscal year DoD lung cancer research program consortium award that we were awarded (W81XWH-10-1-1007) along with three different sites/study centers. One of the goals of this recent grant award is to cross-integrate RNA signatures (tumor/normal airway differentials) with the available gene expression data we have already acquired from the Vanguard cohort.

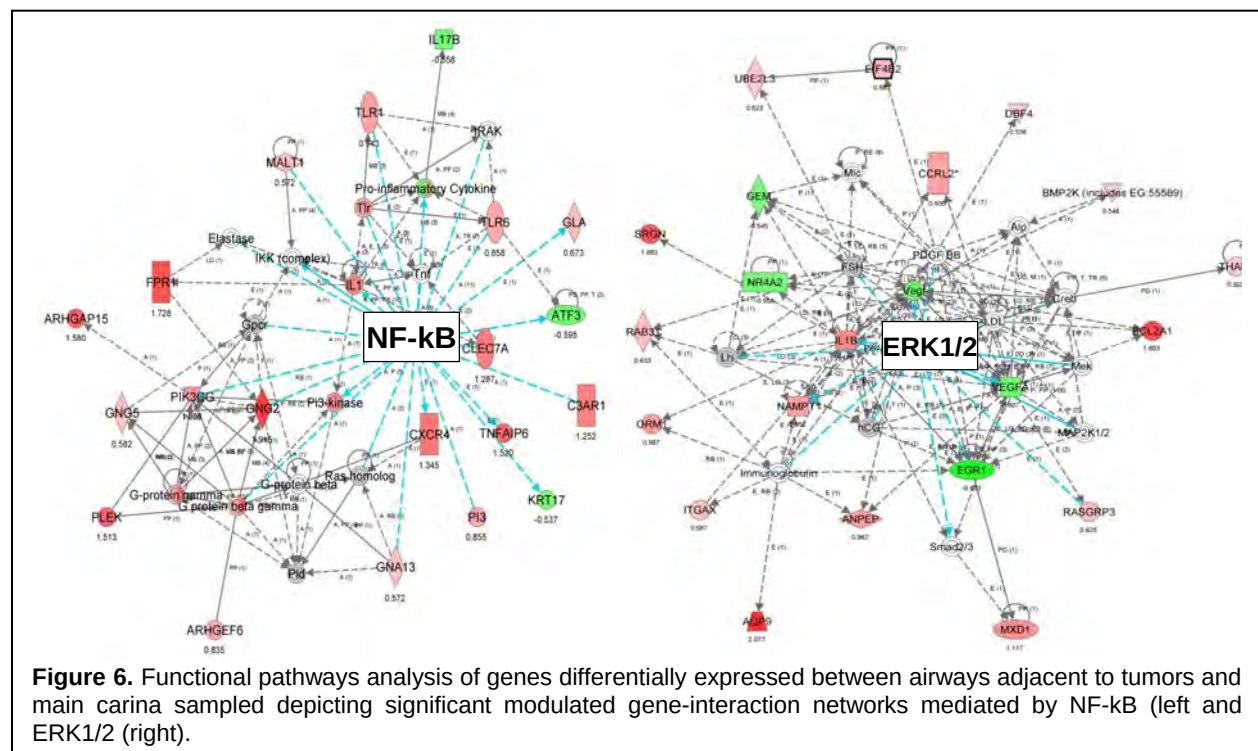


Figure 6. Functional pathways analysis of genes differentially expressed between airways adjacent to tumors and main carina sampled depicting significant modulated gene-interaction networks mediated by NF- κ B (left) and ERK1/2 (right).

In addition and based on the findings outlined in the previous report, we decided to expand the analysis to cover as many eligible Vanguard patients airway samples as possible. We identified nineteen NSCLC patients with available brushings at five to six different airway regions in the lung that have also been sampled at baseline up to 24 or 36 months following the baseline time point yielding 320 samples. Total RNA was isolated from all samples which were subsequently amplified for preparation of target DNA and eventually biotin-labeled fragmented DNA for hybridization onto Affymetrix Human Gene 1.0 ST arrays. The samples are currently being hybridized onto arrays and we anticipate completion of the raw data generation in mid- to late-January 2011 and completion of the data analysis by end of the 2011 spring quarter. As noted previously, a 6-month no-cost extension for this grant has been requested, and further findings from this work will be reported at completion of that period.

Key Research Accomplishments

- Validated the FILM five-gene signature, initially analyzed using public datasets, at the protein level by immunohistochemistry using an independent set of FFPE NSCLC tissue.
- Showed that the FILM protein signature, like its transcript counterpart, is effective in predicting survival in lung adenocarcinoma and demonstrated prognostic specificity towards this subtype of NSCLC.
- Demonstrated that the FILM protein signature is powerful in predicting the survival of early stage (stage-I or stage-IB) lung adenocarcinoma patients who did not receive any form of therapy.
- Analyzed genes differentially expressed among airways sampled at different anatomical locations of the lung and at various distances from the original resected tumors and identified site-dependent differentially expressed genes and pathways in the field of cancerization that have potential value in predicting recurrence or second primary tumor development.
- Processed the remaining of eligible Vanguard patient airway samples (n=320) for molecular profiling analysis to be completed in the spring of 2011.
- Awarded a DoD Lung Cancer Research Program Consortium award (PI Ignacio Wistuba, Director, Thoracic Molecular Pathology Laboratory) based on the preliminary findings of the Vanguard molecular profiling analysis to expand the Vanguard study findings and analyze the field of cancerization in more unprecedented detail.

Conclusions

Our findings highlight an immunohistochemical protein signature (FILM protein signature), corresponding to a five-gene transcript signature we previously derived, that is effective in predicting the survival of non-treated stage-I and in particular stage-IB lung adenocarcinoma patients. Thus, this five-protein FFPE signature shows promise as a clinically relevant classifier in distinguishing early-stage patients who will benefit from adjuvant therapy from those who can be spared therapy.

In addition, our findings demonstrate that gene expression in the lung airway field of cancerization is modulated in a site- (adjacent versus distant from original primary tumor) and time- (following resection of tumor) dependent manner. Moreover, differential gene expression in the airway field of cancerization encompasses canonical cancer-associated cellular pathways that may be 1) biologically crucial for early pathogenesis of lung cancer and 2) clinically relevant for prediction of recurrence and second primary tumor and therefore development of personalized targeted strategies for chemoprevention of the disease.

Project 3: Premalignant Bronchial Epithelia: Molecular and Cellular Characterization of Lung Tumorigenesis

(PI and co-PIs: Walter Hittelman, Ph.D., Ja Seok Koo, Ph.D., Rodolfo C. Morice, M.D.)

Aim 1 Identify and characterize differentially expressed genes in the LIFE bronchoscopy-identified abnormal areas of the bronchial epithelia of enrolled subjects in VITAL trials.

Previous studies have shown that bronchial regions that appear abnormal by light-induced fluorescence endoscopy (LIFE) bronchoscopy show increased genetic changes when compared to normal-appearing sites, even if there are no differences in histological appearance. Since LIFE-positive lesions are at increased risk for cancer development, especially when they contain particular genetic alterations, we hypothesize that these LIFE-positive sites represent lesions at an early stage of tumorigenesis and may differentially express genes important for driving tumorigenesis. Thus, comparative gene expression analyses between LIFE-positive and LIFE-negative sites within the same individual may provide a filter for identifying genes whose levels of expression are important for driving tumorigenesis.

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

Aim 2 Establish an organotypic model system that mimics *in vivo* interactions between normal, premalignant, and malignant bronchial epithelial cells in the lung using cells derived from bronchial biopsies and immortalized bronchial cells.

Our prior studies using chromosome *in situ* hybridization to visualize genetic changes in the bronchial epithelium of current and former smokers suggested that, over years of tobacco smoke exposure, the combination of accumulating genetic damage, ongoing tissue damage, and wound healing results in a mosaic of evolving clonal outgrowths throughout the bronchial epithelium. To better understand the molecular basis of preferential outgrowth of more advanced bronchial epithelial clones, we proposed to utilize a cell culture model whereby normal and abnormal bronchial epithelial cells are grown on collagen or stromal cell-coated, suspended filters and exposed to an air-liquid interface. This organotypic culture environment mimics lung stratified epithelium, complete with basal cells, ciliated columnar cells, and mucus-producing goblet cells. Our group has extended this model system by tagging cell populations with fluorescent probes (e.g., green fluorescence protein, or GFP) that allows us to carry out live cell imaging of mixed clonal populations. This model system permits characterization of the ability of more advanced bronchial epithelial cell populations to expand on the growth surface at the expense of less advanced bronchial epithelial cell populations.

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

Aim 3 Determine the mechanisms of genetic instability and elucidate the signaling pathways associated with clonal outgrowth of premalignant and malignant bronchial epithelial cells using the organotypic model system.

Our prior studies using chromosome *in situ* hybridization to visualize genetic changes in the bronchial epithelium of current and former smokers suggested that current tobacco exposure

was associated with increased levels of ongoing genetic instability (i.e., chromosome polysomy). Upon smoking cessation, while the initiated clonal outgrowths appeared to be maintained over tens of years, the levels of ongoing genetic instability appeared to decrease gradually during the first year following smoking cessation. However, in some cases, we observed evidence for ongoing genetic instability in the bronchial epithelial cells even 10-20 years following smoking cessation. Since nearly half of the newly diagnosed lung cancer cases occur in former smokers, we felt that this finding suggested that an ongoing intrinsic process of genetic instability might exist in the lungs of some former smokers that drives continued genetic evolution toward lung cancer even after cessation of extrinsic carcinogenic exposure.

Our working hypothesis is that years of tobacco exposure induces a chronic damage and wound healing cycle that results both in the accumulation of genetic alterations in the epithelial cells that influences both chromosome stability mechanisms (e.g., loss of cell cycle checkpoint and cell loss mechanisms through loss of p16 expression, p53 mutations, cyclin D1 overexpression, etc) and creates a poor growth environment (e.g., altered stromal signals). The goal of this specific aim was to utilize the lung organotypic model to address this hypothesis *in vitro* utilizing bronchial epithelial cells derived from LIFE bronchoscopically identified “abnormal” and “normal” regions of the lung of current and former smokers participating in the clinical trial of Project 1.

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

Aim 4 Characterize the impact of chemopreventive and/or chemotherapeutic agents on early lung tumorigenesis events in reconstructed bronchial epithelium and in the bronchial biopsies of subjects entered onto the clinical trials in Project 1.

The goals of the first three specific aims of this project are essentially to develop and utilize the lung organotypic culture model to identify the factors that control ongoing clonal expansion and genetic instability in the lungs of current and former smokers. The idea behind this fourth specific aim is to integrate the information garnered from the first three specific aims to identify targeted strategies to slow preferential outgrowth of more advanced bronchial epithelial cells and to decrease the levels of ongoing genetic instability. We also proposed to determine whether treatment of these organotypic cultures with the chemopreventive agents used in the clinical trial of Project 1 would slow these aberrant properties *in vitro* and whether results obtained in the organotypic culture model reflected that seen in the lungs of the participants in the clinical trial.

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

Aim 5. Identify gene expression signatures in epithelial cells detected by LIFE bronchoscopy that determine aggressiveness.

This new aim under the ReVITALization plan will be performed in conjunction with Dr. Wistuba, Core C Director. Our preliminary results from our research in VITAL have shown that epithelial cells isolated from bronchial biopsies of LIFE-abnormal mucosa can be characterized as more aggressive (invasive and migratory) than those of LIFE-normal mucosa. Microarray analysis suggested that several CXCL-chemokine signaling pathways are mainly deregulated in LIFE-abnormal cells. Moreover, we identified that pro-angiogenic ELR+ (glutamic acid, lysine and

arginine motif) chemokines were strongly upregulated by inflammatory cytokines in lung cancer cells.

Summary of Research Findings

This aim is completed. The research findings were submitted in two publications (attached in Appendix 3);

1. Prevention of Bronchial Hyperplasia by EGFR Pathway Inhibitors in an Organotypic Culture Model. Jangsoon Lee^{1,4}, Seung-Hee Ryu^{1,5}, Shin Myung Kang^{1,6}, Wen-Cheng Chung¹, Kathryn Ann Gold², Edward S. Kim¹, Walter N. Hittelman³, Waun Ki Hong¹, and Ja Seok Koo; (Accepted, *Cancer Prevention Research*).
2. CXCR2 expression in tumor cells is associated with a poor prognosis in a large set of non-small-cell lung cancer. Pierre Saintigny¹, Diane Liu², J. Jack Lee², Yuan Ping³, Carmen Behrens³, Luisa M. Solis Soto³, John V. Heymach¹, Edward S. Kim¹, Jonathan M. Kurie¹, Waun Ki Hong⁴, Ignacio I. Wistuba³, and Ja Seok Koo. (Submitted, 2011 AACR Annual Meeting).

A summary of these research findings is presented below.

CXCR2 plays an important role in inflammation, and stimulation of CXCR2-expressing endothelial cells by ELR+ CXC chemokines promotes angiogenesis. Our goal was to study the expression of CXCR2 by tumor cells and its impact on prognosis in NSCLC. CXCR2 expression was determined using immunohistochemistry and microarray with a large set (n=458) of NSCLC tissue . The association between cytoplasmic CXCR2 (cCXCR2) expression in tumor cells and clinico-pathological factors as well as survival was analyzed. Distribution of CXCR2 and its ligands (IL8, CXCL1, CXCL2, CXCL3, CXCL5, CXCL6 and CXCL7) gene expression was studied using publicly available gene expression profiles from 52 NSCLC cell lines (GSE4824) and 444 lung adenocarcinomas (NCI Director's Challenge). To summarize the effect of CXCR2/CXCR2 ligands biological axis, Principal Component Analysis (PCA) and unsupervised hierarchical clustering were performed using CXCR2 and its ligands gene expression in both cell lines and lung adenocarcinomas. The first Principal Component (PC1) was correlated (Pearson) with the whole genome in 52 NSCLC cell lines. All genes were ranked according to their correlation with PC1, and used for Gene Set Enrichment Analysis (GSEA) "pre-ranked" analysis. Using the median of expression to dichotomize the patients in a high versus low expression group, 238 (52.1%) tumors expressed high cCXCR2. No association was observed with gender, race, smoking habits, histology, and stage. High cCXCR2 was associated with overall survival [Hazard ratio (HR) 1.5696; confidence interval (CI)=1.176-2.096, p-value=0.002] and recurrence-free survival (HR 1.321; CI=1.027-1.698, p-value=0.030) in a univariate Cox proportional hazards (CPH) model. High cCXCR2 remained significant for overall in a multivariate CPH after adjusting for age, gender, histology, stage, neoadjuvant chemotherapy for overall survival (HR 1.465; CI=1.088-1.972, p-value=0.012) and a trend was observed for recurrence-free survival (HR 1.261; CI=0.973-1.633, p-value=0.080). Gene expression distribution of CXCR2 and its ligands were strikingly similar in cell lines and lung adenocarcinoma tissue. In both cases, hierarchical clustering showed a cluster mostly driven by CXCR2, CXCL5, and CXCL7, representing 20% of the samples. The KRAS and NFkB oncogenic pathways were the top 2 gene sets associated with PC1. Using the median as a cutoff, PC1 was associated with a worse overall survival in 444 lung adenocarcinomas (Log-rank P=0.006).

These findings strongly suggest that cCXCR2 expression in NSCLC tumor cells is frequent and associated with an adverse outcome. The CXCR2/CXCR2 ligands biological axis may be

associated with activation of the KRAS and NFkB signaling pathways, and poor prognosis, in lung adenocarcinomas.

In addition, our earlier studies showed that CREB is critical for mediating lung inflammatory response stimulated by cytokine and thus play substantial role in lung carcinogenesis. These research findings were previously reported in 3 papers:

1. Cyclic AMP-responsive element binding protein- and nuclear factor-kappaB-regulated CXC chemokine gene expression in lung carcinogenesis. Sun H, Chung WC, Ryu SH, Ju Z, Tran HT, Kim E, Kurie JM, Koo JS. *Cancer Prev Res (Phila)*. 2008 Oct;1(5):316-28.PMID: 19138976
2. Cyclic AMP response element-binding protein overexpression: a feature associated with negative prognosis in never smokers with non-small cell lung cancer. Seo HS, Liu DD, Bekele BN, Kim MK, Pisters K, Lippman SM, Wistuba II, Koo JS. *Cancer Res*. 2008 Aug 1;68(15):6065-73.PMID: 18676828.
3. Growth suppression of lung cancer cells by targeting cyclic AMP response element-binding protein. Aggarwal S, Kim SW, Ryu SH, Chung WC, Koo JS. *Cancer Res*. 2008 Feb 15;68(4):981-8.

Grants/programs awarded based on VITAL results

An R01 application was successfully submitted and awarded in the last year, titled “Role of CREB in Lung Cancer Development” (1R01CA126801; project period, 4/1/2010–3/31/2015; \$1,597,750 total direct costs). In this proposal, we hypothesized that aberrantly regulated CREB plays a critical role in the abnormal proliferation and survival of NSCLC cells. To test this hypothesis, we will pursue three specific aims: 1) To determine whether CREB promotes the pathogenesis of NSCLC by analyzing preneoplastic NSCLC lesions and modulating CREB expression and activity in an in vitro lung carcinogenesis model cell and in vivo animal models; 2) To identify the biologic and molecular consequences of modulating CREB activity using small molecule inhibitors in lung cancer development; and 3) To establish the genetic role of CREB in the development of lung cancer. These studies will increase our understanding of the mechanisms of abnormal survival and proliferation of NSCLC cells, in particular, the role of CREB in lung cancer development. We are hopeful that this study will identify a novel target for preventive and/or therapeutic strategies in patients with NSCLC.

Also based on the VITAL project findings, we recently submitted an R01 grant proposal entitled “Cytokine-CREB Network in Lung Disease” that is pending review. Earlier studies showed that IL-1b, one of several loci that are significantly associated with COPD and lung cancer, activates transcription factors, including NF-kB and cAMP response element-binding protein (CREB), and then increases expression of numerous proinflammatory cytokines, proangiogenic chemokines, and matrix metalloproteinases (MMPs). We also showed previously that CREB is critical for the survival and proliferation of lung cancer cells. Moreover, overexpression of CREB is a poor prognosis factor for lung cancer patients. Based on our preliminary findings and in concert with the known functions of cytokines and chemokine signaling systems in COPD and lung cancer, we established a working hypothesis that overly activated cytokine and chemokine signaling systems lead to activation of CREB, and that cytokine/chemokine to CREB signaling networks play a critical role in the development of COPD and lung cancer. Our proposal will test this hypothesis using animal models of COPD and lung cancer and also tissues and blood specimens from lung cancer patients.

Key Research Accomplishments

- Established an organotypic culture method to mimic bronchial hyperplasia/dysplasia *in vitro*. We also tested a potential preventive method that blocks hyperplasia/dysplasia using the EGFR inhibitor erlotinib.
- Demonstrated that chemokine receptor CXCR2 expression in tumor cells is associated with a poor prognosis in NSCLC.
- Demonstrated that CREB and NF- κ B-regulated CXC chemokine gene expression in lung carcinogenesis.
- Showed that CREB overexpression is associated with negative prognosis in never-smokers with NSCLC.
- Demonstrated growth suppression of lung cancer cells by targeting CREB.

Conclusion

Our studies produced several important findings that implicate a potential for future use in clinical settings. As a chemoprevention strategy, the combination of erlotinib with a MEK inhibitor may be more efficient in suppression of early abnormal changes in bronchial epithelium. In particular, we identified the chemopreventive effect of EGFR and MEK inhibitors, identified CXCR2 as a new potential therapeutic/preventive target, identified a small molecule inhibitor of a new class of anti-angiogenic method, and demonstrated that CREB could act as a novel therapeutic and preventive strategy. CXCR2 overexpression in NSCLC tumor cells is frequent and associated with an adverse outcome. The CXCR2/CXCR2 ligands biological axis may be associated with activation of the KRAS and NF κ B signaling pathways and poor prognosis in lung adenocarcinoma. Given the fact that there is no efficient method to treat NSCLC with KRAS mutation, targeting CXCR2 in combination with conventional chemotherapy or targeted therapeutics could lead to potential modes of treatment in the future. Targeting CREB using a small molecule inhibitor successfully blocked angiogenic effect of CXCLs, suggesting that targeting CREB may be a new anti-angiogenic strategy for lung cancer treatment and prevention. Further studies outside the scope of this grant are planned to translate these findings into clinical applications.

Project 4: Modulation of Death Receptor-Mediated Apoptosis for Chemoprevention

(Project Leader and co-leaders: Shi-Yong Sun, Ph.D.; Taofeek Owonikoko, M.D., Ph.D.)

The objective of Project 4 is to understand the role of death receptor (DR)-mediated apoptotic pathways in lung carcinogenesis, cancer prevention, and therapy in order to develop mechanism-driven combination regimens by modulating DR-mediated apoptosis for chemoprevention and therapy of lung cancer. It should be noted that Dr. Taofeek Owonikoko assumed the leadership of Project 4 following Dr. Fadlo Khuri's decision to step down due to increased administrative responsibilities; this administrative change was approved in November 2009. Following is a summary of our research progress:

Aim 1: To determine whether decoy receptor (DcR) and tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) expression are reduced or lost while DR remains largely expressed and whether procaspase-8 and FLIP expression and Akt activity are increased during lung carcinogenesis.

Summary of Research Findings

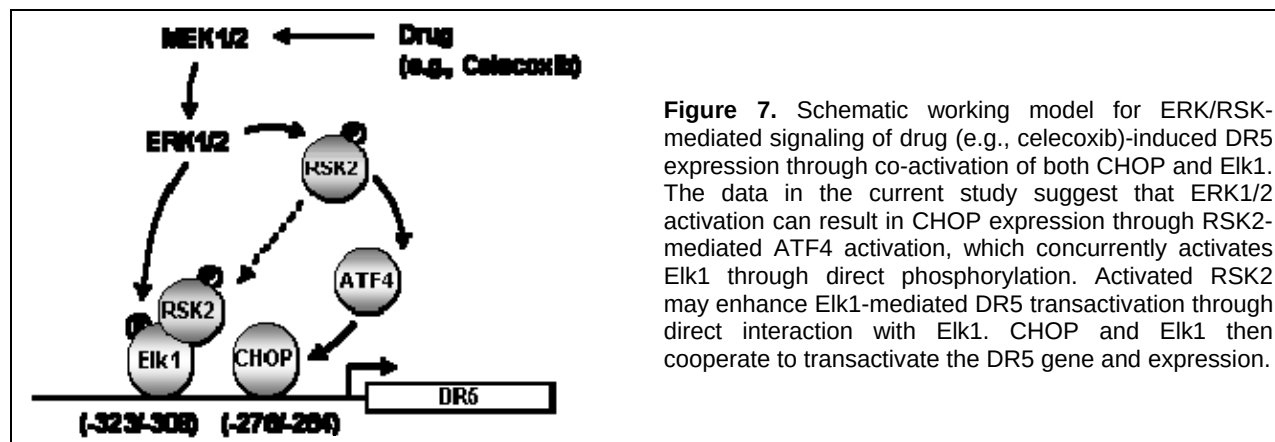
As an extension of the proposed aim, we have successfully constructed lentiviral-inducible expression systems for the following genes, whose products are key components in the extrinsic apoptotic pathway: DR4, DR5, DcR1, DcR2 and caspase-8. In addition, we also made lentiviral viral expression constructs for wild-type FADD, dominant-negative FADD and FADD-S194A (phosphorylation-resistant mutant). The expression of these constructs has been validated. These valuable resources will be very helpful and useful for our future studies in this direction to understand and dissect this pathway in regulation of cancer development and in mediation of cancer therapy.

Aim 2: To establish TRAIL-resistant cell lines from a TRAIL-sensitive lung cancer cell line and determine whether levels of DcRs, DRs, procaspase-8, TRAIL and FLIPs and Akt activity are altered and are associated with cell resistance to TRAIL and DR-inducing agents.

As noted in the previous report, we were not able to demonstrate that the TRAIL-resistant lung cancer cell lines exhibited cross-resistance to some DR-inducing agents. Alternatively, we have focused on addressing the question of whether these agents modulate the DR-mediated apoptotic pathway and, if so, how they modulate the DR-mediated apoptotic pathway and whether the modulations impact apoptosis by these DR-inducing agents.

Summary of Research Findings

ERK/RSK signaling positively regulates DR5 expression through co-activation of CHOP and Elk1. One of our long-term interests is to understand the general mechanisms by which DR5 expression is regulated. In our previous report, we showed that celecoxib and other agents induce DR5 expression. Using celecoxib as a DR5 inducer, we have revealed a novel ERK/RSK-dependent mechanism that regulates DR5 expression. Both CHOP and Elk1 are required for celecoxib-induced DR5 expression based on promoter deletion and mutation analysis and siRNA-mediated gene silencing results. Co-expression of both CHOP and Elk1 exhibited enhanced effects on increasing DR5 promoter activity and DR5 expression, indicating that CHOP and Elk1 cooperatively regulate DR5 expression. Since Elk1 is an ERK-regulated protein, we found that celecoxib increased the levels of phosphorylated ERK1/2, RSK2, and Elk1. Inhibition of either ERK signaling with a MEK inhibitor or ERK1/2 siRNA, or RSK2 signaling with a RSK2 inhibitor or RSK2 siRNA abrogated DR5 upregulation by celecoxib as well as by other agents. Moreover, these inhibitions suppressed celecoxib-induced CHOP upregulation. Thus, ERK/RSK-dependent, CHOP, and Elk1-mediated mechanisms are critical for DR5 induction. Additionally, celecoxib increased CHOP promoter activity in an ATF4-dependent manner and siRNA-mediated blockade of ATF4 abrogated both CHOP induction and DR5 upregulation, indicating that ATF4 is involved in celecoxib-induced CHOP and DR5 expression. Collectively, we conclude that small molecules such as celecoxib induce DR5 expression through activating ERK/RSK signaling and subsequent Elk1 activation and ATF4-dependent CHOP induction (Figure 7).



Oncogenic Ras and B-Raf increase DR5 expression involving co-activation of ERK/RSK and JNK signaling via enhancing CHOP, Elk1, and c-Jun-dependent transcription. Oncogenic mutations of Ras and B-Raf frequently occur in many types of cancers and play critical roles in cell transformation and oncogenesis. Interestingly, the oncogenic Ras has been documented to induce the expression of DR5 with undefined mechanisms. The above described findings on ERK/RSK regulation of DR5 led us to further ask whether this is also the mechanism by which the oncogenic Ras induces DR5 expression. Indeed, enforced expression of the oncogenic H-Ras, K-Ras (e.g., RasV12), or B-Raf (E600V) increased DR5 promoter activity and DR5 expression accompanied with activation of both ERK/RSK and JNK signaling pathways evidenced by increased levels of p-ERK, p-RSK, p-JNK, and p-c-Jun as well as their downstream proteins, CHOP, Elk1, and c-Jun (Figure 8). Gene silencing-mediated inhibition of ERK, RSK, or JNK activation blocked RasV12-induced DR5 expression (Figure 9). Moreover, knockdown of CHOP, Elk1, or c-Jun abrogated RasV12-induced DR5 expression as well. These data collectively indicate that both ERK/RSK and JNK signaling pathways are required for RasV12-induced DR5 expression. When CHOP, Elk1, and c-Jun were co-expressed together, the highest DR5 promoter transactivation and protein expression were detected in comparison to any single or double gene expression. Consistently, oligonucleotide pull-down assay could detect Elk1, CHOP, and c-Jun proteins bound to the DR5 promoter in cells co-transfected with CHOP, Elk1, and c-Jun genes, or transfected with Ras-V12 gene, indicating that these proteins together bind to the DR5 promoter region (Figure 10). Thus, we suggest that these three proteins cooperatively participate in RasV12-induced DR5 expression by enhancing its transcription. Importantly, we found that RasV12-transformed human ovarian cells were more sensitive than their non-transformed counterparts to undergo TRAIL-induced apoptosis (Figure 11). In summary, this study has demonstrated that the oncogenic Ras induces DR5 expression through activation of both ERK/RSK and JNK signaling pathways and subsequent CHOP, Elk1, and c-Jun-mediated gene transactivation. Cancers with Ras mutation may be suitable for TRAIL-based cancer therapy. The manuscript on this part of work is in preparation. The abstract on this part of work has been submitted to the 2010 AACR meeting.

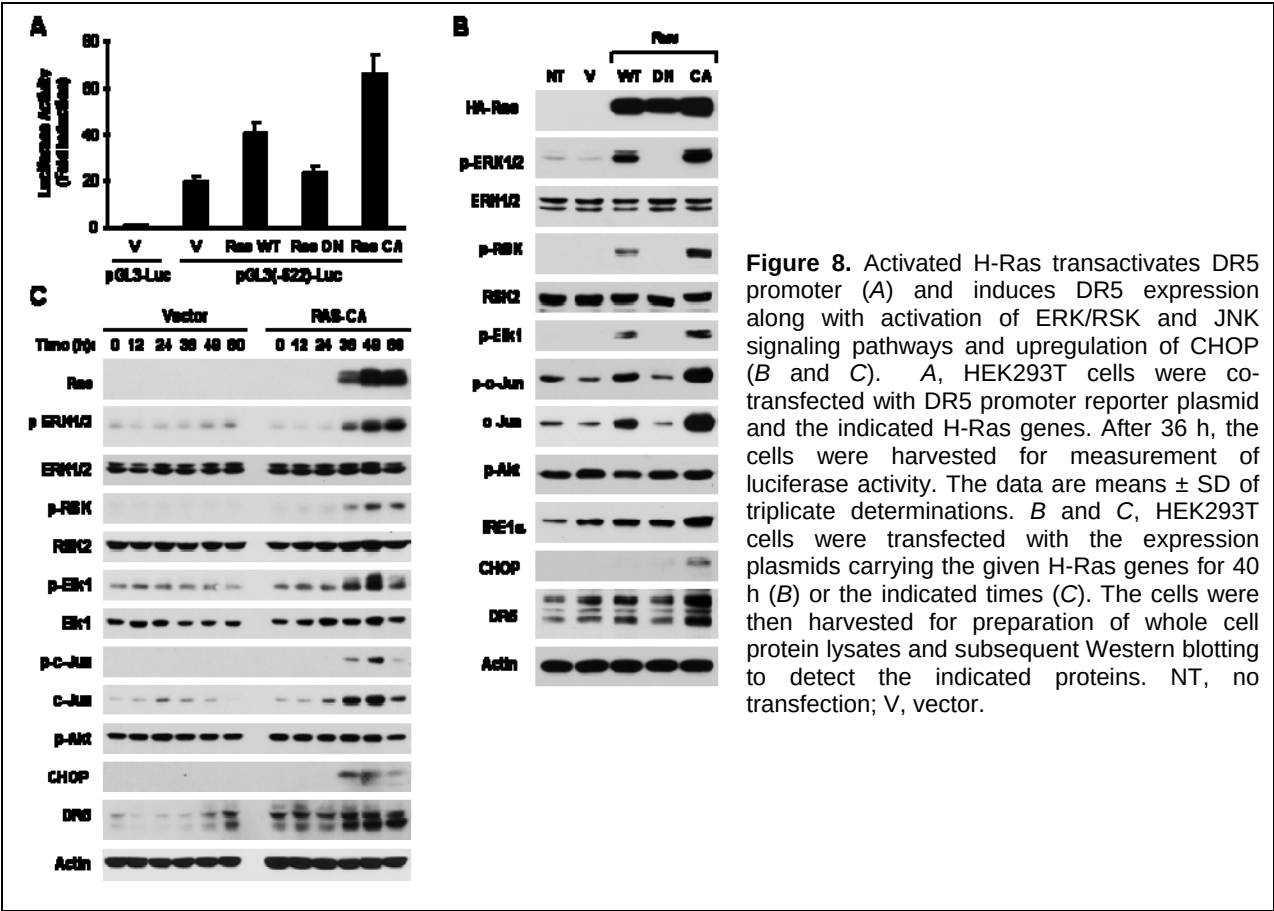


Figure 8. Activated H-Ras transactivates DR5 promoter (A) and induces DR5 expression along with activation of ERK/RSK and JNK signaling pathways and upregulation of CHOP (B and C). A, HEK293T cells were co-transfected with DR5 promoter reporter plasmid and the indicated H-Ras genes. After 36 h, the cells were harvested for measurement of luciferase activity. The data are means $\pm$ SD of triplicate determinations. B and C, HEK293T cells were transfected with the expression plasmids carrying the given H-Ras genes for 40 h (B) or the indicated times (C). The cells were then harvested for preparation of whole cell protein lysates and subsequent Western blotting to detect the indicated proteins. NT, no transfection; V, vector.

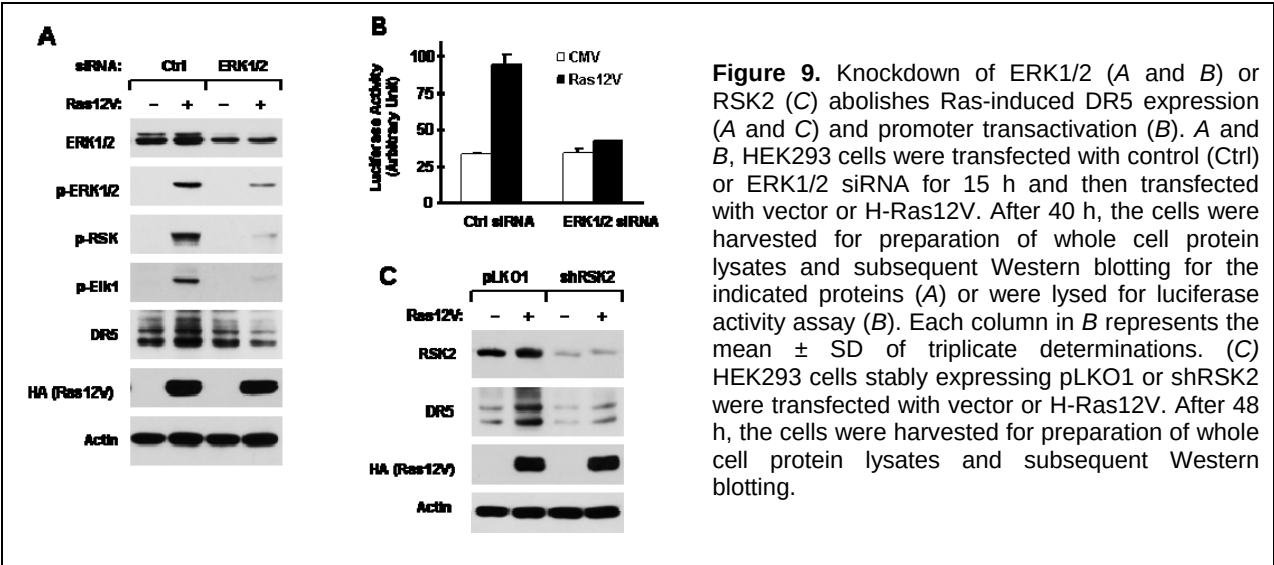


Figure 9. Knockdown of ERK1/2 (A and B) or RSK2 (C) abolishes Ras-induced DR5 expression (A and C) and promoter transactivation (B). A and B, HEK293 cells were transfected with control (Ctrl) or ERK1/2 siRNA for 15 h and then transfected with vector or H-Ras12V. After 40 h, the cells were harvested for preparation of whole cell protein lysates and subsequent Western blotting for the indicated proteins (A) or were lysed for luciferase activity assay (B). Each column in B represents the mean $\pm$ SD of triplicate determinations. (C) HEK293 cells stably expressing pLKO1 or shRSK2 were transfected with vector or H-Ras12V. After 48 h, the cells were harvested for preparation of whole cell protein lysates and subsequent Western blotting.

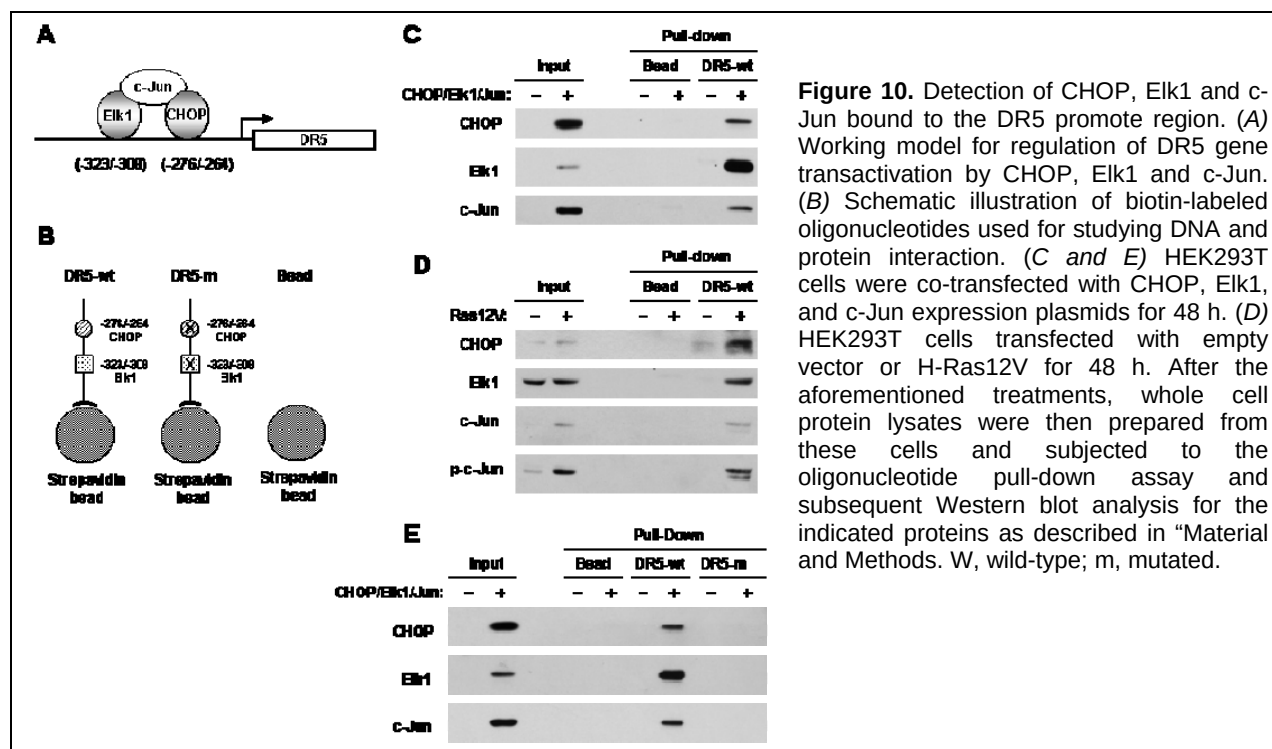


Figure 10. Detection of CHOP, Elk1 and c-Jun bound to the DR5 promote region. (A) Working model for regulation of DR5 gene transactivation by CHOP, Elk1 and c-Jun. (B) Schematic illustration of biotin-labeled oligonucleotides used for studying DNA and protein interaction. (C and E) HEK293T cells were co-transfected with CHOP, Elk1, and c-Jun expression plasmids for 48 h. (D) HEK293T cells transfected with empty vector or H-Ras12V for 48 h. After the aforementioned treatments, whole cell protein lysates were then prepared from these cells and subjected to the oligonucleotide pull-down assay and subsequent Western blot analysis for the indicated proteins as described in "Material and Methods. W, wild-type; m, mutated.

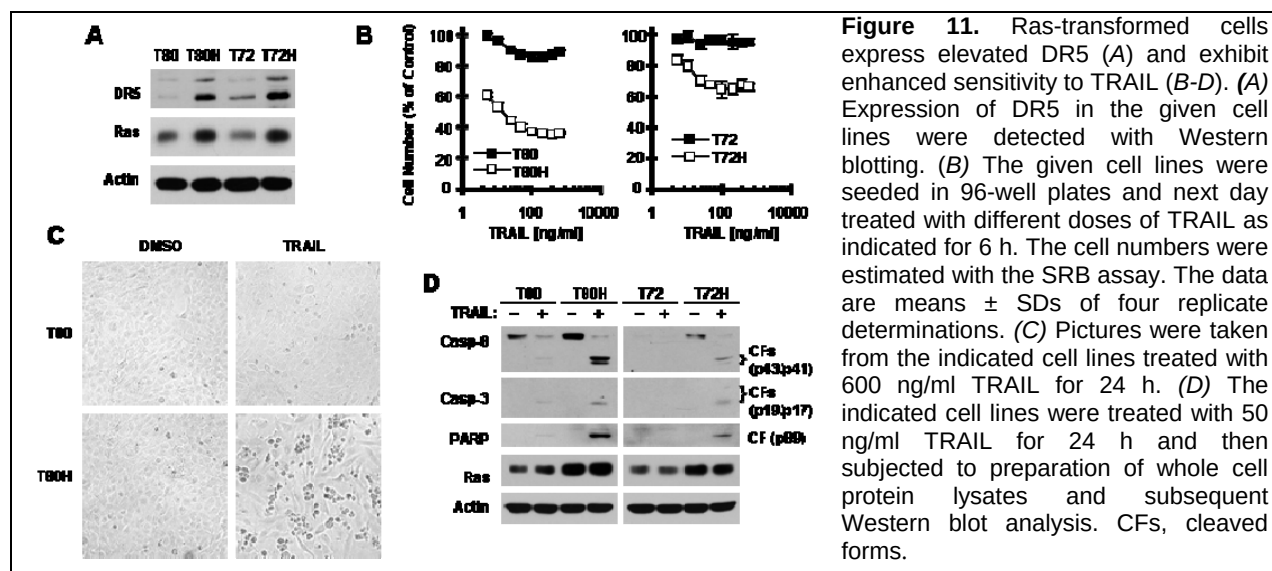


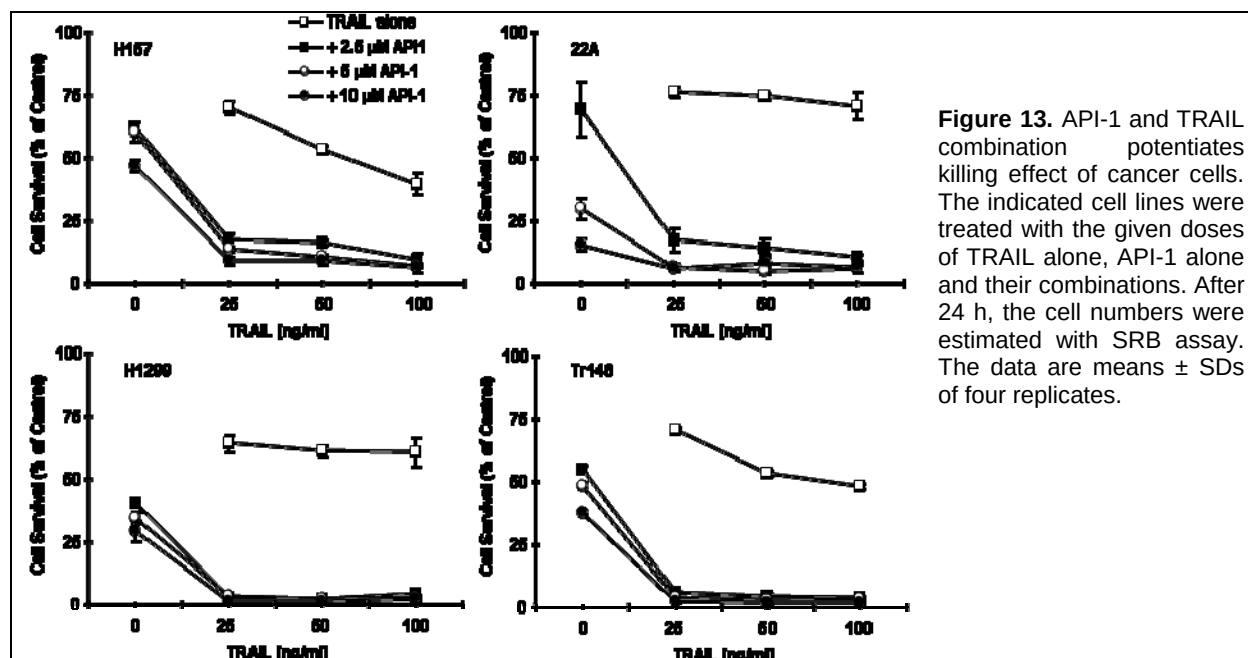
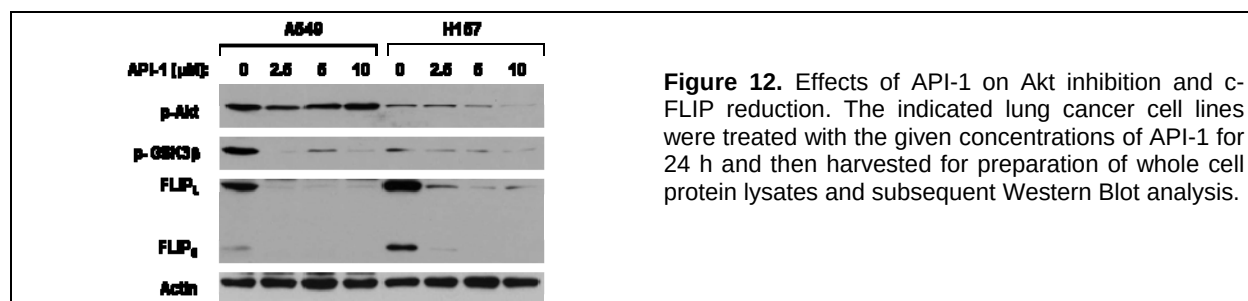
Figure 11. Ras-transformed cells express elevated DR5 (A) and exhibit enhanced sensitivity to TRAIL (B-D). (A) Expression of DR5 in the given cell lines were detected with Western blotting. (B) The given cell lines were seeded in 96-well plates and next day treated with different doses of TRAIL as indicated for 6 h. The cell numbers were estimated with the SRB assay. The data are means ± SDs of four replicate determinations. (C) Pictures were taken from the indicated cell lines treated with 600 ng/ml TRAIL for 24 h. (D) The indicated cell lines were treated with 50 ng/ml TRAIL for 24 h and then subjected to preparation of whole cell protein lysates and subsequent Western blot analysis. CFs, cleaved forms.

Aim 3: To determine whether suppression of PI3K/Akt activity sensitizes premalignant and/or malignant airway epithelial cells to apoptosis induced by DR-induced agents via enhancement of TRAIL/DR-mediated mechanism.

Summary of Research Findings

The novel AKT inhibitor API-1 potently reduces c-FLIP levels and enhances TRAIL-induced apoptosis. We recently studied the effects of the novel Akt inhibitor API-1 on c-FLIP expression and TRAIL-induced apoptosis. API-1 potently reduced the levels of c-FLIP (Figure 12) and enhanced TRAIL's cell killing effect in both lung and head and neck cancer cells (Figure 13). The ongoing works will be conducted outside the scope of this grant: 1) further demonstrate the

cooperative effect of API-1 and TRAIL combination on induction of apoptosis; 2) understand the mechanism(s) by which API-1 induces c-FLIP downregulation; and 3) study the relationship between Akt inhibition and c-FLIP reduction.



The novel PI3 kinase inhibitor BKM120 sensitizes lung cancer cells to TRAIL-induced apoptosis. We recently got the novel PI3K inhibitor BKM120 from Novartis and tested its effect on TRAIL-mediated cell killing effect in human lung cancer cells. We found that BKM120 in combination with TRAIL exhibited enhanced killing effects in the tested cell lines. We are further demonstrating the efficacy of BKM120 and TRAIL combination on induction of apoptosis and tumor growth. Moreover, we will reveal the mechanism by which BKM120 enhances TRAIL-induced apoptosis with ongoing research external to this mechanism.

Aim 4: To determine whether DRs, DcRs, c-FLIP, and procaspase-8 serve as biomarkers for lung cancer chemoprevention and therapy.

Summary of Research Findings

We deferred analysis of these markers in the limited clinical specimens due to the lack of compelling data, when reviewed in collaboration with the project investigators, statisticians and molecular pathologist, to justify such work at this time. However, there are TMA histology sections for further analysis outside the scope of this grant, should more compelling data be obtained in the future.

Key Research Accomplishments

- Demonstrated ERK/RSK-dependent regulation of DR5 expression through co-activation of CHOP and EIK1 transcriptional factors.
- Conducted the first detailed experiment to demonstrate the mechanism by which the oncogenic Ras induces DR5 expression. This finding provides a scientific rationale for targeting tumors with Ras or B-Raf mutation with TRAIL- or DR5 agonistic antibody-based therapies.
- Demonstrated that the Akt inhibitor API-1 downregulates c-FLIP and enhances TRAIL-induced apoptosis.
- Demonstrated that the PI3K inhibitor BKM120 sensitizes lung cancer cells to TRAIL-induced apoptosis.

Conclusion

Our findings on Ras or B-Raf regulation of DR5 expression suggest that the cancer types with Ras or B-Raf mutations may be appropriate populations for TRAIL- or DR5 agonistic antibody-based therapies.

Project 5: Molecular Strategies Targeting the AKT Signaling Pathway for Lung Cancer Chemoprevention and Therapy

(PI and co-PI: Ho-Young Lee, Ph.D., Edward S. Kim, M.D.)

Our goal is to find novel chemopreventive/therapeutic agents that can prevent lung carcinogenesis effectively. Results from our work and others' have demonstrated that Akt, which has a clear role in cellular survival and transformation, is constitutively active in premalignant and malignant HBEs and in NSCLC cell lines. These findings suggest an importance of PI3K/Akt signaling pathway in lung carcinogenesis. The purpose of our studies is to determine whether activation of Akt induces malignant transformation of HBE cells and to develop novel agents inhibiting Akt activity as a strategy to prevent lung carcinogenesis.

Aim 1 Develop a retroviral vector expressing constitutively active Akt and characterize the *in vitro* and *in vivo* effects of Akt activation on the malignant transformation of HBE cells.

Summary of Research Findings

Results from this completed Specific Aim were presented previously.

Aim 2 Evaluate the ability of chemopreventive agents used in VITAL trials (gefitinib, erlotinib, SCH66336, and celecoxib, alone and in combination) to inhibit Akt activity and induce apoptosis in transformed HBE and NSCLC cell lines.

Summary of Research Findings

This aim was completed as previously reported.

Aim 3 Determine whether Akt is activated in bronchial specimens from enrolled patients in VITAL trials and whether treatment with chemopreventive agents suppresses Akt level or activity in these patients.

Summary of Research Findings

This aim was completed as previously reported.

Core B: Biostatistics & Data Management Core

(Core Director: J. Jack Lee, Ph.D.)

Core Goals:

1. To provide statistical design, sample size/power calculations, and integrated, comprehensive analysis for each basic science, pre-clinical, and clinical study.
2. To develop a data management system that provides tracking, quality control, and integration of clinical, pathological, and basic science data. New database modules will be developed and integrated to the existing VITAL web-based database and with the clinical database from the Department of Thoracic/Head and Neck Medical Oncology.
3. To provide statistical and data management support for genomic and imaging studies including microarray, proteomics, protein antibody array, and spiral CT.
4. To develop and adapt innovative statistical methods pertinent to biomarker-integrated translational lung cancer studies.
5. To generate statistical reports for all projects.
6. To collaborate and assist all project investigators in the publication of scientific results.

Summary of Research Findings

A major database revamping effort has been accomplished to integrate the retrospectively collected data with the prospectively conducted clinical trial data. We have developed a consent form to collect further patient data and continue to provide enhancement of a web-enabled database system to facilitate the research activities of the VANGUARD trial. New database modules have been developed and integrated into the existing VITAL database with the clinical database from the Department of Thoracic/Head and Neck Medical Oncology. The key activities are summarized below.

- The ReVITALization database has been developed to extend and integrate with the VITAL database. Additional tissue, clinical, and pathological data were acquired. Tissue repository and tracking were provided. The ReVITALization effort incorporates retrospectively collected tissue sample data. Previously collected tissue data has been evaluated and validated to be included into the VITAL project. Patient consent was sought for eligible patients with available samples to obtain clinical data.
- The SQL Server 2005 database and ASP.NET web application were implemented with VB.net language. Several supporting features were created including query tool and SQL 2005 reports. Secure Socket Layer (SSL) and secured database password were used to keep data transactions protected and confidential. The tissue data includes clinical and pathological data. The main components are listed below.

1) Clinical Module

The database's clinical module provides user friendly input interfaces for entering and viewing patients' clinical data and facilitates the patient search by medical record number (MRN), path number and name. The module contains the following Web forms:

- Patient Information

- Social History (Alcohol and Smoking history)
 - Medical History
 - Other Malignancy
 - Treatments (Surgery, Chemotherapy, Radiotherapy and Other Treatments)
 - Clinical Staging
 - Follow up
- 2) The ReVITALization's pathological module collects primary and metastasis general data, their histology and tumor features which can be used by our tool to automatically determine the cancer's staging. The tissue bank module of frozen (tissue, blood and pleural) and paraffin keep the tissue storage location and the concentration, volume and quality of the DNA, RNA and protein. The module contains the following Web forms:
- Primary and Metastasis data (Diagnosis and Surgery Specimens)
 - Histology
 - Staging and Tumor Information: Cancer staging (TNM classification) is automatically determined by the system based on the tumor information provided.
 - Tissue Bank (Frozen Tissue and Paraffin)
- 3) Query Tool: This module offers the capability to provide an extensive search of the stored data efficiently by internal query template and various reports. The Query Tools allow:
- Query for various types of SQL commands such as in, not in, like and =.
 - Specify multiple criterions for the same fields.
 - Form a complex query by using 'and', 'or' and parenthesis.
 - Simplify the Sample Type Selection section and Query Criteria Section.
 - The query can be saved for the future usage. The saved query can also be deleted, overwritten and renamed.
 - The query can be sorted by any field specified.
 - The selected columns can be ordered and re-ordered by our user-friendly tool.
 - The query results can be exported to the excel format.
- 4) Reports: Several Excel reports are provided for clinical and pathological module.
1. Clinical Report
 2. Pathological Report
 3. Patient Report
 4. Accession Report
 5. General Information Report
 6. Other Malignancy Report
 7. Surgery Report
 8. Chemotherapy Report
 9. Radiotherapy Report
 10. Other Treatment Report
 11. Staging Report
 12. Follow up Report
 13. Histology Diagnosis Report
 14. Patient Summary Report

- 5) Dictionaries: The database gives control for the users to conveniently add and update dictionaries; however to prevent data loss, the dictionary deletion is prohibited for users. The deletion of dictionaries is only allowed for the database administrators.

The ReVITALization database application has been constantly updated to assist users in the data entering process. Selected screen shots are provided in the Appendix 1. The retrospective data is checked and imported into the database system, while some new data was entered by the users. The database is maintained and constantly backed up by the database administrator.

- **Additional Data Management Activities**

- 1) The data has been examined thoroughly for data quality control. Data cleaning process involving clinicians, nurses, programmers, and the statisticians is finalized.
- 2) The data were consistently downloaded for the analysis.
- 3) The database has been monitored, maintained, and backed up routinely by the database administrators.
- 4) The retrospective data is checked and imported into the database system. New data have been entered by the users.
- 5) Routine database maintenance is performed by the database administrator.

We have generated the statistical report for ReVITALization data incorporating the important patients' clinicopathological characteristics, treatment, and 21 biomarkers located in the histological analysis.

Key Research Accomplishments

- Completed the data cleaning and data uploading into the database.
- Analyzed data for Project investigators to predict patient outcome.
- Started final statistical analysis on the ReVITALization data to build a risk model.

Conclusion

We will continue our efforts in finalizing the statistical models to predict overall survival and recurrence-free survival, and validating the models in the VITAL data set. A six-month no-cost extension for this component of the grant has been requested to allow protected time for our biostatisticians to complete these critical statistical analyses.

Core C: Pathology and Specimen Procurement Core

(Core Director: Ignacio Wistuba, M.D.)

Aim 1. Develop and maintain a repository of tissue and other biologic specimens from patients enrolled on the clinical trials in Project 1.

Summary of Research Findings

In the past year, we continued enrolling patients in the VITAL/Vanguard trials (Project 1). As of December 2010, 128 bronchoscopies have been performed at baseline, 12-, 24- and 36-month periods. From these patients, Core C has acquired, processed, and banked a total of 1,840 specimens obtained during these bronchoscopies (Table 2).

Table 2. Summary of specimens collected and banked in the Pathology Core

Type of Specimen	Number
Sputum	122
Buccal Brush	122
Bronchial Brush	737
Bronchial Wash	123
Tissue Specimens	736
Total	1,840

In collaboration with Project 2, mRNA has been extracted from 320 bronchial epithelial samples for mRNA Affymetrix profiling, including RNA quality control and RNA amplification.

Aim 2. Maintain a comprehensive database of tissue and specimen characteristics from patients enrolled in the clinical trials of Project 1, including pathologic characteristics of each specimen, inventory and distribution.

Summary of Research Findings

The Biostatistics Core (Dr. J.J. Lee) has developed a Web-based database that has been used by the Pathology Core members to catalogue all the specimens obtained and banked in the Core C and to report pathology diagnosis (see previous Core C report). From the Vanguard patients, 1,104 cytological specimens and 736 bronchial biopsies have been tracked and inventoried using the Web-based database.

Aim 3. Provide comprehensive pathologic characterization of all tissues and other biologic specimens and assist in preparation and evaluation of studies involving these tissues.

Summary of Research Findings

Core C has processed and performed histopathological diagnosis of 1,608 histology sections from patient bronchoscopies (Table 3). We examined two H&E-stained tissue sections per bronchial biopsy.

Table 3. Summary of the histopathology diagnoses made in 1,608 histology sections from bronchial biopsies obtained from Vanguard clinical trial (Project 1).

Diagnosis	N	%
No Tissue/Denuded Epithelium	45	2.8%
Normal Epithelium	1,102	68.5%
Goblet Cell Metaplasia (GCM)	94	5.8%
Basal Cell Hyperplasia (BCH)	292	18.2%
Combined GCM/BCH	386	24.0%
Squamous Metaplasia	57	3.5%
Angiogenic Squamous Dysplasia (ASD)	2	0.1%
Mild Dysplasia	8	0.5%
Moderate Dysplasia	8	0.5%

Aim 4. Provide centralized immunohistochemistry and laser capture microdissection services, nucleic acid extractions and assistance with construction and evaluation of tissue arrays.

Summary of Research Findings

This Aim was completed and reported in the previous year.

Aim 5. Identify ~600 surgically resected tissue specimens from stages I/II NSCLC (tumor, normal and abnormal adjacent bronchial epithelium specimens) and their complete clinical and pathologic information.

Summary of Research Findings

This Aim was completed and reported in the previous year.

Aim 6. Examine over 40 biomarkers in those specimens by immunohistochemical (IHC) and tissue microarrays (TMAs).

Summary of Research Findings

Our initial aim was to study over 40 markers by IHC using TMA methodology. However, as part of the ReVITALization project, we proposed to study 25 IHC markers in total requested by the research projects: Project 2 (J. Minna, L. Mao, and R. Lotan), six markers; Project 3 (P. Koo and W. Hittelman), five markers; Project 4 (F. Khuri), four markers; Project 5 (H-Y Lee), ten markers. Based in the data available from the research projects, 21 markers were examined and reported last year. The data generated by Core C were reported to the Biostatistics Core (Core B) for analysis (see related Project and Biostatistics Core reports for additional detail). We did not expand the number of IHC markers to be studied during the last year of the no-cost extension for this grant due to the lack of compelling data, when analyzed in collaboration with the research project investigators, to support such work. However, there are TMA histology sections for future IHC analysis, including those that will be generated by the analysis of the data obtained during the mRNA profiling of the Vanguard bronchoscopy biopsies.

Key Research Accomplishments

- Processed and diagnosed 736 bronchoscopy tissue specimens from 128 bronchoscopies from patients enrolled in the Vanguard trial (Project 1).
- In collaboration with Project 2, mRNA has been extracted from 320 bronchial epithelial samples for mRNA Affymetrix profiling, including RNA quality control using RNA bioanalyzer.

Conclusion

During the last year, we have acquired and banked 736 tissue specimens and 1,104 cytology specimens, mostly bronchial brushes, from 128 bronchoscopies and resected specimens from lung cancer and head/neck tumor patients, and used the database developed by the Biostatistics Core to track and inventory the bronchoscopy specimens and report the histopathological features.

KEY RESEARCH ACCOMPLISHMENTS (IN SUMMARY)

Project 1: Biologic Approaches for Adjuvant Treatment of Aerodigestive Tract Cancer

- Increased number of evaluable patients to 38, with 27 patients remaining on study.
- Continued to collect patient clinical data and tissues for distribution to support research projects in the VITAL grant.
- Completed tissue analysis of patients who have completed their 12-month bronchoscopy.
- Performed statistical analyses on data from 542 patients (including 10 from the Vanguard trial) whose tumor specimens were analyzed for 21 markers.
- Completed tissue identification and clinical data collection of over 500 archived tissue specimens from the pathology database for the ReVITALization plan.

Project 2: Identification of Biomarkers of Response to Chemoprevention Agents in Lung Epithelium

- Validated the FILM five-gene signature, initially analyzed using public datasets, at the protein level by immunohistochemistry using an independent set of FFPE NSCLC tissue.
- Showed that the FILM protein signature, like its transcript counterpart, is effective in predicting survival in lung adenocarcinoma and demonstrated prognostic specificity towards this subtype of NSCLC.
- Demonstrated that the FILM protein signature is powerful in predicting the survival of early stage (stage-I or stage-IB) lung adenocarcinoma patients who did not receive any form of therapy.
- Analyzed genes differentially expressed among airways sampled at different anatomical locations of the lung and at various distances from the original resected tumors and identified site-dependent differentially expressed genes and pathways in the field of cancerization that have potential value in predicting recurrence or second primary tumor development.
- Processed the remaining of eligible Vanguard patient airway samples (n=320) for molecular profiling analysis to be completed in the spring of 2011.
- Awarded a DoD Lung Cancer Research Program Consortium award (PI Ignacio Wistuba, Director, Thoracic Molecular Pathology Laboratory) based on the preliminary findings of the Vanguard molecular profiling analysis to expand the Vanguard study findings and analyze the field of cancerization in more unprecedented detail.

Project 3: Premalignant Bronchial Epithelia: Molecular and Cellular Characterization of

- Established an organotypic culture method to mimic bronchial hyperplasia/dysplasia *in vitro*. We also tested a potential preventive method that blocks hyperplasia/dysplasia using the EGFR inhibitor erlotinib.
- Demonstrated that chemokine receptor CXCR2 expression in tumor cells is associated with a poor prognosis in NSCLC.
- Demonstrated that CREB and NF- κ B-regulated CXC chemokine gene expression in lung carcinogenesis.
- Showed that CREB overexpression is associated with negative prognosis in never-smokers with NSCLC.
- Demonstrated growth suppression of lung cancer cells by targeting CREB.

Project 4: Modulation of Death Receptor-Mediated Apoptosis for Chemoprevention

- Demonstrated ERK/RSK-dependent regulation of DR5 expression through co-activation of CHOP and Elk1 transcriptional factors.

- Conducted the first detailed experiment to demonstrate the mechanism by which the oncogenic Ras induces DR5 expression. This finding provides a scientific rationale for targeting tumors with Ras or B-Raf mutation with TRAIL- or DR5 agonistic antibody-based therapies.
- Demonstrated that the Akt inhibitor API-1 downregulates c-FLIP and enhances TRAIL-induced apoptosis.
- Demonstrated that the PI3K inhibitor BKM120 sensitizes lung cancer cells to TRAIL-induced apoptosis.

Core B: Biostatistics & Data Management Core

- Completed the data cleaning and data uploading into the database.
- Analyzed data for Project investigators to predict patient outcome.
- Started final statistical analysis on the ReVITALization data to build a risk model.

Core C: Pathology and Specimen Procurement Core

- Processed and diagnosed 736 bronchoscopy tissue specimens from 128 bronchoscopies from patients enrolled in the Vanguard trial (Project 1).
- In collaboration with Project 2, mRNA has been extracted from 320 bronchial epithelial samples for mRNA Affymetrix profiling, including RNA quality control using RNA bioanalyzer.

REPORTABLE OUTCOMES

Publications (attached in Appendix 3):

List of Publications (published and in progress)

Chen S, Fu L, Raja SM, Yue P, Khuri FR, Sun SY. Dissecting the roles of DR4, DR5 and c-FLIP in the regulation of geranylgeranyltransferase I inhibition-mediated augmentation of TRAIL-induced apoptosis. *Molecular Cancer*. 2010 Jan 29;9:23. PMCID: PMC2824632.

Fan S, Li Y, Yue P, Khuri FR, Sun SY. The eIF4E/eIF4G interaction inhibitor 4EGI-1 augments TRAIL-induced apoptosis through DR5 induction and c-FLIP downregulation independent of inhibition of cap-dependent protein translation. *Neoplasia*. 2010 Apr;12(4):346-56. PMCID: PMC2847742.

Jeong Y, Xie Y, Xiao G, Behrens C, Girard L, Wistuba II, Minna JD, Mangelsdorf DJ. Nuclear receptor expression defines a set of prognostic biomarkers for lung cancer. *PLoS Medicine*. 2010 Dec 14;7(12):e1000378. PMCID: PMC3001894.

Kadara H, Behrens C, Yuan P, Solis LM, Liu D, Gu X, Minna JD, Lee JJ, Kim ES, Hong WK, Wistuba II, Lotan R. A five-gene and corresponding-protein signature for stage-I lung adenocarcinoma prognosis. *Clinical Cancer Research*. Dec 16, 2010, epub ahead of print. PMID: 21163870.

Kim ES, Hong WK, Lee JJ, Mao L, Morice RC, Liu DD, Jimenez CA, Eapen GA, Lotan R, Tang X, Newman RA, Wistuba II, Kurie JM. Biological activity of celecoxib in the bronchial epithelium of current and former smokers. *Cancer Prevention Research*. 2010 Feb;3(2):148-59. PMID: 20103722.

Lee J, Ryu SH, Kang SM, Chung WC, Gold KA, Kim ES, Hittelman WN, Hong WK, Koo JS. Prevention of Bronchial Hyperplasia by EGFR Pathway Inhibitors in an Organotypic Culture Model. (Accepted, *Cancer Prevention Research*).

Li C, Chen S, Yue P, Lonial S, Khuri FR, Sun SY. The proteasome inhibitor PS-341 (Bortezomib) induces calpain-dependent I κ B α degradation. *Journal of Biological Chemistry*. 2010 May 21;285(21):16096-104. PMCID: PMC2871478.

Oh YT, Liu X, Yue P, Kang S, Chen J, Taunton J, Khuri FR, Sun SY. ERK/RSK signaling positively regulates death receptor 5 expression through co-activation of CHOP and Elk1. *Journal of Biological Chemistry*. 2010 Dec 31;285(53):41310-9. PMCID: PMC3009856.

Abstracts (attached in Appendix 2)

Behrens C, Lin H, Nunez M, Yuan P, Solis LM, Raso MG, Prudkin L, Sun M, Li X, Tang X, Roth JA, Minna JD, Stewart D, Hong WK, Lee JJ, Wistuba II. Differences in protein expression patterns in lung adenocarcinomas arising in never versus ever smokers. Proceedings of the 101st Annual Meeting of the American Association for Cancer Research; 2010 Apr 17-21; Washington, DC. Philadelphia (PA): AACR; 2010. Abstract 787.

Behrens C, Yuan P, Solis L, Saintigny P, Kadara H, Fujimoto J, Moran C, Swisher SG, Heymach JV, Wistuba II. EZH2 expression is an early event in the pathogenesis of non-small

cell lung cancer (NSCLC) and correlates with tumor progression. Submitted to the 2011 AACR Meeting.

Fujimoto J, Kadara H, Behrens C, Liu D, Lee JJ, Solis LM, Kim ES, Shalafkhane A, Wistuba II, Lotan R. Implication of GPRC5A loss in lung carcinogenesis in patients with and without chronic obstructive pulmonary disease. Submitted to the 2011 AACR Meeting.

Kadara H, Hong WK, Wistuba II, Lotan R. A five-gene signature predictive of survival in lung adenocarcinoma but not in squamous cell carcinoma. Proceedings of the 101st Annual Meeting of the American Association for Cancer Research; 2010 Apr 17-21; Washington, DC. Philadelphia (PA): AACR; 2010. Abstract 4817.

Kadara H, Saintigny P, Fan Y, Chow CW, Chu ZM, Lang W, Behrens C, Gold K, Liu D, Lee JJ, Mao L, Kim ES, Hong WK, Wistuba II. Gene expression analysis of field of cancerization in early stage NSCLC patients towards development of biomarkers for personalized prevention. Submitted to the 2011 AACR Meeting.

Kim WY, Jin Q, Prudkin L, Kim JS, Morgillo F, Feng L, Kim ES, Hennessy B, Lee JS, Mills M, Lee J, Glisson B, Lippman SM, Wistuba II, Lee HY. *EGFR* and *K-Ras* mutations and resistance of lung cancer to the IGF-1R tyrosine kinase inhibitors. Proceedings of the 101st Annual Meeting of the American Association for Cancer Research; 2010 Apr 17-21; Washington, DC. Philadelphia (PA): AACR; 2010. Abstract 4127.

Saintigny P, Liu J, Lee JJ, Ping Y, Behrens C, Solis Soto LM, Heymach JV, Kim ES, Hong WK, Kurie JM, Wistuba II, Koo JS. CXCR2 expression in tumor cells is associated with an adverse outcome in a large set of non-small-cell lung cancer (NSCLC). Submitted to the 2011 AACR Meeting.

Yuan P, Kadara H, Behrens C, Tang X, Woods D, Solis LM, Huang J, Spinola M, Dong W, Yin G, Fujimoto J, Kim E, Xie Y, Girard L, Moran C, Hong WK, Minna JD, Wistuba II. Sex determining region Y-box 2 is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung. Proceedings of the 101st Annual Meeting of the American Association for Cancer Research; 2010 Apr 17-21; Washington, DC. Philadelphia (PA): AACR; 2010. Abstract 5166.

Grants/programs awarded

DoD Lung Cancer Research Program collaborative Award (W81XWH-10-1-1007, PI: Ignacio I. Wistuba).

NIH/NCI "Role of CREB in Lung Cancer Development" (1R01CA126801 PI; Ja Seok Koo).

CONCLUSIONS

Project 1: Completion of the development of the pathological and clinical database for the ReVITALization Proposal was a key milestone for this program. The analysis of the tissue for important biomarkers has been completed, and results were correlated with the clinical outcomes. With expertise from the Pathology and Biostatistics Cores, we are currently in the development of a risk model for development of SPT and recurrence. This model has served as an example for other ongoing projects to effectively integrate pathology and clinical databases. Continued support for the completion of the statistical risk modeling, the final component of proposed work for this grant, has been requested in the form of a six-month no-cost extension.

Project 2: Our findings highlight an immunohistochemical protein signature (FILM protein signature), corresponding to a five-gene transcript signature we previously derived, that is effective in predicting the survival of non-treated stage-I and in particular stage-IB lung adenocarcinoma patients. Thus, this five-protein FFPE signature shows promise as a clinically relevant classifier in distinguishing early-stage patients who will benefit from adjuvant therapy from those who can be spared therapy.

In addition, our findings demonstrate that gene expression in the lung airway field of cancerization is modulated in a site- (adjacent versus distant from original primary tumor) and time- (following resection of tumor) dependent manner. Moreover, differential gene expression in the airway field of cancerization encompasses canonical cancer-associated cellular pathways that may be 1) biologically crucial for early pathogenesis of lung cancer and 2) clinically relevant for prediction of recurrence and second primary tumor and therefore development of personalized targeted strategies for chemoprevention of the disease.

Project 3: Our studies produced several important findings that implicate a potential for future use in clinical settings. As a chemoprevention strategy, the combination of erlotinib with a MEK inhibitor may be more efficient in suppression of early abnormal changes in bronchial epithelium. In particular, we identified the chemopreventive effect of EGFR and MEK inhibitors, identified CXCR2 as a new potential therapeutic/preventive target, identified a small molecule inhibitor of a new class of anti-angiogenic method, and demonstrated that CREB could act as a novel therapeutic and preventive strategy. CXCR2 overexpression in NSCLC tumor cells is frequent and associated with an adverse outcome. The CXCR2/CXCR2 ligands biological axis may be associated with activation of the KRAS and NF κ B signaling pathways and poor prognosis in lung adenocarcinoma. Given the fact that there is no efficient method to treat NSCLC with KRAS mutation, targeting CXCR2 in combination with conventional chemotherapy or targeted therapeutics could lead to potential modes of treatment in the future. Targeting CREB using a small molecule inhibitor successfully blocked angiogenic effect of CXCLs, suggesting that targeting CREB may be a new anti-angiogenic strategy for lung cancer treatment and prevention. Further studies outside the scope of this grant are planned to translate these findings into clinical applications.

Project 4: Our findings on Ras or B-Raf regulation of DR5 expression suggest that the cancer types with Ras or B-Raf mutations may be appropriate populations for TRAIL- or DR5 agonistic antibody-based therapies.

Core B: We will continue our efforts in finalizing the statistical models to predict overall survival and recurrence-free survival, and validating the models in the VITAL data set. A six-month no-cost extension for this component of the grant has been requested to allow protected time for our biostatisticians to complete these critical statistical analyses.

Core C: During the last year, we have acquired and banked 736 tissue specimens and 1,104 cytology specimens, mostly bronchial brushes, from 128 bronchoscopies and resected specimens from lung cancer and head/neck tumor patients, and used the database developed by the Biostatistics Core to track and inventory the bronchoscopy specimens and report the histopathological features.

APPENDICES

APPENDIX 1

APPENDIX 2

APPENDIX 3

Appendix 1: Screen Shots

1) ReVITALization's clinical module: Patient Information, Social History, Medical History

Admin Projects Histo-Pathology Lab Logout

MRN Name Benjamin G Sample

Path# ☒ Patient Completed

662526

Status: Ready...

Other Malignancy Treatment Staging Follow up Pathology All Clinical TO EXCEL

Patient Information EXPORT TO EXCEL

Last Sample Middle G First Benjamin MDAH 662526

Gender Male Race White DOB 12/31/1936

City Montgomery State TX Enter Date 06/22/2006

Zip Code 77356 Country USA Age

Last Visit Date Death or Alive Alive

Lost To Follow up n/a Chemotherapy n/a

Radiotherapy n/a Recurrence n/a

Social History EXPORT TO EXCEL

Smoking History ☐ Yes ☐ No Pack Years 0

Are you currently smoking? ☐ Yes ☒ No

If no, date quit smoking

Age started smoking regularly 18

Average number of cigarettes smoked per day? 20

If there was a quit-smoking period, total time during the smoking years. 0 yrs., 0 mos.

Age Quit Smoking

Overall Smoking Years

Actual Smoking Years

Pack Years

Asbestos Exposure ☐ Yes ☐ No

Alcohol History ☒ Yes ☐ No No. of Drinks / Month 0

Medical History EXPORT TO EXCEL

Hypertension ☐ Yes ☒ No Heart Problem ☐ Yes ☒ No Thyroid ☐ Yes ☐ No

Diabetes ☐ Yes ☒ No Renal Insufficiency ☐ Yes ☐ No Asthma ☐ Yes ☒ No

DVT ☐ Yes ☒ No Pulmonary Embolism ☐ Yes ☒ No Mild hemoptysis ☐ Yes ☐ No

Radiation Fibrosis ☐ Yes ☐ No Hepatic Problem ☐ Yes ☐ No COPD ☐ Yes ☐ No

Other Hypertension, Hyperlipidemia, Depression, Gastric Reflux Date COPD Dx

2) ReVITALization's clinical module: Other Malignancy

OtherMalignancy: 499975				Status: Ready...	
Click here to add 1 more row(s).				Save it	Cancel it
				Open it	Save and Close
EXPORT TO EXCEL					
Malig. ID	Patient ID	Dx Date	Malig. Detail		Treatment
 45	113	1/1/1900	Period: Synchronous to lung cancer (within a year) Organ: Lung Histology: Adenocarcinoma Comments: NED: ☐ NED Date: 1/1/1900		Surgery: ☐ Date: 3/12/2003 Chemo: ☐ Date: 1/1/1900 Radio: ☐ Date: 1/1/1900
 46	113	1/1/1900	Period: n/a Organ: Skin Histology: Melanoma Comments: NED: ☐ NED Date: 1/1/1900		Surgery: ☐ Date: 4/20/2004 Chemo: ☐ Date: 1/1/1900 Radio: ☐ Date: 1/1/1900
 49	113	1/1/1900	Period: After lung cancer Organ: Skin Histology: Squamous Cell Carcinoma Comments: NED: ☐ NED Date: 1/1/1900		Surgery: ☒ Date: 2/9/2005 Chemo: ☐ Date: 1/1/1900 Radio: ☐ Date: 1/1/1900

3) ReVITALization's clinical module: Treatment: Surgery, Chemotherapy, Radiotherapy and Other Treatments.

Status: Ready...

Surgery: 523912

[Click here](#) to add 1 more row(s). [Save It](#) [Cancel It](#) [Open It](#) [Save and Close](#) [EXPORT TO EXCEL](#)

Surgery ID	Patient ID	Surgery Date	Is MDA	Surgery Procedure	Comments	Margin Left
29	142	3/24/2003	☐	n/a		☐

Chemotherapy

[Click here](#) to add 1 more row(s). [Save It](#) [Cancel It](#) [Open It](#) [Save and Close](#) [EXPORT TO EXCEL](#)

Chemo ID	Chemo Type	Chemo Date	Drug	Tumor Size	Response	Comments
29	☐ Is MDA:	Start: 1/1/1900	A: Carboplatin	Before(CT)	0	
142	Chemo. Type: Neoadjuvant	Stop: 1/1/1900	B: Paclitaxel	After(CT)	0	
			C: n/a	% Reduction	NaN	
			#crs: 0	Before(Patho)	0	
				After(Patho)	0	
				% Reduction	NaN	
143	☐ Is MDA:	Start: 7/29/2006	A: Cisplatin	Before(CT)	0	
142	Chemo. Type: 1st Line	Stop: 9/16/2003	B: Etoposide	After(CT)	0	
			C: n/a	% Reduction	NaN	
				Before(Patho)	0	
				After(Patho)	0	
				% Reduction	NaN	

Radiotherapy

[Click here](#) to add 1 more row(s). [Save It](#) [Cancel It](#) [Open It](#) [Save and Close](#) [EXPORT TO EXCEL](#)

Radio ID	Treatment Option	Radio Date	Tumor Size	Response	Comments
	☐ Is MDA:		Before(CT)	0	
	Site1: L-lung	Start: 7/29/2003	After(CT)	0	
37	Site2: n/a	Stop: 9/16/2003	% Reduction	NaN	
142	Site3: n/a		Before(Patho)	0	
	Treatment Option: 1st Line		After(Patho)	0	
			% Reduction	NaN	

Other Treatment

[Click here](#) to add 1 more row(s). [Save It](#) [Cancel It](#) [Open It](#) [Save and Close](#) [EXPORT TO EXCEL](#)

Treatment ID	Patient ID	Surgery Date	Other Treatment	Comments
8	142	5/5/2004	Craniotomy	

4) ReVITALization's clinical module: Staging

Staging:				Status: Ready...							
Click here to add 1 more row(s). Save it Cancel it Open it Save and Close Excel Report											
Stage ID	Patient ID	Staging Date	Current Situation	Clin. T	Clin. N	Clin. M	Clin. Stage	Pleu Eff	Malign PI Eff		
 1	44	1/1/1900	Multiple primary tumor different histology	T2	N0	M1	IB	☒	☒		

Pathology											
Tumor Specimens											
Patient ID	Accession	Surgical Date	Single Wedge	Multiple Wedge	Single Segmentectomy	Multiple Segmentectomy	Lobectomy	Bilobectomy	Pneumonectomy	# Nodules	Tumor ID
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	23
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	26
44	s-04-23495	05/11/2005	False	False	False	False	True	False	False	2	27

Dx Specimens							
Patient ID	AccessionNo	Path Type	Event	Dx Specimen Date	Specimen Type	Tumor Site	Specimen Avail
44	SB-1111	Primary	Dx Specimen	01/19/2007	CORE BIOPSY	LLL	

Metastasis Specimens							
Patient ID	AccessionNo	Path Type	Event	Met Date	Specimen Type	Tumor Site	Specimen Avail

5) ReVITALization's clinical module: Follow up

Follow up: 127855				Status: Ready...	
Click here to add 1 more row(s). Save it Cancel it Open it Save and Close EXPORT TO EXCEL					
Person Review n/a				Date Updated	
Fu ID	Patient ID	Fu Date	Fu Detail		
	422	55	2/4/2004	Status	recurrence
				If "no change or no recurrence":	Form of contact:
				If "recurrence":	Form of contact: visit to clinic Date of recurr: 1/1/1900
					Thoracic Lymph
				Site of Recur:	n/a If Lung: n/a
					n/a
				Recur Biopsy:	☐ Image Date: 1/1/1900
				If "death":	Death Date: Cause of Death:
					Info by:
				Comments:	
	298	55	1/4/2006	Status	no change
				If "no change or no recurrence":	Form of contact: visit to clinic
				If "recurrence":	Form of contact: Date of recurr:
				Site of Recur:	If Lung:
				Recur Biopsy:	☐ Image Date:
				If "death":	Death Date: Cause of Death:
					Info by:
				Comments:	

6) ReVITALization's pathological module: Tissue Pathological Data

- Primary Diagnosis Specimen
- Primary Surgical Specimen
- Metastasis Diagnosis Specimen
- Metastasis Surgical specimen

Admin
Projects
Histo-Pathology Lab
Logout

Select a participant
Status: Ready...

Hist Dx TO EXCEL
All Pathology TO EXCEL

[Click here](#) to add more row(s).
[Save it](#)
[Cancel it](#)
[Open it](#)
[Save and Close](#)
EXPORT TO EXCEL

				Accession	Biopsy	Event
				563 SB-1111	Primary	Surgical Specimen
				564 SB-2222	Primary	Dx Specimen
				566 SS-1112	Metastasis	Surgical Specimen
				567 S-04-23495	Primary	Surgical Specimen
				568 SB-01-1234	Metastasis	Dx Specimen

Dx Specimen

[Click here](#) to add more row(s).
[Save it](#)
[Cancel it](#)
[Open it](#)
[Save and Close](#)
EXPORT TO EXCEL

			Dx Specimen ID	Obtained Date	Accession ID	Accession No	Specimen Type	Tumor Site	Specimen Avail

Surgical Specimen

[Click here](#) to add more row(s).
[Save it](#)
[Cancel it](#)
[Open it](#)
[Save and Close](#)
EXPORT TO EXCEL

			Surg ID	Acc. ID	Accession No	Surgical Date	Single Wedge	Multi Wedge	Single Segmen tectomy	Multi Segmen tectomy	Lobec tomy	Bilo bectomy	Pneumo nectomy	No of Nodules	T	N	M
			6	567	S-04-23495	5/11/2005	☐	☐	☐	☐	☒	☐	☐	2	n/a	n/a	n/a

Met Specimen

[Click here](#) to add more row(s).
[Save it](#)
[Cancel it](#)
[Open it](#)
[Save and Close](#)
EXPORT TO EXCEL

			Met Specimen ID	Met Date	Accession ID	Accession No	Specimen Type	Tumor Site	Specimen Avail
			DX 6	1/1/1900	566	SS-1112	n/a	n/a	☐

7) ReVITALization's pathological module: Histology

Histology		Status: Ready
UpdateDelete		
Hist ID:	5	Dx Specimen ID: 22Met Specimen ID: 0Tumor ID: 0
Histology Dx:	Metastasis to lung	
Metastasis (to the lung) Dx:	Carcinoma	Other Tumoral Characteristics (Hist Dx)
Carcinoma Met Site:	Breast	Necrosis % 23
		Fibrosis % 3
		Inflammation Severe

8) ReVITALization's pathological module: Staging and Tumor Information

Tumor: s-04-23495 (6)				Status: Ready...			
Click here to add 1 more row(s). Save it Cancel it Open it Save and Close Excel Report							
			Tumor ID	Surgical ID	Specimen Type	Tumor Site	Localization
		DX	23	6	n/a	n/a	n/a
		DX	26	6	Bilobectomy	n/a	n/a
		DX	27	6	Lobectomy	n/a	n/a

Type of Tissue Available	Tumor Size(cm) 0	Pathological T
☐ Bronchus Extrapulmonary ☐ Principal ☐ Lobar Intrapulmonary ☐ with Cartilage ☐ w/o Cartilage ☐ Bronchiole ☐ Alveoli	☐ Tumor Invasion ☐ Pleural ☐ Neural ☐ Vascular ☐ Other ☐ Margin Positive ☐ Bronchial ☐ Parenchymal ☐ Soft Tissue ☐ Other	Pathological T: n/a ☐ Tx T1: <= 3 cm. T2: > 3 cm or <= 3cm and/or attached to visceral pleura ☐ Pleural Attached T3: ☐ Parietal Pleura ☐ Mediastinal Pleura ☐ Chest Wall ☐ Mediastinal Fat ☐ Pericardium ☐ Phrenic Nerve ☐ Vagus Nerve ☐ Sympathetic Chain ☐ Atelectasis Entire Lung T4: ☐ >1 Nodal ☐ Invades to Great Vessel ☐ Invades to Heart ☐ Invades to Trachea ☐ Invades to Carina ☐ Invades to Esophagus ☐ Invades to Vertebral Bones ☐ Associated with Malignant Pleural Effusion ☐ Satellite Tumor Nodule (size)

Field Study	Lymph Node Metastasis ☐ Lymph Node Involvement																																																												
Normal bronchial epithelium and premalignant lesions ☐ Normal Bronchial Epithelium ☐ Normal Lung Parenchyma ☐ Hyperplastic Alveoli ☐ Bronchial Hyperplasia ☐ Squamous Metaplasia ☐ Mild Squamous Dysplasia ☐ Moderately Squamous Dysplasia ☐ High Squamous Dysplasia ☐ In Situ Squamous Carcinoma ☐ Atypical Adenomatous Hyperplasia ☐ Tumorlet	Analysis of Lymph Node Station <table border="1"> <thead> <tr> <th></th> <th>+</th> <th>Total</th> <th></th> <th>+</th> <th>Total</th> <th></th> <th>+</th> <th>Total</th> <th></th> <th>+</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td>S1</td> <td> 0 </td> <td> 0 </td> <td>S2</td> <td> 0 </td> <td> 0 </td> <td>S3</td> <td> 0 </td> <td> 0 </td> <td>S4</td> <td> 0 </td> <td> 0 </td> </tr> <tr> <td>S5</td> <td> 0 </td> <td> 0 </td> <td>S6</td> <td> 0 </td> <td> 0 </td> <td>S7</td> <td> 0 </td> <td> 0 </td> <td>S8</td> <td> 0 </td> <td> 0 </td> </tr> <tr> <td>S9</td> <td> 0 </td> <td> 0 </td> <td>S10</td> <td> 0 </td> <td> 0 </td> <td>S11</td> <td> 0 </td> <td> 0 </td> <td>S12</td> <td> 0 </td> <td> 0 </td> </tr> <tr> <td>S13</td> <td> 0 </td> <td> 0 </td> <td>S14</td> <td> 0 </td> <td> 0 </td> <td>NS</td> <td> 0 </td> <td> 0 </td> <td>Total</td> <td> 0 </td> <td> 0 </td> </tr> </tbody> </table> ☐ Positive Contralateral LN mtt Path N: n/a Type of LN metastasis: ☐ Intranodal ☐ Capsular ☐ Subcapsular ☐ Perinodal Size of Metastasis: Min. Size: 0 Max. Size: 0		+	Total		+	Total		+	Total		+	Total	S1	0	0	S2	0	0	S3	0	0	S4	0	0	S5	0	0	S6	0	0	S7	0	0	S8	0	0	S9	0	0	S10	0	0	S11	0	0	S12	0	0	S13	0	0	S14	0	0	NS	0	0	Total	0	0
	+	Total		+	Total		+	Total		+	Total																																																		
S1	0	0	S2	0	0	S3	0	0	S4	0	0																																																		
S5	0	0	S6	0	0	S7	0	0	S8	0	0																																																		
S9	0	0	S10	0	0	S11	0	0	S12	0	0																																																		
S13	0	0	S14	0	0	NS	0	0	Total	0	0																																																		

9) ReVITALization's pathological module: Tissue Bank (Frozen and Paraffin)

Tissue Bank: s-04-23495				Status: Ready...				
Click here to add 1 more row(s). Save it Cancel it Open it Save and Close Excel Report								
		Frozen	FFPE	TBID	Collection Date	Frozen Avail	FFPE Avail	
				6	4/12/2008	☒	☐	
				8	4/13/2008	☐	☐	
				9	4/12/2008	☐	☐	
Frozen								
Click here to add 1 more row(s). Save it Cancel it Open it Save and Close Excel Report								
		ID		Tissue	Blood	Pleural		
		Frozen ID: 0 TB ID: 6 SPORE No: 0 TID No: 0	Normal Lung	☐	DNA Conc	0	DNA Conc	0
			Tumor	☐	DNA Vol	0	DNA Vol	0
			Bronchus	☐	DNA Quality	n/a	DNA Quality	n/a
			LN	☐	RNA Conc	0	RNA Conc	0
			Serum	☐	RNA Vol	0	RNA Vol	0
			Lymphocyte	☐	RNA Quality	n/a	RNA Quality	n/a
			Pleural	☐	Prot Conc	0	Prot Conc	0
						Prot Vol	0	Prot Vol
			Prot Quality	n/a	Prot Quality	n/a		
FFPE								
Click here to add 1 more row(s). Save it Cancel it Open it Save and Close Excel Report								
		FFPE ID	TB ID	Cabinet	Tray	Block	Slide	
		0	6			0	0	

10) ReVITALization's dictionary module

Admin
Projects
Histo-Pathology Lab
Logout

Dictionaries

Country
Update

Page 1 [2] [3] [4] [5] [...]

	Dictionaries	Set Order
+		
	Afghanistan	1
	Albania	2
	Algeria	3
	American Samoa	4
	Andorra	5
	Angola	6
	Anguilla	7
	Antarctica	8
	Antigua and Barbuda	9
	Arctic Ocean	10
	Argentina	11
	Armenia	12
	Aruba	13
	Ashmore and Cartier Islands	14
	Atlantic Ocean	15
	Australia	16
	Austria	17
	Azerbaijan	18
	Bahamas	19

Page 1 [2] [3] [4] [5] [...]

Page 1 of 14

11) Query Tool.

Admin ▾ Projects ▾ Histo-Pathology Lab ▾ Logout

Query Generator

Query Template Management

select query: Select ▾ open enter query name: private

Run Save Overwrite Delete Rename

Status: Ready

Query Criteria Reset Create a new query Go!

Field ID	Field Name	Operator	Field Value	And/Or
+	0	Select...	Select...	and

Field Selection Reset

Selected Fields	Unselect	Clear	Ordering Fields	Unselect	Clear	Sorting Fields	Unselect	Clear

asc desc

Clear Reset Select All

☒ All
 ☐ Accession
 ☐ Chemotherapy
 ☐ FFPE
 ☐ Frozen Bank
 ☐ General Info
 ☐ Histological Dx
 ☐ Other Malignancy

☐ Other Treatment
 ☐ Patient
 ☐ Radiology
 ☐ Staging
 ☐ Surgery
 ☐ Tissue Bank
 ☐ Tumor

☐ Acc No	☐ Date COPD Dx	☐ Invasive Heart	☐ Middle Name	☐ Preferred ID	☐ Site
☐ Accession No	☐ Date Other Tx	☐ Invasive Trachea	☐ MIM Hemoptysis	☐ Race	☐ Slide
☐ AccessionID	☐ Date Quit Smoke	☐ Invasive Vessel Boxes	☐ MIM Squamous Dysplasia	☐ Radiation Fibrosis	☐ Smoked >= 100
☐ Address	☐ Date Screening	☐ Invasive Vessel	☐ MIM Squamous Dysplasia	☐ Radio Clin Response	☐ Soft Tissue Margin
☐ Address1	☐ Date Staging	☐ Is Chemo	☐ More Than 1 Nodal	☐ Radio Comments	☐ Solid
☐ Address2	☐ Date Surgery	☐ Is NED	☐ Multiple Segmentectomy	☐ Radio CT Response	☐ Specimen Avail
☐ Age Start Smoke	☐ Diabetes	☐ Is Radio	☐ Multiple Wedge	☐ Radio MDA	☐ Specimen Type
☐ Alcohol	☐ DOB	☐ Is Surgery	☐ N Clin	☐ Radio Patho Response	☐ SPOB_Na
☐ Anesthesia	☐ Dx	☐ Last Name	☐ Necrosis	☐ Radio Site	☐ Squamous Metaplasia
☐ Asthma	☐ Dx Date	☐ LN	☐ NED Date	☐ Radio Start Date	☐ State
☐ Asthmatic Lung	☐ Dx Specimen Date	☐ LN Involve	☐ No Desks x Met	☐ Radio Stop Date	☐ Study ID
☐ Atypical Adeno Type	☐ Dx Variant	☐ LN Met Size	☐ No of Nodules	☐ Radio Treatment Option	☐ Sub Type
☐ Avg Cig a Day	☐ DxSp Access	☐ LN Met	☐ No Smoke Mo	☐ Radio Tumor Size After CT	☐ Surgery
☐ BAC	☐ DxSp BAC	☐ LN Met Cigarette	☐ No Smoke Years	☐ Radio Tumor Size After Patho	☐ Surgery Comments
☐ Biopsy	☐ DxSp Carcinoma Met Site	☐ LN Met Intramural	☐ Normal Branch Eps	☐ Radio Tumor Size Before Patho	☐ Surgery Procedure
☐ Black	☐ DxSp Dx	☐ LN Met Pericard	☐ Normal Lung Parenchyma	☐ Radiotherapy	☐ Surgical Date
☐ Blood DNA Conc	☐ DxSp Dx Variant	☐ LN Met Subepicard	☐ Normal Lung	☐ Radio Tumor Size Before CT	☐ Symp Chain
☐ Blood DNA Quality	☐ DxSp Fibrosis	☐ LN Met Size	☐ NS +	☐ Radical Ineff	☐ T Clin
☐ Blood DNA Volume	☐ DxSp Grade	☐ Laboratory	☐ NS Analyzed	☐ S1	☐ TMD
☐ Blood Protein Conc	☐ DxSp Inflammation	☐ Lymphocyte	☐ NSCLC Type	☐ S1 +	☐ Thyroid
☐ Blood Protein Quality	☐ DxSp Met Dx	☐ M Clin	☐ On Study	☐ S10	☐ TMD_No
☐ Blood Protein Volume	☐ DxSp Microcapillary	☐ Mdx P HIF	☐ Other Margins	☐ S10 +	☐ Time Line
☐ Blood RNA Conc	☐ DxSp Necrosis	☐ Malignant Pleural Eff	☐ Other MDA Ra	☐ S11	☐ Tissue DNA Conc
☐ Blood RNA Quality	☐ DxSp NSCLC Type	☐ Margin +	☐ Other Med History	☐ S11 +	☐ Tissue DNA Quality
☐ Blood RNA Volume	☐ DxSp Papillary	☐ Margin Residual	☐ Other Treatment Comments	☐ S12	☐ Tissue DNA Volume
☐ Branch Hyperplasia	☐ DxSp Sarcoma Met Site	☐ MDA	☐ Papillary	☐ S12 +	☐ Tissue Protein Conc
☐ Branchial Margin	☐ DxSp Solid	☐ MDAH	☐ Parenchymal Margins	☐ S13	☐ Tissue Protein Quality
☐ Branchial	☐ DxSp Sub Type	☐ Mediastinal Pleura	☐ Parenchymal Pleura	☐ S13 +	☐ Tissue Protein Volume
☐ Branchial	☐ DxSp Sub Type	☐ Mediastinal Fat	☐ Path M	☐ S14	☐ Tissue RNA Conc
☐ Branchial Sample	☐ DxSp Sub Type	☐ Met Date	☐ Path N	☐ S14 +	☐ Tissue RNA Quality
☐ Cabazit	☐ DxSp Sub Type	☐ Met Dx	☐ Path T	☐ S2	☐ Tissue RNA Volume
☐ Carcinoma Met Site	☐ DxSp Sub Type	☐ Met Specimen Avail	☐ Path Type	☐ S2 +	☐ Total +
☐ Cardio	☐ DxSp Sub Type	☐ Met Specimen Type	☐ Patho N	☐ S3	☐ Total Analyzed
☐ Cardio DVT	☐ DxSp Sub Type	☐ Met Tumor Site	☐ Patho T	☐ S3 +	☐ Ttry
☐ Cell Phone	☐ DxSp Sub Type	☐ Metyp Access	☐ Patient Comments	☐ S4	☐ Tumor
☐ Chemo Drug A	☐ DxSp Sub Type	☐ Metyp BAC	☐ Patient ID	☐ S4 +	☐ Tumor Invasions
☐ Chemo Drug B	☐ DxSp Sub Type	☐ Metyp Carcinoma Met Site	☐ Pericardium	☐ S5	☐ Tumor Invasions Nodal
☐ Chemo Drug C	☐ DxSp Sub Type	☐ Metyp Dx	☐ Phrenic Nerve	☐ S5 +	☐ Tumor Invasions Other
☐ Chemo MDA	☐ DxSp Sub Type	☐ Metyp Dx Variant	☐ Pleural	☐ S6	☐ Tumor Invasions Pleural
☐ Chemo Start Date	☐ DxSp Sub Type	☐ Metyp Fibrosis	☐ Pleural Attached	☐ S6 +	☐ Tumor Invasions Vascular
☐ Chemo Stop Date	☐ DxSp Sub Type	☐ Metyp Grade	☐ Pleural DNA Conc	☐ S7	☐ Tumor Location
☐ Chemo Type	☐ DxSp Sub Type	☐ Metyp Inflammation	☐ Pleural DNA Quality	☐ S7 +	☐ Tumor Organ Site
☐ Chemotherapy	☐ DxSp Sub Type	☐ Metyp Met Dx	☐ Pleural DNA Volume	☐ S8	☐ Tumor Size
☐ Chest Wall	☐ DxSp Sub Type	☐ Metyp Microcapillary	☐ Pleural Eff	☐ S8 +	☐ Tumor Size
☐ City	☐ DxSp Sub Type	☐ Metyp Necrosis	☐ Pleural Protein Conc	☐ S9	☐ Tumor Type
☐ Clin Stage	☐ DxSp Sub Type	☐ Metyp NSCLC Type	☐ Pleural Protein Quality	☐ S9 +	☐ Tumorlet
☐ Collection Date	☐ DxSp Sub Type	☐ Metyp Papillary	☐ Pleural Protein Volume	☐ Sarcoma Met Site	☐ Tx
☐ COPD	☐ DxSp Sub Type	☐ Metyp Sarcoma Met Site	☐ Pleural RNA Conc	☐ Satellite Tumor	☐ Vagus Nerve
☐ Country	☐ DxSp Sub Type	☐ Metyp Solid	☐ Pleural RNA Quality	☐ Sarcom	☐ Work Phone
☐ Current Situation	☐ DxSp Sub Type	☐ Metyp Sub Type	☐ Pleural RNA Volume	☐ Single Segmentectomy	☐ Zip
☐ Current Smoker	☐ DxSp Sub Type	☐ Metyp Sub Type	☐ Pneumothorax	☐ Single Wedge	

12) The query results page

Query Name: Adenocarcinoma																										
Record Count: 13																										
Accession No	Biopsy	Dx Specimen Date	Event	Location	Met Date	Met Specimen Avail	Met Specimen Type	Met Tumor Site	Multiple Segmentectomy	Multiple Wedge	No. of Nodes	Path M	Path N	Path I	Path Type	Pneumonectomy	Single Segmentectomy	Single Wedge	Specimen Avail	Specimen Type	Surgical Date	Tumor Site	Acinar	Ductal	Carcinoma Met Site	Diagnosis
S-03-000995	False		Surgical Specimen	False					False	False	1	M0	N0	T2	Primary	False	False	True			1/8/2003 12:00:00 AM	35	5	n/a	NSCLC	
S-03-015760	False		Surgical Specimen	False					False	False	1	M0	N0	T1	Primary	False	False	True			4/2/2003 12:00:00 AM	80	0	n/a	NSCLC	
S-03-018420	False		Surgical Specimen	True					False	False	1	M0	N2	T1	Primary	False	False	False			4/16/2003 12:00:00 AM	0	0	n/a	NSCLC	n
S-03-020423	False		Surgical Specimen	False					False	False	1	M0	N2	T2	Primary	False	False	True			4/29/2003 12:00:00 AM	60	0	n/a	NSCLC	
S-03-024985	False		Surgical Specimen	True					False	False	1	M0	N0	T1	Primary	False	False	False			5/23/2003 12:00:00 AM	20	80	n/a	NSCLC	
S-03-052181	False		Surgical Specimen	True					False	False	3	M0	N2	T4	Primary	False	False	False			10/23/2003 12:00:00 AM	0	0	n/a	NSCLC	n
S-04-051605	False		Surgical Specimen	True					False	False	1	M0	N0	T2	Primary	False	False	False			10/6/2004 12:00:00 AM	25	75	n/a	NSCLC	
S-04-056678	False		Surgical Specimen	True					False	False	2	M1	N0	T2	Primary	False	False	True			11/2/2004 12:00:00 AM	35	50	n/a	NSCLC	
S-05-020799	False		Surgical Specimen	True					False	False	1	M0	N0	T2	Primary	False	False	False			4/18/2005 12:00:00 AM	0	0	n/a	NSCLC	
S-05-029775	False		Surgical Specimen	True					False	False	1	M0	N1	T3	Primary	False	False	False			6/6/2005 12:00:00 AM	0	0	n/a	NSCLC	
S-05-050475	False		Surgical Specimen	True					False	False	1	M0	N2	T2	Primary	False	False	False			9/19/2005 12:00:00 AM	0	0	n/a	NSCLC	n
S-05-065030	False		Surgical Specimen	False					False	False	1	M0	N0	T2	Primary	False	False	True			12/7/2005 12:00:00 AM	0	0	n/a	NSCLC	
S-05-068150	False		Surgical Specimen	True					False	False	1	M0	N1	T2	Primary	False	False	False			12/27/2005 12:00:00 AM	0	0	n/a	NSCLC	e

13) The example of the ReVITALization's Excel reports.

	K	L	M	N	O	P	Q
1	Specimen Avail	Surgical Date	Single Wedge	Multiple Wedge	Single Segmentectomy	Multiple Segmentectomy	Lobectomy
2							
3							
4							
5		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
6		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
7		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
8		5/11/2005 0:00	FALSE	FALSE	FALSE	FALSE	TRUE
9							
10							
11							
12							
13		1/1/1900 0:00	FALSE	TRUE	TRUE	FALSE	FALSE
14		1/1/1900 0:00	FALSE	TRUE	TRUE	FALSE	FALSE
15	FALSE						
16	TRUE						
17							
18							
19		1/1/1900 0:00	FALSE	FALSE	FALSE	FALSE	FALSE
20							
21							
22							
23							
24							
25							
26		1/1/1900 0:00	FALSE	FALSE	FALSE	FALSE	FALSE
27							
28							
29							
30							
31							
32							
33							
34							
35							

14). Patient Summary Report

PATIENT SUMMARY

Vital Status

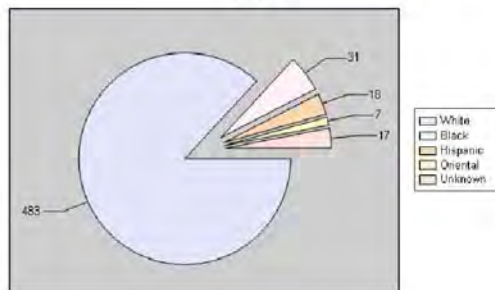
12/28/2009 11:02:10 AM

	Number of Patients	Percentage
Deceased Patients	230	41.37 %
Alive Patients	326	58.63 %
Total	556	100.00 %

Gender

	Number of Patients	Percentage
Male	272	48.92 %
Female	284	51.08 %
Unknown	0	0.00 %
Total	556	100.00 %

Race



Statistical Report for REVITALIZATION Study

November 22, 2010

Table of Contents

Title	Page
I. Data	2
II. Statistical Methods	2
III. Main findings	3
IV. Results	5
1. Patient Population	5
2. Dr. HY Lee	10
2.1. Markers	10
2.2. Overall Survival	35
2.3. Recurrence Free Survival	41
3. Dr. Koo	48
3.1. Markers	48
3.2. Overall Survival	59
3.3. Recurrence Free Survival	63
4. Dr. Lotan	67
4.1. Markers	67
4.2. Overall Survival	78
4.3. Recurrence Free Survival	82
5. All Markers	85

I. Data:

We received biomarker readings from 542 patients. A total of 21 markers have been studied by Drs. HY Lee (IGF1R, IGFBP3, InsulinR, pAKT, pIGF1R, pSRC, pmTOR, pAMPK, pEGFR, pS6), Lotan (FEN1, MCM2, MCM6, SFN, TPX2, UBE2C) and Koo (CASK, CD51, CXCR2, EpCAM, SPP1).

Merging patients' demographic, diagnosis/histology, treatment info and follow-up information with marker data, excluding patients with stage IIIB and IV or wedge resection, a total of 370 patients are included in this analysis.

II. Statistical Methods:

Summary statistics, including frequency tabulation, means, standard deviations, median, and range, were given to describe subject characteristics and biomarkers. Wilcoxon rank sum test or Kruskal-Wallis test was used to test the difference of markers between/among categorical variable levels. The association between gender, histology and smoking status, treatment and stage is test using χ^2 test or Fisher's exact test when appropriate. The continuous markers were dichotomized by either 0 vs positive or median. The Kaplan-Meier method was used to construct overall and progression-free survival curves and log-rank test was used to test the difference in survival by covariates. Univariate and multivariate Cox model were fitted to estimate the effect of prognostic factors, including age, gender, histology, stage, markers (both continuous and dichotomized levels) on time to event endpoints, including overall survival (OS) and recurrence free survival (RFS). All statistical tests were two-sided, and p values of 0.05 or less were considered to be statistically significant.

The predictive accuracy of Cox regression models is quantified by C-index, which provides the area under the receiver operating characteristics curve for censored data [1, 2]. A C-index of 0.5 indicates that outcomes are completely random, whereas a C-index of 1 indicates that the model is a perfect predictor. To protect against overfitting during stepwise regression, we used bootstrap internal-validation, which allows for computation of an unbiased estimate of predictive accuracy, C-index. We also validate the models for calibration accuracy in predicting the probability of surviving 1 year, 3 years or 5 years, or probability of recurrence-free recurrence free at 1 year, 3 years or 5 years. We use 200 bootstrap samples in both bootstrap validation and calibration.

[1] Harrell F. Regression modeling strategies. New York: Springer-Verlag; 2001.

[2] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.

Notation for marker names:

Original readings:	marker name _*	*=c	cytoplasmic
		=m	membrane
		=n	nucleus
Dichotomized 1:	marker name _*_01	0 = 0	

Dichotomized 2:

marker name _*_01m

1 = positive reading

0 = < median

1 = ≥median

III. Main Findings:

- 1) Patients demographic information, disease characteristics and treatment information are summarized in Table 1 (page 5).
- 2) Gender and histology are significantly associated with smoking status, more former/current smokers in male than female patients, and more former/current smokers in patients with SCC than with ADENO. Treatment is significantly associated with stage of the disease. Patients with more severe diseases receive more adjuvant or neo-adjuvant treatment.
- 3) Please find the correlation between markers and comparison of markers by patient characteristics and disease status in the report.
- 4) With median follow-up time of 5.3 years, 160 deaths have been observed. Median survival=6.4 years. A total of 209 cases with recurrence or deaths have been recorded. In univariate analysis, age, gender, stage of the disease and certain adjuvant or neo-adjuvant treatment are significantly associated with OS (Table 4, page 8). Gender, stage, necrosis, inflammation, and certain adjuvant or neo-adjuvant treatment are found to be significantly associated with RFS (Table 5, page 9).
- 5) **Dr. HY Lee's data:** Adjusted for age, gender, stage, neoadjuvant (including chemo and radiation or concurrent therapy), both positive cytoplasmic pAMPK (HR=0.68, 95% CI: (0.49, 0.95), p=0.02) and positive cytoplasmic pmTOR (HR=0.62, 95% CI: (0.44, 0.89), p=0.0085) predict longer OS (page 41). When assessing effects of covariates on RFS, adjusted for age, stage and neoadjuvant treatment, cytoplasmic IGF1R (HR=1.003, 95% CI: (0.999, 1.007), p=0.10) and positive membrane insulin (HR=1.49, 95% CI: (1.12, 1.98), p=0.006) are associated with shorter RFS, whereas both positive cytoplasmic pAMPK (HR=0.61, 95% CI: (0.46, 0.82), p=0.001) and positive cytoplasmic pmTOR (HR=0.67, 95% CI: (0.49, 0.93), p=0.015) are significant predictors for longer RFS (page 47).
- 6) **Dr. Koo's data:** Adjusted for age, gender, stage, neoadjuvant (including chemo and radiation or concurrent therapy), higher cytoplasmic CXCR2 (larger than median) (HR=1.55, 95% CI: (1.12, 2.15), p=0.008) is associated with poor OS and positive cytoplasmic EpCAM (HR=0.61, 95% CI: (0.44, 0.84), p=0.003) predicts longer OS (page 63). When assessing effects of covariates on RFS, adjusted for age, stage and neoadjuvant treatment, higher cytoplasmic CXCR2 (larger than median) (HR=1.35, 95% CI: (1.01, 1.80), p=0.04) is associated with poor RFS, while positive cytoplasmic EpCAM (HR=0.69, 95% CI: (0.52, 0.92), p=0.01) and membrane CASK (HR=0.996, 95% CI: (0.99, 1.00) predict longer RFS (page 66).
- 7) **Dr. Lotan's data:** Adjusted for age, stage, neoadjuvant (including chemo and radiation or concurrent therapy), both higher nuclear FEN1 (larger than median) (HR=1.27, 95% CI: (0.90, 1.80), p=0.18) and positive nuclear MCM6 (HR=1.62, 95% CI: (0.91, 2.88), p=0.098) are associated, however not significantly, with poor OS (page 81). No significant marker was identified in predicting RFS in multicovariate analysis.

- 8) **All markers: OS** (page 85) – Adjusted for age, histology, stage and neoadjuvant therapy, positive cytoplasmic pAMPK, positive cytoplasmic pmTOR and positive cytoplasmic EpCAM are significant predictors in longer OS, whereas higher cytoplasmic CXCR2 (larger than median) and higher nuclear FEN1 (larger than median) are significant predictors in shorter OS. Predictive accuracy of the model from internal validation is 0.67. **RFS** (page 86) -- Adjusted for age, histology, stage and neoadjuvant therapy, positive membrane insulin and higher cytoplasmic CXCR2 (larger than median) are significant predictors in poor RFS, whereas positive cytoplasmic pAMPK, positive cytoplasmic pmTOR, positive cytoplasmic EpCAM and higher membrane CASK are associated with longer RFS. Predictive accuracy of the model from internal validation is 0.66.

IV. Results:

1. Patient population (N=370):

Table 1. Demographic and pathological information

<i>covariate</i>	<i>Levels</i>	<i>N(%)</i>	<i>covariate</i>	<i>Levels</i>	<i>N(%)</i>
Gender	F	184(49.7%)	Adj_Chem	No	249(70.7%)
	M	186(50.3%)		Yes	103(29.3%)
Race	Black	21(5.7%)	Adj_XRT	Unknown	18
	Hispanic	14(3.8%)		No	307(87.0%)
	Oriental	5(1.4%)		Yes	46(13.0%)
	White	330(89.2%)	Adj_Concurrent	Unknown	17
Tobacco	No	38(10.3%)		No	346(97.7%)
	Yes	332(89.7%)		Yes	8(2.3%)
Smoker	Current	162(43.8%)	Adjuvant	Unknown	16
	Former	170(45.9%)		No	224(63.6%)
	Never	38(10.3%)		Yes	128(36.4%)
pathT	T1	135(36.5%)	Neo_Chemo	Unknown	18
	T2	214(57.8%)		No	318(86.6%)
	T3	21(5.7%)		Yes	49(13.4%)
pathN	N0	246(67.2%)	Neo_XRT	Unknown	3
	N1	69(18.9%)		No	365(99.5%)
	N2	51(13.9%)		Yes	2(0.5%)
	Unknown	4	Neo_Concurrent	Unknown	3
Path stage	IA	103(27.8%)		No	362(98.6%)
	IB	131(35.4%)		Yes	5(1.4%)
	IIA	22(5.9%)	NeoAdjuvant	Unknown	3
	IIB	53(14.3%)		No	313(85.3%)
	IIIA	61(16.5%)		Yes	54(14.7%)
Histology	ADENO	227(61.4%)	Adjvant/NeoAdjvant	Unknown	3
	Other	17(4.6%)		No	194(54.8%)
	SCC	126(34.1%)		Yes	160(45.2%)
Grade	Poorly	122(34.2%)	Diabetes	Unknown	16
	Moderately	199(55.7%)		No	341(92.9%)

<i>covariate</i>	<i>Levels</i>	<i>N(%)</i>	<i>covariate</i>	<i>Levels</i>	<i>N(%)</i>
Inflammation	Well	36(10.1%)	Metformin	Yes	26(7.1%)
	Unknown	9		Unknown	3
	Mild	154(42.7%)		No	276(95.5%)
	Moderately	145(40.2%)		Yes	13(4.5%)
	Severe	62(17.2%)		Unknown	81
	Unknown	9			

<i>Smoking Status</i>					
<i>covariate</i>	<i>levels</i>	<i>Never</i>	<i>Former</i>	<i>Current</i>	<i>p_value</i>
Gender	F	30(16.3%)	78(42.4%)	76(41.3%)	.0007
	M	8(4.3%)	92(49.5%)	86(46.2%)	
Histology	ADENO	37(16.3%)	103(45.4%)	87(38.3%)	<.0001
	SCC	1(0.8%)	61(48.4%)	64(50.8%)	
	Other	0	6(35.3%)	11(64.7%)	

Table 2. Treatment by Stage

<i>Stage</i>					
<i>covariate</i>	<i>levels</i>	<i>I</i>	<i>II</i>	<i>IIIA</i>	<i>p_value</i>
adj_chem	No	182(80.9%)	37(52.1%)	30(53.6%)	<.0001
	Yes	43(19.1%)	34(47.9%)	26(46.4%)	
adj_XRT	No	219(97.3%)	63(88.7%)	25(43.9%)	<.0001
	Yes	6(2.7%)	8(11.3%)	32(56.1%)	
neo_chemo	No	211(90.6%)	66(89.2%)	41(68.3%)	<.0001
	Yes	22(9.4%)	8(10.8%)	19(31.7%)	
Neoadj	No	208(89.3%)	65(87.8%)	40(66.7%)	<.0001
	Yes	25(10.7%)	9(12.2%)	20(33.3%)	
AdjNeoAdj	No	160(70.8%)	25(35.2%)	9(15.8%)	<.0001
	Yes	66(29.2%)	46(64.8%)	48(84.2%)	

Table 3. Continuous variables

<i>covariate</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>
Age	370	65.71 ± 10.17, 66.31 (32.24, 89.96)
Necrosis	369	11.95 ± 15.67, 5 (0, 90)
Fibrosis	369	21.82 ± 18.85, 20 (0, 95)

<i>Table of RECI by OSI</i>			
<i>RECURRENCE</i>	<i>OS</i>		
<i>Frequency</i>	<i>Alive</i>	<i>Dead</i>	<i>Total</i>
<i>No</i>	161	56	217
<i>Yes</i>	49	104	153
<i>Total</i>	210	160	370

Table 4. Univariate Cox model assessing effects of common covariates on overall survival (OS), median survival=6.4 years, median follow-up time = 5.3 Years.

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
Age	0.0154	0.0082	1.0155	0.9994	1.0320	0.0596	370	160	210
Gender (M vs F)	0.3940	0.1598	1.4829	1.0842	2.0282	0.0137	370	160	210
Race (White vs Other)	-0.2604	0.2446	0.7707	0.4772	1.2447	0.2869	370	160	210
Tobacco (Yes vs No)	0.1098	0.2716	1.1160	0.6554	1.9003	0.6861	370	160	210
Smoke							370	160	210
Former vs Never	0.0193	0.2851	1.019	0.583	1.783	0.9461			
Current vs Never	0.2006	0.2830	1.222	0.702	2.128	0.4784			
Histology							370	160	210
ADENO vs SCC	-0.1855	0.1661	0.831	0.600	1.150	0.2639			
Other vs SCC	-0.3688	0.4287	0.692	0.299	1.602	0.3896			
Grade							357	155	202
Poor vs Well	0.3337	0.3101	1.396	0.760	2.564	0.2818			
Moderate vs Well	0.3554	0.2967	1.427	0.798	2.552	0.2310			
pathN (N1-3 vs N0)	0.5782	0.1621	1.7829	1.2976	2.4497	0.0004	366	157	209
Stage							370	160	210
II vs I	0.4009	0.1971	1.493	1.015	2.197	0.0420			

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
III vs I	0.8296	0.1973	2.292	1.557	3.375	<.0001			
Fibrosis	-0.0067	0.0045	0.9933	0.9845	1.0021	0.1373	369	159	210
Necrosis	0.0076	0.0049	1.0076	0.9980	1.0173	0.1212	369	159	210
Inflammation							361	155	206
Moderate vs Mild	-0.0557	0.1755	0.946	0.671	1.334	0.7509			
Severe vs Mild	-0.2722	0.2368	0.762	0.479	1.212	0.2503			
Neo_Chemo (Yes vs No)	0.4800	0.2085	1.6161	1.0739	2.4321	0.0213	367	158	209
Adj_chem (Yes vs No)	-0.2177	0.1894	0.8044	0.5550	1.1659	0.2504	352	153	199
Adj_XRT (Yes vs No)	0.7453	0.2070	2.1072	1.4043	3.1618	0.0003	353	153	200
Neoadj (Yes vs No)	0.5462	0.2006	1.7266	1.1653	2.5583	0.0065	367	158	209
Adjuvant (Yes vs No)	0.0009	0.1725	1.0009	0.7138	1.4035	0.9957	352	153	199
AdjNeoAdj (Yes vs No)	0.1661	0.1631	1.1807	0.8576	1.6255	0.3085	354	154	200

Table 5. Univariate Cox model assessing effects of common covariates on recurrence free survival (RFS)

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
Age	0.0081	0.0070	1.0082	0.9944	1.0222	0.2475	370	209	161
Gender (M vs F)	0.3123	0.1395	1.3665	1.0397	1.7962	0.0252	370	209	161
Race (White vs Other)	-0.1981	0.2212	0.8203	0.5317	1.2654	0.3703	370	209	161
Tobacco (Yes vs No)	0.0747	0.2353	1.0775	0.6795	1.7087	0.7510	370	209	161
Smoke							370	209	161
Former vs Never	0.0519	0.2457	1.053	0.651	1.705	0.8328			
Current vs Never	0.0993	0.2469	1.104	0.681	1.792	0.6874			
Histology									
ADENO vs SCC	-0.2131	0.1448	0.808	0.608	1.073	0.1413	370	209	161
Other vs SCC	-0.4543	0.3713	0.635	0.307	1.315	0.2211			
Grade							357	203	154
Poor vs Well	0.2713	0.2713	1.312	0.771	2.232	0.3172			
Moderate vs Well	0.3402	0.2594	1.405	0.845	2.336	0.1897			
pathN (N1-3 vs N0)	0.7049	0.1422	2.0237	1.5313	2.6743	<.0001	366	205	161
Stage							370	209	161
II vs I	0.5203	0.1719	1.683	1.201	2.357	0.0025			

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
III vs I	0.9498	0.1739	2.585	1.839	3.635	<.0001			
Fibrosis	-0.0055	0.0039	0.9945	0.9870	1.0021	0.1568	369	208	161
Necrosis	0.0090	0.0042	1.0090	1.0008	1.0173	0.0306	369	208	161
Inflammation							361	203	158
Moderate vs Mild	-0.2823	0.1538	0.754	0.558	1.019	0.0665			
Severe vs Mild	-0.4402	0.2094	0.644	0.427	0.971	0.0356			
Neo_Chemo (Yes vs No)	0.4614	0.1879	1.5864	1.0977	2.2926	0.0141	367	207	160
Adj_chem (Yes vs No)	0.0579	0.1564	1.0596	0.7799	1.4396	0.7114	352	199	153
Adj_XRT (Yes vs No)	0.7165	0.1850	2.0472	1.4247	2.9417	0.0001	353	199	154
Neoadj (Yes vs No)	0.5253	0.1799	1.6909	1.1885	2.4058	0.0035	367	207	160
Adjuvant (Yes vs No)	0.2109	0.1469	1.2348	0.9258	1.6469	0.1512	352	199	153
AdjNeoAdj (Yes vs No)	0.3134	0.1422	1.3681	1.0354	1.8077	0.0275	354	200	154

2. Dr. HY Lee

2.1 Markers

Table 6. Descriptive summary of markers

<i>covariate</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>covariate</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>
IGF1R_c	367	22.37 ± 36.14, 6.67 (0, 190)	pSRC_c	369	3.05 ± 7.01, 0 (0, 40)
IGF1R_m	367	3.67 ± 14.25, 0 (0, 120)	pSRC_m	370	17.24 ± 35.28, 0 (0, 190)
IGFBP3_c	370	24.61 ± 33.07, 10 (0, 180)	pmTOR_c	369	39.12 ± 51.89, 16.67 (0, 230)
IGFBP3_m	370	11.64 ± 31.5, 0 (0, 240)	pmTOR_m	370	61.52 ± 68.79, 36.67 (0, 240)
Insulin_c	367	44.11 ± 46.47, 30 (0, 186.67)	pAMPK_c	370	40.46 ± 49.02, 16.67 (0, 200)
Insulin_m	366	14.64 ± 28.41, 0 (0, 180)	pAMPK_m	370	2.33 ± 10.36, 0 (0, 110)
pAKT_c	369	3.68 ± 13.16, 0 (0, 86.67)	pEGFR_c	368	18.93 ± 37.45, 0 (0, 190)
pAKT_m	370	0.23 ± 1.96, 0 (0, 30)	pEGFR_m	369	16.57 ± 30.11, 3.33 (0, 160)
pAKT_n	370	1.1 ± 4.33, 0 (0, 40)	pEGFR_n	369	31.24 ± 38.83, 13.33 (0, 160)
pIGF1R_c	368	5.83 ± 13.58, 0 (0, 120)	pS6_c	370	44.55 ± 42.49, 33.33 (0, 200)
pIGF1R_m	368	13.76 ± 39.12, 0 (0, 240)	pS6_m	370	6.05 ± 20.32, 0 (0, 180)

Figure 1. Distribution of markers, Dr. HY Lee (Red line – Mean, Green line – Median)

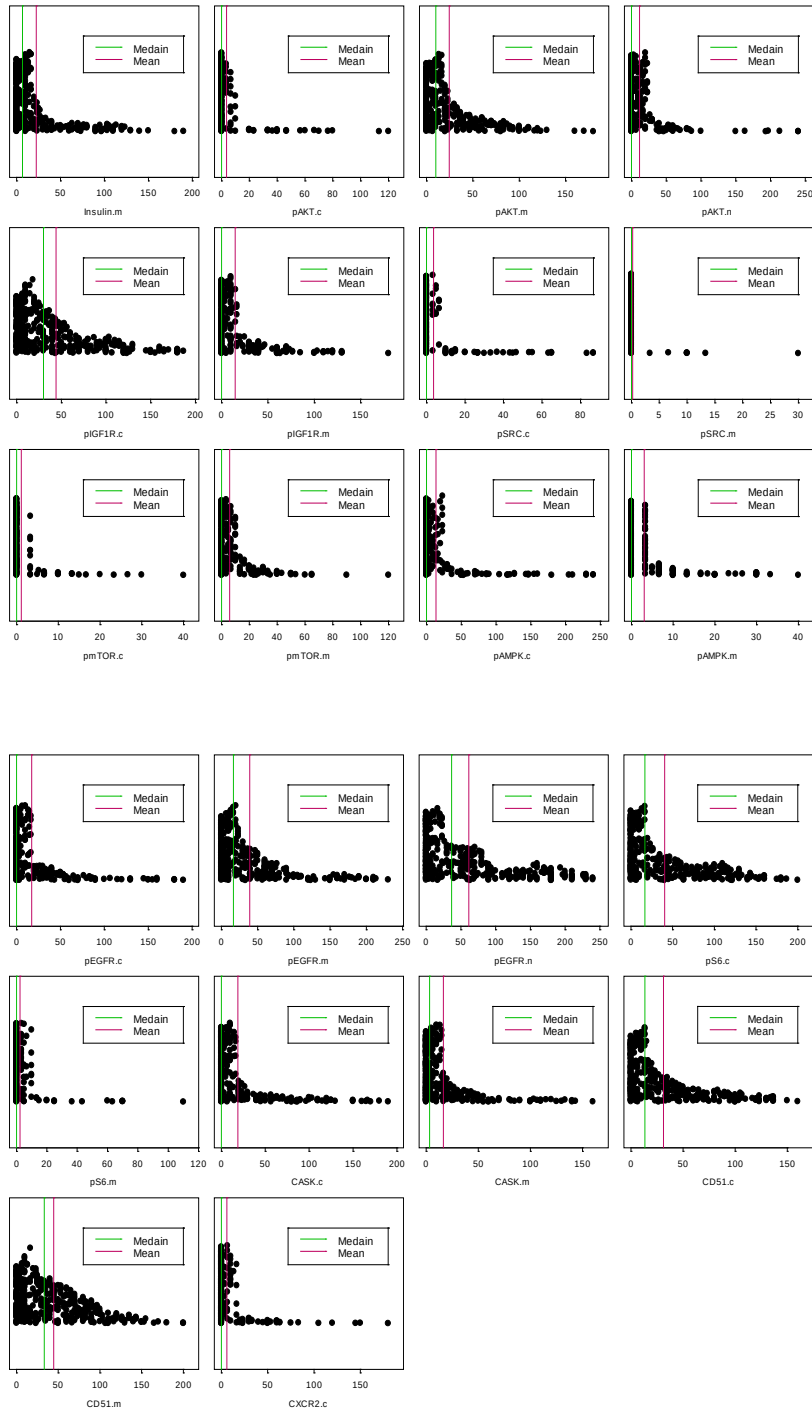


Table 7. Correlation between markers

Spearman Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations									
	IGF1R_c	IGF1R_m	IGFBP3_c	IGFBP3_m	Insulin_c	Insulin_m	pAKT_c	pAKT_m	pAKT_n
IGF1R_c	1	0.54200	0.12708	0.15158	-0.02499	-0.06343	0.22535	0.04874	0.19148
	<.0001	<.0001	0.0148	0.0036	0.6346	0.2280	<.0001	0.3518	0.0002
	367	367	367	367	364	363	366	367	367
IGF1R_m	0.54200	1	0.14013	0.12340	0.02801	-0.01045	0.04542	-0.05391	0.05867
	<.0001	<.0001	0.0072	0.0180	0.5943	0.8427	0.3863	0.3030	0.2622
	367	367	367	367	364	363	366	367	367
IGFBP3_c	0.12708	0.14013	1	0.60804	0.04104	-0.13357	0.07723	0.06800	0.00828
	0.0148	0.0072	<.0001	<.0001	0.4332	0.0105	0.1387	0.1919	0.8738
	367	367	370	370	367	366	369	370	370
IGFBP3_m	0.15158	0.12340	0.60804	1	-0.06140	-0.10337	0.12932	0.04957	0.05407
	0.0036	0.0180	<.0001	<.0001	0.2406	0.0481	0.0129	0.3416	0.2996
	367	367	370	370	367	366	369	370	370
Insulin_c	-0.02499	0.02801	0.04104	-0.06140	1	0.39421	0.03701	-0.03746	-0.06347
	0.6346	0.5943	0.4332	0.2406	<.0001	<.0001	0.4803	0.4744	0.2251
	364	364	367	367	367	366	366	367	367
Insulin_m	-0.06343	-0.01045	-0.13357	-0.10337	0.39421	1	-0.00196	0.05561	0.02261
	0.2280	0.8427	0.0105	0.0481	<.0001	<.0001	0.9703	0.2887	0.6664
	363	363	366	366	366	366	365	366	366
pAKT_c	0.22535	0.04542	0.07723	0.12932	0.03701	-0.00196	1	0.31983	0.55645
	<.0001	0.3863	0.1387	0.0129	0.4803	0.9703	<.0001	<.0001	<.0001
	366	366	369	369	366	365	369	369	369
pAKT_m	0.04874	-0.05391	0.06800	0.04957	-0.03746	0.05561	0.31983	1	0.22558
	0.3518	0.3030	0.1919	0.3416	0.4744	0.2887	<.0001	<.0001	<.0001
	367	367	370	370	367	366	369	370	370
pAKT_n	0.19148	0.05867	0.00828	0.05407	-0.06347	0.02261	0.55645	0.22558	1
	0.0002	0.2622	0.8738	0.2996	0.2251	0.6664	<.0001	<.0001	<.0001
	367	367	370	370	367	366	369	370	370
pIGF1R_c	0.04115	0.07205	0.06407	0.13376	0.10203	0.07733	0.07635	0.05123	0.01247
	0.4332	0.1696	0.2201	0.0102	0.0515	0.1409	0.1443	0.3270	0.8116
	365	365	368	368	365	364	367	368	368
pIGF1R_m	-0.00883	0.03090	0.02782	0.10603	0.06208	0.03266	0.06443	0.05009	0.02256
	0.8665	0.5562	0.5948	0.0421	0.2368	0.5345	0.2182	0.3379	0.6662
	365	365	368	368	365	364	367	368	368
pSRC_c	0.17186	0.13887	0.08455	0.18250	0.01820	0.05065	0.13423	0.10755	0.08164
	0.0010	0.0078	0.1049	0.0004	0.7286	0.3345	0.0099	0.0389	0.1175
	366	366	369	369	366	365	368	369	369

Spearman Correlation Coefficients									
Prob > r under H0: Rho=0									
Number of Observations									
	IGF1R_c	IGF1R_m	IGFBP3_c	IGFBP3_m	Insulin_c	Insulin_m	pAKT_c	pAKT_m	pAKT_n
pSRC_m	0.12782	0.11213	0.07093	0.20326	0.08381	0.09932	0.11583	0.12257	0.06602
	0.0143	0.0318	0.1734	<.0001	0.1089	0.0577	0.0261	0.0183	0.2051
	367	367	370	370	367	366	369	370	370
pmTOR_c	-0.29987	-0.22251	0.02532	-0.09283	0.14796	-0.01566	-0.01393	-0.06009	-0.03707
	<.0001	<.0001	0.6277	0.0749	0.0046	0.7656	0.7900	0.2495	0.4778
	366	366	369	369	366	365	368	369	369
pmTOR_m	-0.31648	-0.22256	-0.06881	-0.14827	0.11020	0.14343	0.00272	-0.03206	0.01129
	<.0001	<.0001	0.1866	0.0043	0.0348	0.0060	0.9584	0.5387	0.8286
	367	367	370	370	367	366	369	370	370

Spearman Correlation Coefficients						
Prob > r under H0: Rho=0						
Number of Observations						
	pIGF1R_c	pIGF1R_m	pSRC_c	pSRC_m	pmTOR_c	pmTOR_m
IGF1R_c	0.04115	-0.00883	0.17186	0.12782	-0.29987	-0.31648
	0.4332	0.8665	0.0010	0.0143	<.0001	<.0001
	365	365	366	367	366	367
IGF1R_m	0.07205	0.03090	0.13887	0.11213	-0.22251	-0.22256
	0.1696	0.5562	0.0078	0.0318	<.0001	<.0001
	365	365	366	367	366	367
IGFBP3_c	0.06407	0.02782	0.08455	0.07093	0.02532	-0.06881
	0.2201	0.5948	0.1049	0.1734	0.6277	0.1866
	368	368	369	370	369	370
IGFBP3_m	0.13376	0.10603	0.18250	0.20326	-0.09283	-0.14827
	0.0102	0.0421	0.0004	<.0001	0.0749	0.0043
	368	368	369	370	369	370
Insulin_c	0.10203	0.06208	0.01820	0.08381	0.14796	0.11020
	0.0515	0.2368	0.7286	0.1089	0.0046	0.0348
	365	365	366	367	366	367
Insulin_m	0.07733	0.03266	0.05065	0.09932	-0.01566	0.14343
	0.1409	0.5345	0.3345	0.0577	0.7656	0.0060
	364	364	365	366	365	366
pAKT_c	0.07635	0.06443	0.13423	0.11583	-0.01393	0.00272
	0.1443	0.2182	0.0099	0.0261	0.7900	0.9584
	367	367	368	369	368	369
pAKT_m	0.05123	0.05009	0.10755	0.12257	-0.06009	-0.03206
	0.3270	0.3379	0.0389	0.0183	0.2495	0.5387
	368	368	369	370	369	370

Spearman Correlation Coefficients						
Prob > r under H0: Rho=0						
Number of Observations						
	<i>pIGF1R_c</i>	<i>pIGF1R_m</i>	<i>pSRC_c</i>	<i>pSRC_m</i>	<i>pmTOR_c</i>	<i>pmTOR_m</i>
<i>pAKT_n</i>	0.01247 0.8116 368	0.02256 0.6662 368	0.08164 0.1175 369	0.06602 0.2051 370	-0.03707 0.4778 369	0.01129 0.8286 370
<i>pIGF1R_c</i>	1 <.0001 368	0.81730 <.0001 368	0.65536 <.0001 367	0.66182 <.0001 368	0.12378 0.0177 367	-0.03251 0.5341 368
<i>pIGF1R_m</i>	0.81730 <.0001 368	1 <.0001 368	0.60625 <.0001 367	0.59282 <.0001 368	0.15051 0.0039 367	0.02149 0.6812 368
<i>pSRC_c</i>	0.65536 <.0001 367	0.60625 <.0001 367	1 <.0001 369	0.82219 <.0001 369	0.09120 0.0806 368	-0.02066 0.6925 369
<i>pSRC_m</i>	0.66182 <.0001 368	0.59282 <.0001 368	0.82219 <.0001 369	1 <.0001 370	0.04124 0.4296 369	-0.06998 0.1792 370
<i>pmTOR_c</i>	0.12378 0.0177 367	0.15051 0.0039 367	0.09120 0.0806 368	0.04124 0.4296 369	1 <.0001 369	0.73911 <.0001 369
<i>pmTOR_m</i>	-0.03251 0.5341 368	0.02149 0.6812 368	-0.02066 0.6925 369	-0.06998 0.1792 370	0.73911 <.0001 369	1 <.0001 370

Spearman Correlation Coefficients							
Prob > r under H0: Rho=0							
Number of Observations							
	pAMPK_c	pAMPK_m	pEGFR_c	pEGFR_m	pEGFR_n	pS6_c	pS6_m
pAMPK_c	1	0.41941	0.25539	0.14190	-0.03197	-0.05288	-0.11848
	<.0001	<.0001	<.0001	0.0063	0.5404	0.3104	0.0226
	370	370	368	369	369	370	370
pAMPK_m	0.41941	1	0.14041	0.12172	0.07484	-0.08216	0.04847
	<.0001	<.0001	0.0070	0.0193	0.1513	0.1146	0.3525
	370	370	368	369	369	370	370
pEGFR_c	0.25539	0.14041	1	0.67646	0.02341	0.19825	-0.01865
	<.0001	0.0070	<.0001	<.0001	0.6549	0.0001	0.7213
	368	368	368	368	367	368	368
pEGFR_m	0.14190	0.12172	0.67646	1	0.06551	0.18835	0.09597
	0.0063	0.0193	<.0001	<.0001	0.2099	0.0003	0.0656
	369	369	368	369	368	369	369
pEGFR_n	-0.03197	0.07484	0.02341	0.06551	1	0.21804	0.20692
	0.5404	0.1513	0.6549	0.2099	<.0001	<.0001	<.0001
	369	369	367	368	369	369	369
pS6_c	-0.05288	-0.08216	0.19825	0.18835	0.21804	1	0.29661
	0.3104	0.1146	0.0001	0.0003	<.0001	<.0001	<.0001
	370	370	368	369	369	370	370
pS6_m	-0.11848	0.04847	-0.01865	0.09597	0.20692	0.29661	1
	0.0226	0.3525	0.7213	0.0656	<.0001	<.0001	<.0001
	370	370	368	369	369	370	370

<i>covariate</i>	<i>levels</i>	<i>N(%)</i>	<i>covariate</i>	<i>levels</i>	<i>N(%)</i>
	1	48(13.1%)		0	209(56.8%)
IGFBP3_c_01	0	118(31.9%)		1	159(43.2%)
	1	252(68.1%)	pEGFR_m_01	missing	1
IGFBP3_m_01	0	245(66.2%)		0	178(48.2%)
	1	125(33.8%)		1	191(51.8%)
Insulin_c_01	missing	3	pEGFR_n_01	missing	1
	0	51(13.9%)		0	107(29%)
	1	316(86.1%)		1	262(71%)
Insulin_m_01	missing	4	pS6_c_01	0	54(14.6%)
	0	212(57.9%)		1	316(85.4%)
	1	154(42.1%)	pS6_m_01	0	297(80.3%)
pAKT_c_01	missing	1		1	73(19.7%)
	0	318(86.2%)	IGF1R_c_01m	Missing	3
	1	51(13.8%)		0	182(49.6%)
pAKT_m_01	0	363(98.1%)		1	185(50.4%)
	1	7(1.9%)	IGFBP3_c_01m	0	172(46.5%)
pAKT_n_01	0	332(89.7%)		1	198(53.5%)
	1	38(10.3%)	Insulin_c_01m	missing	3
pIGF1R_c_01	missing	2		0	179(48.8%)
	0	247(67.1%)		1	188(51.2%)
	1	121(32.9%)	pmTOR_c_01m	missing	1
pIGF1R_m_01	missing	2		0	181(49.1%)
	0	257(69.8%)		1	188(50.9%)
	1	111(30.2%)	pmTOR_m_01m	0	181(48.9%)
pSRC_c_01	missing	1		1	189(51.1%)
	0	268(72.6%)	pEGFR_m_01m	missing	1
	1	101(27.4%)		0	178(48.2%)
pSRC_m_01	0	225(60.8%)		1	191(51.8%)
	1	145(39.2%)	pEGFR_n_01m	missing	1
pmTOR_c_01	missing	1		0	174(47.2%)
	0	91(24.7%)		1	195(52.8%)
	1	278(75.3%)	pS6_c_01m	0	180(48.6%)

<i>covariate</i>	<i>levels</i>	<i>N(%)</i>	<i>covariate</i>	<i>levels</i>	<i>N(%)</i>
pmTOR_m_01	0	90(24.3%)		1	190(51.4%)
	1	280(75.7%)			

Table 9. Markers by gender

<i>covariate</i>	<i>Gender</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
IGF1R_c	F	182	17.55 ± 30.6, 3.33 (0, 140)	.0088
	M	185	27.11 ± 40.39, 10 (0, 190)	
IGF1R_m	F	182	2.16 ± 9.41, 0 (0, 66.67)	.0623
	M	185	5.15 ± 17.67, 0 (0, 120)	
IGFBP3_c	F	184	25.79 ± 32.96, 10 (0, 160)	.5033
	M	186	23.44 ± 33.22, 10 (0, 180)	
IGFBP3_m	F	184	13.4 ± 37.42, 0 (0, 240)	.6277
	M	186	9.91 ± 24.24, 0 (0, 196.67)	
Insulin_c	F	182	44.45 ± 47.54, 25.83 (0, 180)	.9066
	M	185	43.78 ± 45.52, 30 (0, 186.67)	
Insulin_m	F	182	13.54 ± 24.19, 0 (0, 115)	.7634
	M	184	15.72 ± 32.07, 0 (0, 180)	
pAKT_c	F	183	4.16 ± 13.46, 0 (0, 83.33)	.3981
	M	186	3.2 ± 12.88, 0 (0, 86.67)	
pAKT_m	F	184	0.34 ± 2.62, 0 (0, 30)	.6791
	M	186	0.11 ± 0.91, 0 (0, 10)	
pAKT_n	F	184	1.16 ± 4.66, 0 (0, 40)	.7469
	M	186	1.05 ± 3.99, 0 (0, 30)	
pAMPK_c	F	184	44.56 ± 51.44, 25.83 (0, 200)	.0997
	M	186	36.42 ± 46.3, 10 (0, 180)	
pAMPK_m	F	184	3.24 ± 13.32, 0 (0, 110)	.5768
	M	186	1.43 ± 6.07, 0 (0, 60)	
pEGFR_c	F	183	22.64 ± 40.73, 0 (0, 190)	.0153
	M	185	15.26 ± 33.6, 0 (0, 166.67)	
pEGFR_m	F	183	18.86 ± 31.8, 3.33 (0, 160)	.0931
	M	186	14.31 ± 28.26, 0 (0, 143.33)	

<i>covariate</i>	<i>Gender</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
pEGFR_n	F	183	30.73 ± 37.44, 15 (0, 150)	.8814
	M	186	31.74 ± 40.25, 13.33 (0, 160)	
pIGF1R_c	F	183	5.93 ± 13.26, 0 (0, 90)	.4689
	M	185	5.73 ± 13.93, 0 (0, 120)	
pIGF1R_m	F	183	10.18 ± 28.01, 0 (0, 205)	.6528
	M	185	17.31 ± 47.45, 0 (0, 240)	
pS6_c	F	184	47.64 ± 42.23, 40 (0, 180)	.0840
	M	186	41.49 ± 42.64, 30 (0, 200)	
pS6_m	F	184	5.6 ± 15.68, 0 (0, 120)	.3488
	M	186	6.49 ± 24.08, 0 (0, 180)	
pSRC_c	F	184	3.35 ± 7.81, 0 (0, 40)	.8307
	M	185	2.75 ± 6.12, 0 (0, 33.33)	
pSRC_m	F	184	17.57 ± 36.39, 0 (0, 180)	.7911
	M	186	16.91 ± 34.24, 0 (0, 190)	
pmTOR_c	F	183	45.11 ± 54.91, 23.33 (0, 230)	.0014
	M	186	33.22 ± 48.15, 10 (0, 210)	
pmTOR_m	F	184	69.49 ± 69.36, 53.33 (0, 240)	.0019
	M	186	53.63 ± 67.48, 21.67 (0, 240)	

Table 10. Markers by Race

<i>covariate</i>	<i>Race</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
IGF1R_c	Black	21	30.56 ± 40.83, 10 (0, 120)	.2912
	Hispanic	14	20.48 ± 32.02, 0 (0, 96.67)	
	Oriental	5	18 ± 20.63, 10 (0, 53.33)	
	White	327	21.99 ± 36.24, 6.67 (0, 190)	
IGF1R_m	Black	21	4.6 ± 10.41, 0 (0, 40)	.0941
	Hispanic	14	0 ± 0, 0 (0, 0)	
	Oriental	5	6.67 ± 14.91, 0 (0, 33.33)	
	White	327	3.72 ± 14.75, 0 (0, 120)	
IGFBP3_c	Black	21	15.2 ± 30.61, 2.5 (0, 123.33)	.0670
	Hispanic	14	12.62 ± 16.35, 9.17 (0, 56.67)	

<i>covariate</i>	<i>Race</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
IGFBP3_m	Oriental	5	54 ± 47.98, 60 (0, 106.67)	.5420
	White	330	25.27 ± 33.25, 10 (0, 180)	
	Black	21	4.13 ± 9.54, 0 (0, 40)	
	Hispanic	14	1.9 ± 3.86, 0 (0, 13.33)	
	Oriental	5	5.33 ± 7.67, 0 (0, 16.67)	
Insulin_c	White	330	12.63 ± 33.12, 0 (0, 240)	.0757
	Black	21	37.7 ± 44.44, 25 (0, 186.67)	
	Hispanic	14	73.93 ± 52.38, 67.5 (10, 165)	
	Oriental	5	28.67 ± 30.97, 10 (0, 70)	
	White	327	43.49 ± 46.23, 30 (0, 180)	
Insulin_m	Black	21	19.76 ± 42.38, 0 (0, 180)	.3724
	Hispanic	14	4.17 ± 9.58, 0 (0, 35)	
	Oriental	5	2 ± 4.47, 0 (0, 10)	
	White	326	14.95 ± 27.98, 0 (0, 130)	
	Black	21	1.19 ± 2.75, 0 (0, 10)	.7927
pAKT_c	Hispanic	14	1.67 ± 4.29, 0 (0, 13.33)	
	Oriental	5	0 ± 0, 0 (0, 0)	
	White	329	3.98 ± 13.87, 0 (0, 86.67)	
	Black	21	0 ± 0, 0 (0, 0)	.8345
pAKT_m	Hispanic	14	0 ± 0, 0 (0, 0)	
	Oriental	5	0 ± 0, 0 (0, 0)	
	White	330	0.25 ± 2.07, 0 (0, 30)	
	Black	21	0.79 ± 2.56, 0 (0, 10)	.5099
pAKT_n	Hispanic	14	0 ± 0, 0 (0, 0)	
	Oriental	5	0 ± 0, 0 (0, 0)	
	White	330	1.19 ± 4.54, 0 (0, 40)	
	Black	21	43.57 ± 50.51, 20 (0, 160)	.7634
pAMPK_c	Hispanic	14	24.88 ± 27.27, 18.33 (0, 80)	
	Oriental	5	43.33 ± 40.69, 53.33 (0, 93.33)	
	White	330	40.88 ± 49.8, 16.67 (0, 200)	
	Black	21	3.49 ± 15.26, 0 (0, 70)	.7012
pAMPK_m	Hispanic	14	0.6 ± 1.55, 0 (0, 5)	

<i>covariate</i>	<i>Race</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
pEGFR_c	Oriental	5	0 ± 0, 0 (0, 0)	.3142
	White	330	2.37 ± 10.28, 0 (0, 110)	
	Black	20	17.08 ± 36.94, 0 (0, 150)	
	Hispanic	14	14.4 ± 32.37, 0 (0, 120)	
pEGFR_m	Oriental	5	0 ± 0, 0 (0, 0)	.0718
	White	329	19.52 ± 37.97, 0 (0, 190)	
	Black	21	8.33 ± 14.4, 0 (0, 55)	
	Hispanic	14	6.9 ± 12.62, 0 (0, 45)	
pEGFR_n	Oriental	5	0 ± 0, 0 (0, 0)	.8714
	White	329	17.76 ± 31.38, 3.33 (0, 160)	
	Black	21	40.08 ± 43.39, 26.67 (0, 136.67)	
	Hispanic	14	27.2 ± 34.92, 17.08 (0, 116.67)	
pIGF1R_c	Oriental	5	25.67 ± 24.08, 16.67 (0, 56.67)	.8670
	White	329	30.93 ± 38.93, 13.33 (0, 160)	
	Black	21	7.78 ± 14.54, 0 (0, 40)	
	Hispanic	14	2.62 ± 3.5, 0 (0, 10)	
pIGF1R_m	Oriental	5	1 ± 2.24, 0 (0, 5)	.8617
	White	328	5.91 ± 13.87, 0 (0, 120)	
	Black	21	16.03 ± 34.95, 0 (0, 143.33)	
	Hispanic	14	4.29 ± 8.91, 0 (0, 30)	
pS6_c	Oriental	5	1 ± 2.24, 0 (0, 5)	.9868
	White	328	14.22 ± 40.41, 0 (0, 240)	
	Black	21	48.13 ± 43.84, 46.67 (0, 155)	
	Hispanic	14	46.67 ± 49.65, 21.67 (0, 126.67)	
pS6_m	Oriental	5	56 ± 72.47, 10 (0, 143.33)	.3139
	White	330	44.06 ± 41.75, 33.33 (0, 200)	
	Black	21	6.9 ± 26.1, 0 (0, 120)	
	Hispanic	14	0 ± 0, 0 (0, 0)	
pSRC_c	Oriental	5	6.67 ± 14.91, 0 (0, 33.33)	.0899
	White	330	6.24 ± 20.43, 0 (0, 180)	
	Black	21	1.27 ± 3.07, 0 (0, 10)	
	Hispanic	14	0.24 ± 0.89, 0 (0, 3.33)	

<i>covariate</i>	<i>Race</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
pSRC_m	Oriental	5	0 ± 0, 0 (0, 0)	.3901
	White	329	3.33 ± 7.33, 0 (0, 40)	
	Black	21	12.7 ± 30.78, 0 (0, 115)	
	Hispanic	14	5.6 ± 11.43, 0 (0, 40)	
	Oriental	5	1 ± 2.24, 0 (0, 5)	
	White	330	18.27 ± 36.35, 0 (0, 190)	
	Black	21	28.33 ± 38.74, 10 (0, 150)	
	Hispanic	14	55.71 ± 65.77, 45 (0, 213.33)	
pmTOR_c	Oriental	5	46.67 ± 86.06, 10 (0, 200)	.8374
	White	329	38.98 ± 51.45, 16.67 (0, 230)	
	Black	21	68.81 ± 74.85, 50 (0, 230)	
	Hispanic	14	33.1 ± 31.52, 37.5 (0, 70)	
pmTOR_m	Oriental	5	28.67 ± 30.51, 13.33 (0, 63.33)	.4879
	White	330	62.76 ± 69.72, 36.67 (0, 240)	

Table 10.1. Markers by Smoking Status

<i>covariate</i>	<i>smoker</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
IGF1R_c	1 Never	38	5.53 ± 8.72, 0 (0, 40)	.0138
	2 Former	168	22.6 ± 35.34, 6.67 (0, 140)	
	3 Current	161	26.1 ± 39.81, 10 (0, 190)	
IGF1R_m	1 Never	38	0.18 ± 1.08, 0 (0, 6.67)	.1261
	2 Former	168	3.13 ± 11.43, 0 (0, 80)	
	3 Current	161	5.05 ± 17.96, 0 (0, 120)	
IGFBP3_c	1 Never	38	19.47 ± 25.54, 8.33 (0, 106.67)	.8088
	2 Former	170	22.73 ± 29.12, 10 (0, 123.33)	
	3 Current	162	27.78 ± 38.03, 10 (0, 180)	
IGFBP3_m	1 Never	38	2.85 ± 7.34, 0 (0, 40)	.0715
	2 Former	170	10.45 ± 30.26, 0 (0, 240)	
	3 Current	162	14.96 ± 35.64, 0 (0, 240)	
Insulin_c	1 Never	38	43.25 ± 47.48, 21.67 (0, 165)	.0876
	2 Former	169	48.13 ± 46.83, 40 (0, 186.67)	

<i>covariate</i>	<i>smoker</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
Insulin_m	3 Current	160	40.08 ± 45.78, 23.33 (0, 180)	.3680
	1 Never	38	8.64 ± 20.24, 0 (0, 100)	
	2 Former	168	14.98 ± 29.26, 0 (0, 180)	
pAKT_c	3 Current	160	15.71 ± 29.12, 0 (0, 130)	.2861
	1 Never	38	2.11 ± 10.76, 0 (0, 65)	
	2 Former	170	4.01 ± 14.28, 0 (0, 86.67)	
pAKT_m	3 Current	161	3.7 ± 12.47, 0 (0, 86.67)	.6018
	1 Never	38	0 ± 0, 0 (0, 0)	
	2 Former	170	0.31 ± 2.62, 0 (0, 30)	
pAKT_n	3 Current	162	0.19 ± 1.25, 0 (0, 10)	.6878
	1 Never	38	0.83 ± 2.84, 0 (0, 15)	
	2 Former	170	1.03 ± 4.37, 0 (0, 30)	
pAMPK_c	3 Current	162	1.25 ± 4.59, 0 (0, 40)	.1287
	1 Never	38	51.18 ± 52.5, 35.83 (0, 180)	
	2 Former	170	40.08 ± 47.48, 16.67 (0, 200)	
pAMPK_m	3 Current	162	38.35 ± 49.77, 10 (0, 186.67)	.7775
	1 Never	38	3.03 ± 11.87, 0 (0, 70)	
	2 Former	170	2.02 ± 10.55, 0 (0, 110)	
pEGFR_c	3 Current	162	2.5 ± 9.81, 0 (0, 70)	.5448
	1 Never	38	12.81 ± 23.43, 0 (0, 100)	
	2 Former	169	19.81 ± 37.89, 0 (0, 180)	
pEGFR_m	3 Current	161	19.45 ± 39.65, 0 (0, 190)	.0260
	1 Never	38	9.04 ± 21.12, 0 (0, 100)	
	2 Former	170	18.01 ± 30.88, 5 (0, 140)	
pEGFR_n	3 Current	161	16.82 ± 30.97, 0 (0, 160)	.6516
	1 Never	37	25 ± 32.4, 10 (0, 136.67)	
	2 Former	170	30.62 ± 39.48, 13.33 (0, 136.67)	
pIGF1R_c	3 Current	162	33.31 ± 39.53, 15 (0, 160)	.2733
	1 Never	38	2.98 ± 7.93, 0 (0, 40)	
	2 Former	169	5.3 ± 11.19, 0 (0, 65)	
pIGF1R_m	3 Current	161	7.06 ± 16.53, 0 (0, 120)	.5620
	1 Never	38	7.06 ± 25.81, 0 (0, 155)	

<i>covariate</i>	<i>smoker</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	2 Former	169	14.41 ± 42.3, 0 (0, 240)	
	3 Current	161	14.67 ± 38.29, 0 (0, 230)	
pS6_c	1 Never	38	48.2 ± 37.85, 43.33 (0, 126.67)	.3319
	2 Former	170	43.61 ± 45.7, 30 (0, 200)	
	3 Current	162	44.69 ± 40.15, 35.83 (0, 200)	
pS6_m	1 Never	38	2.41 ± 7.95, 0 (0, 45)	.3852
	2 Former	170	6.31 ± 22.66, 0 (0, 180)	
	3 Current	162	6.62 ± 19.72, 0 (0, 145)	
pSRC_c	1 Never	38	0.7 ± 3.3, 0 (0, 20)	.0068
	2 Former	169	3.14 ± 6.21, 0 (0, 33.33)	
	3 Current	162	3.51 ± 8.25, 0 (0, 40)	
pSRC_m	1 Never	38	7.06 ± 22.57, 0 (0, 120)	.0356
	2 Former	170	18.19 ± 35.15, 0 (0, 180)	
	3 Current	162	18.63 ± 37.57, 0 (0, 190)	
pmTOR_c	1 Never	38	83.66 ± 71.32, 70 (0, 230)	<.0001
	2 Former	169	32.58 ± 43.49, 13.33 (0, 210)	
	3 Current	162	35.48 ± 49.85, 13.33 (0, 230)	
pmTOR_m	1 Never	38	96.47 ± 69.99, 86.25 (0, 230)	.0003
	2 Former	170	62.72 ± 70.83, 36.67 (0, 240)	
	3 Current	162	52.06 ± 63.82, 23.33 (0, 230)	

Table 11. Markers by Histology

<i>covariate</i>	<i>histology0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
IGF1R_c	ADENO	226	8.49 ± 21.54, 0 (0, 190)	<.0001
	Other	17	20.2 ± 31.07, 5 (0, 95)	
	SCC	124	47.96 ± 43.56, 33.33 (0, 180)	
IGF1R_m	ADENO	226	0.84 ± 6.31, 0 (0, 76.67)	<.0001
	Other	17	0.78 ± 1.87, 0 (0, 6.67)	
	SCC	124	9.22 ± 21.99, 0 (0, 120)	
IGFBP3_c	ADENO	227	23.85 ± 30.45, 10 (0, 170)	.0099
	Other	17	6.47 ± 15.34, 0 (0, 63.33)	

<i>covariate</i>	<i>histology0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	SCC	126	28.41 ± 38.22, 10 (0, 180)	
IGFBP3_m	ADENO	227	6.72 ± 18.75, 0 (0, 196.67)	.0001
	Other	17	3.53 ± 14.55, 0 (0, 60)	
	SCC	126	21.61 ± 45.97, 0 (0, 240)	
Insulin_c	ADENO	226	50.83 ± 50.4, 33.33 (0, 186.67)	.0053
	Other	17	40 ± 39.02, 30 (0, 120)	
	SCC	124	32.45 ± 36.88, 20 (0, 170)	
Insulin_m	ADENO	225	16.98 ± 31.06, 0 (0, 180)	.1205
	Other	17	11.96 ± 22.49, 0 (0, 85)	
	SCC	124	10.77 ± 23.37, 0 (0, 130)	
pAKT_c	ADENO	226	3.13 ± 13.09, 0 (0, 86.67)	.0447
	Other	17	2.55 ± 7.02, 0 (0, 26.67)	
	SCC	126	4.81 ± 13.91, 0 (0, 86.67)	
pAKT_m	ADENO	227	0.09 ± 0.94, 0 (0, 10)	.1058
	Other	17	0 ± 0, 0 (0, 0)	
	SCC	126	0.5 ± 3.1, 0 (0, 30)	
pAKT_n	ADENO	227	0.82 ± 4.07, 0 (0, 40)	.1495
	Other	17	2.35 ± 7.52, 0 (0, 30)	
	SCC	126	1.44 ± 4.22, 0 (0, 23.33)	
pAMPK_c	ADENO	227	41.11 ± 51.47, 16.67 (0, 200)	.6716
	Other	17	46.08 ± 46.74, 33.33 (0, 140)	
	SCC	126	38.54 ± 44.9, 15.83 (0, 160)	
pAMPK_m	ADENO	227	3.1 ± 12.75, 0 (0, 110)	.5209
	Other	17	0.29 ± 1.21, 0 (0, 5)	
	SCC	126	1.23 ± 4.43, 0 (0, 43.33)	
pEGFR_c	ADENO	226	13.72 ± 31.78, 0 (0, 170)	.0002
	Other	17	17.16 ± 43.98, 0 (0, 180)	
	SCC	125	28.6 ± 43.86, 6.67 (0, 190)	
pEGFR_m	ADENO	226	12.12 ± 23.58, 0 (0, 140)	.0005
	Other	17	11.18 ± 24.06, 0 (0, 90)	
	SCC	126	25.28 ± 38.45, 6.67 (0, 160)	
pEGFR_n	ADENO	226	31.56 ± 39.55, 13.33 (0, 150)	.6955

<i>covariate</i>	<i>histology0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	Other	17	34.9 ± 34.9, 20 (0, 103.33)	
	SCC	126	30.17 ± 38.28, 14.17 (0, 160)	
pIGF1R_c	ADENO	226	4.56 ± 12.64, 0 (0, 120)	.0007
	Other	17	4.12 ± 10.64, 0 (0, 40)	
	SCC	125	8.35 ± 15.22, 0 (0, 90)	
pIGF1R_m	ADENO	226	9.49 ± 32.56, 0 (0, 240)	.0220
	Other	17	16.86 ± 46.91, 0 (0, 160)	
	SCC	125	21.07 ± 47.27, 0 (0, 240)	
pS6_c	ADENO	227	49.68 ± 42.89, 40 (0, 200)	.0007
	Other	17	43.73 ± 42.62, 35 (0, 143.33)	
	SCC	126	35.42 ± 40.49, 25 (0, 155)	
pS6_m	ADENO	227	6.12 ± 17.82, 0 (0, 150)	.1589
	Other	17	2.06 ± 7.3, 0 (0, 30)	
	SCC	126	6.44 ± 25.2, 0 (0, 180)	
pSRC_c	ADENO	226	1.63 ± 4.51, 0 (0, 30)	<.0001
	Other	17	1.86 ± 3.81, 0 (0, 13.33)	
	SCC	126	5.75 ± 9.74, 0 (0, 40)	
pSRC_m	ADENO	227	9.93 ± 24.41, 0 (0, 160)	<.0001
	Other	17	12.06 ± 25.34, 0 (0, 90)	
	SCC	126	31.11 ± 47.12, 6.67 (0, 190)	
pmTOR_c	ADENO	227	53.53 ± 55.98, 35 (0, 230)	<.0001
	Other	17	48.92 ± 69.33, 13.33 (0, 200)	
	SCC	125	11.61 ± 22.41, 3.33 (0, 110)	
pmTOR_m	ADENO	227	87.12 ± 69.99, 66.67 (0, 240)	<.0001
	Other	17	63.73 ± 74.59, 25 (0, 240)	
	SCC	126	15.09 ± 32.2, 0 (0, 180)	

Table 12. Marker by stage

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
IGF1R_c	IA	102	22.14 ± 39.2, 3.33 (0, 190)	.5720
	IB	130	20.54 ± 33.45, 6.67 (0, 140)	
	IIA	21	31.19 ± 44.61, 16.67 (0, 180)	
	IIB	53	23.9 ± 34.78, 10 (0, 130)	
	IIIA	61	22.28 ± 35.04, 3.33 (0, 150)	
IGF1R_m	IA	102	3.17 ± 13.61, 0 (0, 80)	.1912
	IB	130	2.13 ± 8.39, 0 (0, 56.67)	
	IIA	21	12.06 ± 34.87, 0 (0, 120)	
	IIB	53	4.21 ± 14.5, 0 (0, 70)	
	IIIA	61	4.43 ± 12.48, 0 (0, 66.67)	
IGFBP3_c	IA	103	20.19 ± 28.61, 6.67 (0, 120)	.1203
	IB	131	27.97 ± 35.66, 10 (0, 160)	
	IIA	22	17.2 ± 17.16, 10 (0, 56.67)	
	IIB	53	24.27 ± 42.09, 3.33 (0, 180)	
	IIIA	61	27.8 ± 29.15, 16.67 (0, 100)	
IGFBP3_m	IA	103	8.22 ± 28.37, 0 (0, 240)	.0463
	IB	131	16.64 ± 40.99, 0 (0, 240)	
	IIA	22	7.12 ± 13.27, 1.67 (0, 60)	
	IIB	53	8.81 ± 26.74, 0 (0, 163.33)	
	IIIA	61	10.79 ± 17.99, 0 (0, 60)	
Insulin_c	IA	100	46.23 ± 48.13, 30 (0, 180)	.2914
	IB	131	45.2 ± 47.1, 30 (0, 186.67)	
	IIA	22	62.05 ± 59.44, 38.33 (0, 180)	
	IIB	53	33.18 ± 36.53, 23.33 (0, 150)	
	IIIA	61	41.34 ± 43.72, 26.67 (0, 180)	
Insulin_m	IA	99	13.28 ± 24.12, 0 (0, 110)	.6856
	IB	131	14.16 ± 27.75, 0 (0, 130)	
	IIA	22	11.44 ± 17.55, 1.67 (0, 60)	
	IIB	53	15.85 ± 36.8, 0 (0, 180)	
	IIIA	61	17.98 ± 31.44, 0 (0, 120)	
pAKT_c	IA	102	3.12 ± 12.3, 0 (0, 83.33)	.0324

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	IB	131	5.23 ± 15.6, 0 (0, 86.67)	
	IIA	22	0 ± 0, 0 (0, 0)	
	IIB	53	2.52 ± 13.2, 0 (0, 86.67)	
	IIIA	61	3.61 ± 10.86, 0 (0, 65)	
	IIIA	61	3.61 ± 10.86, 0 (0, 65)	
pAKT_m	IA	103	0.16 ± 1.35, 0 (0, 13.33)	.6326
	IB	131	0.46 ± 3, 0 (0, 30)	
	IIA	22	0 ± 0, 0 (0, 0)	
	IIB	53	0.13 ± 0.92, 0 (0, 6.67)	
	IIIA	61	0 ± 0, 0 (0, 0)	
pAKT_n	IA	103	0.7 ± 2.51, 0 (0, 16.67)	.5173
	IB	131	1.5 ± 5.69, 0 (0, 40)	
	IIA	22	0 ± 0, 0 (0, 0)	
	IIB	53	1.07 ± 3.74, 0 (0, 20)	
	IIIA	61	1.37 ± 4.62, 0 (0, 26.67)	
pAMPK_c	IA	103	46.57 ± 56.15, 16.67 (0, 200)	.5691
	IB	131	41.49 ± 46.37, 30 (0, 186.67)	
	IIA	22	41.59 ± 49.14, 26.67 (0, 160)	
	IIB	53	33.08 ± 42.75, 10 (0, 145)	
	IIIA	61	33.96 ± 46.64, 6.67 (0, 180)	
pAMPK_m	IA	103	3.75 ± 15.91, 0 (0, 110)	.6270
	IB	131	2.19 ± 7.95, 0 (0, 63.33)	
	IIA	22	0.3 ± 0.98, 0 (0, 3.33)	
	IIB	53	1.7 ± 5.18, 0 (0, 25)	
	IIIA	61	1.53 ± 7.93, 0 (0, 60)	
pEGFR_c	IA	103	20.79 ± 41.86, 0 (0, 190)	.9401
	IB	131	20.33 ± 38.81, 0 (0, 180)	
	IIA	22	19.7 ± 39.46, 0 (0, 150)	
	IIB	52	16.51 ± 34.21, 0 (0, 130)	
	IIIA	60	14.5 ± 27.79, 0 (0, 120)	
pEGFR_m	IA	103	16.49 ± 30.59, 3.33 (0, 160)	.4141
	IB	131	18.38 ± 32, 3.33 (0, 140)	
	IIA	22	9.77 ± 24.46, 0 (0, 100)	

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
pEGFR_n	IIB	52	15.1 ± 30.33, 3.33 (0, 143.33)	.2147
	IIIA	61	16.5 ± 27.1, 6.67 (0, 130)	
	IA	103	28.29 ± 38.02, 10 (0, 150)	
	IB	130	31.79 ± 37.67, 15 (0, 160)	
	IIA	22	20.08 ± 28.53, 3.33 (0, 90)	
	IIB	53	39.97 ± 43.66, 25 (0, 136.67)	
pIGF1R_c	IIIA	61	31.46 ± 40.92, 15 (0, 136.67)	.8486
	IA	101	5.43 ± 12.33, 0 (0, 65)	
	IB	131	7.23 ± 17.01, 0 (0, 120)	
	IIA	22	5.91 ± 9.42, 0 (0, 33.33)	
	IIB	53	4.59 ± 9.99, 0 (0, 43.33)	
	IIIA	61	4.54 ± 11.07, 0 (0, 65)	
pIGF1R_m	IA	101	10.38 ± 35.24, 0 (0, 240)	.4719
	IB	131	14.95 ± 41.21, 0 (0, 230)	
	IIA	22	28.03 ± 56.55, 0 (0, 210)	
	IIB	53	17.83 ± 47, 0 (0, 240)	
	IIIA	61	8.14 ± 20.91, 0 (0, 120)	
	IA	103	46.59 ± 47.78, 25 (0, 200)	.4861
pS6_c	IB	131	44.26 ± 38.7, 36.67 (0, 143.33)	
	IIA	22	29.47 ± 34.05, 18.33 (0, 126.67)	
	IIB	53	48.49 ± 49.01, 35 (0, 200)	
	IIIA	61	43.77 ± 37.22, 40 (0, 153.33)	
	IA	103	5.79 ± 22.76, 0 (0, 150)	.4219
	IB	131	5.7 ± 20.27, 0 (0, 180)	
pS6_m	IIA	22	1.52 ± 5.32, 0 (0, 23.33)	
	IIB	53	5.88 ± 14.06, 0 (0, 60)	
	IIIA	61	8.99 ± 23.87, 0 (0, 120)	
	IA	103	2.72 ± 6.48, 0 (0, 33.33)	.5796
	IB	130	3.63 ± 7.69, 0 (0, 40)	
	IIA	22	3.33 ± 9.32, 0 (0, 40)	
pSRC_c	IIB	53	2.96 ± 6.94, 0 (0, 30)	
	IIIA	61	2.35 ± 5.41, 0 (0, 30)	

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
pSRC_m	IA	103	15.08 ± 30.14, 0 (0, 145)	.9180
	IB	131	20.8 ± 42.13, 0 (0, 190)	
	IIA	22	18.64 ± 41.45, 0 (0, 180)	
	IIB	53	15.19 ± 31.71, 0 (0, 160)	
	IIIA	61	14.51 ± 27.27, 0 (0, 130)	
pmTOR_c	IA	103	40.73 ± 47.98, 30 (0, 210)	.1905
	IB	131	40.35 ± 53.77, 20 (0, 230)	
	IIA	22	46.29 ± 70.54, 7.5 (0, 230)	
	IIB	52	25.96 ± 37.76, 6.67 (0, 160)	
	IIIA	61	42.38 ± 56.5, 15 (0, 200)	
pmTOR_m	IA	103	68.54 ± 71.87, 55 (0, 240)	.1051
	IB	131	63.57 ± 69.76, 40 (0, 240)	
	IIA	22	54.62 ± 73.73, 5 (0, 226.67)	
	IIB	53	47.11 ± 62.01, 13.33 (0, 230)	
	IIIA	61	60.25 ± 65.16, 35 (0, 230)	

<i>covariate</i>	<i>stage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
IGF1R_c	I	232	21.24 ± 36.02, 5 (0, 190)	.3655
	II	74	25.97 ± 37.65, 10 (0, 180)	
	IIIA	61	22.28 ± 35.04, 3.33 (0, 150)	
IGF1R_m	I	232	2.59 ± 10.98, 0 (0, 80)	.1064
	II	74	6.44 ± 22.26, 0 (0, 120)	
	IIIA	61	4.43 ± 12.48, 0 (0, 66.67)	
IGFBP3_c	I	234	24.54 ± 32.91, 10 (0, 160)	.1301
	II	75	22.2 ± 36.59, 6.67 (0, 180)	
	IIIA	61	27.8 ± 29.15, 16.67 (0, 100)	
IGFBP3_m	I	234	12.93 ± 36.16, 0 (0, 240)	.2910
	II	75	8.31 ± 23.52, 0 (0, 163.33)	
	IIIA	61	10.79 ± 17.99, 0 (0, 60)	
Insulin_c	I	231	45.65 ± 47.45, 30 (0, 186.67)	.6956

<i>covariate</i>	<i>stage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	II	75	41.64 ± 45.99, 26.67 (0, 180)	
	IIIA	61	41.34 ± 43.72, 26.67 (0, 180)	
	IIIB	61	41.34 ± 43.72, 26.67 (0, 180)	
Insulin_m	I	230	13.78 ± 26.2, 0 (0, 130)	.5161
	II	75	14.56 ± 32.3, 0 (0, 180)	
	IIIA	61	17.98 ± 31.44, 0 (0, 120)	
pAKT_c	I	233	4.31 ± 14.26, 0 (0, 86.67)	.0187
	II	75	1.78 ± 11.13, 0 (0, 86.67)	
	IIIA	61	3.61 ± 10.86, 0 (0, 65)	
pAKT_m	I	234	0.33 ± 2.42, 0 (0, 30)	.3904
	II	75	0.09 ± 0.77, 0 (0, 6.67)	
	IIIA	61	0 ± 0, 0 (0, 0)	
pAKT_n	I	234	1.15 ± 4.58, 0 (0, 40)	.5370
	II	75	0.76 ± 3.18, 0 (0, 20)	
	IIIA	61	1.37 ± 4.62, 0 (0, 26.67)	
pAMPK_c	I	234	43.73 ± 50.85, 20 (0, 200)	.2845
	II	75	35.58 ± 44.55, 13.33 (0, 160)	
	IIIA	61	33.96 ± 46.64, 6.67 (0, 180)	
pAMPK_m	I	234	2.88 ± 12.11, 0 (0, 110)	.4512
	II	75	1.29 ± 4.42, 0 (0, 25)	
	IIIA	61	1.53 ± 7.93, 0 (0, 60)	
pEGFR_c	I	234	20.53 ± 40.1, 0 (0, 190)	.6924
	II	74	17.45 ± 35.6, 0 (0, 150)	
	IIIA	60	14.5 ± 27.79, 0 (0, 120)	
pEGFR_m	I	234	17.55 ± 31.34, 3.33 (0, 160)	.4237
	II	74	13.51 ± 28.65, 0 (0, 143.33)	
	IIIA	61	16.5 ± 27.1, 6.67 (0, 130)	
pEGFR_n	I	233	30.25 ± 37.79, 13.33 (0, 160)	.7972
	II	75	34.13 ± 40.67, 15 (0, 136.67)	
	IIIA	61	31.46 ± 40.92, 15 (0, 136.67)	
pIGF1R_c	I	232	6.44 ± 15.15, 0 (0, 120)	.8466
	II	75	4.98 ± 9.78, 0 (0, 43.33)	
	IIIA	61	4.54 ± 11.07, 0 (0, 65)	

<i>covariate</i>	<i>stage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
pIGF1R_m	I	232	12.96 ± 38.71, 0 (0, 240)	.7211
	II	75	20.82 ± 49.81, 0 (0, 240)	
	IIIA	61	8.14 ± 20.91, 0 (0, 120)	
pS6_c	I	234	45.28 ± 42.85, 33.33 (0, 200)	.6571
	II	75	42.91 ± 45.74, 30 (0, 200)	
	IIIA	61	43.77 ± 37.22, 40 (0, 153.33)	
pS6_m	I	234	5.74 ± 21.35, 0 (0, 180)	.4846
	II	75	4.6 ± 12.29, 0 (0, 60)	
	IIIA	61	8.99 ± 23.87, 0 (0, 120)	
pSRC_c	I	233	3.23 ± 7.18, 0 (0, 40)	.6189
	II	75	3.07 ± 7.65, 0 (0, 40)	
	IIIA	61	2.35 ± 5.41, 0 (0, 30)	
pSRC_m	I	234	18.28 ± 37.36, 0 (0, 190)	.7722
	II	75	16.2 ± 34.6, 0 (0, 180)	
	IIIA	61	14.51 ± 27.27, 0 (0, 130)	
pmTOR_c	I	234	40.52 ± 51.2, 21.67 (0, 230)	.0625
	II	74	32 ± 50.15, 6.67 (0, 230)	
	IIIA	61	42.38 ± 56.5, 15 (0, 200)	
pmTOR_m	I	234	65.76 ± 70.59, 45 (0, 240)	.0269
	II	75	49.31 ± 65.24, 10 (0, 230)	
	IIIA	61	60.25 ± 65.16, 35 (0, 230)	

Table 13. Marker by Grade

<i>covariate</i>	<i>grade0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
IGF1R_c	1 Poorly	122	19.54 ± 33.95, 5.83 (0, 150)	.0002
	2 Moderately	196	26.47 ± 37.61, 10 (0, 180)	
	3 Well	36	9.12 ± 32.83, 0 (0, 190)	
IGF1R_m	1 Poorly	122	2.62 ± 9.26, 0 (0, 60)	.1748
	2 Moderately	196	4.8 ± 17.18, 0 (0, 120)	
	3 Well	36	2.13 ± 12.78, 0 (0, 76.67)	
IGFBP3_c	1 Poorly	122	26.17 ± 34.9, 10 (0, 170)	.8743

<i>covariate</i>	<i>grade0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	2 Moderately	199	25.15 ± 33.2, 10 (0, 180)	
	3 Well	36	22.41 ± 29.26, 6.67 (0, 116.67)	
IGFBP3_m	1 Poorly	122	16.08 ± 43.48, 0 (0, 240)	.3839
	2 Moderately	199	10.44 ± 24.95, 0 (0, 193.33)	
	3 Well	36	5.83 ± 13.41, 0 (0, 60)	
Insulin_c	1 Poorly	121	47.19 ± 47.34, 33.33 (0, 180)	.6229
	2 Moderately	197	43.17 ± 47.77, 25 (0, 186.67)	
	3 Well	36	39.4 ± 37.58, 28.33 (0, 123.33)	
Insulin_m	1 Poorly	121	21.35 ± 32.89, 0 (0, 120)	.0111
	2 Moderately	196	12.67 ± 27.3, 0 (0, 180)	
	3 Well	36	3.43 ± 6.51, 0 (0, 30)	
pAKT_c	1 Poorly	121	3.15 ± 10.32, 0 (0, 65)	.6710
	2 Moderately	199	3.98 ± 14.34, 0 (0, 86.67)	
	3 Well	36	4.26 ± 16.51, 0 (0, 83.33)	
pAKT_m	1 Poorly	122	0.11 ± 1.21, 0 (0, 13.33)	.2640
	2 Moderately	199	0.35 ± 2.49, 0 (0, 30)	
	3 Well	36	0 ± 0, 0 (0, 0)	
pAKT_n	1 Poorly	122	0.75 ± 3.25, 0 (0, 26.67)	.5510
	2 Moderately	199	1.31 ± 4.82, 0 (0, 40)	
	3 Well	36	0.46 ± 1.62, 0 (0, 6.67)	
pAMPK_c	1 Poorly	122	47.61 ± 53.83, 20 (0, 186.67)	.1009
	2 Moderately	199	34.9 ± 45.52, 10 (0, 200)	
	3 Well	36	48.1 ± 50.69, 35 (0, 175)	
pAMPK_m	1 Poorly	122	3.5 ± 12.11, 0 (0, 70)	.1176
	2 Moderately	199	1.95 ± 10.24, 0 (0, 110)	
	3 Well	36	1.34 ± 4.69, 0 (0, 20)	
pEGFR_c	1 Poorly	121	22.62 ± 42.61, 0 (0, 170)	.0983
	2 Moderately	198	18.62 ± 35.6, 0 (0, 190)	
	3 Well	36	7.82 ± 17.72, 0 (0, 70)	
pEGFR_m	1 Poorly	121	16.83 ± 29.5, 3.33 (0, 140)	.0139
	2 Moderately	199	18.24 ± 31.64, 3.33 (0, 160)	
	3 Well	36	7.78 ± 23.16, 0 (0, 120)	

<i>covariate</i>	<i>grade0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
pEGFR_n	1 Poorly	122	38.74 ± 41.99, 20 (0, 160)	.0144
	2 Moderately	199	27.37 ± 37.14, 10 (0, 150)	
	3 Well	35	28.69 ± 37.57, 6.67 (0, 123.33)	
pIGF1R_c	1 Poorly	121	6.74 ± 15.96, 0 (0, 120)	.1089
	2 Moderately	198	6.18 ± 13.17, 0 (0, 90)	
	3 Well	36	1.57 ± 4.1, 0 (0, 20)	
pIGF1R_m	1 Poorly	121	14.02 ± 38.38, 0 (0, 240)	.4286
	2 Moderately	198	15.76 ± 43.04, 0 (0, 240)	
	3 Well	36	3.38 ± 7.69, 0 (0, 30)	
pS6_c	1 Poorly	122	44 ± 42.96, 34.17 (0, 200)	.0795
	2 Moderately	199	42.44 ± 41.63, 30 (0, 155)	
	3 Well	36	56.39 ± 43.22, 50 (3.33, 180)	
pS6_m	1 Poorly	122	6.43 ± 18.65, 0 (0, 120)	.3905
	2 Moderately	199	6.62 ± 23.08, 0 (0, 180)	
	3 Well	36	2.5 ± 9.61, 0 (0, 53.33)	
pSRC_c	1 Poorly	122	3.28 ± 7.02, 0 (0, 40)	.0360
	2 Moderately	198	3.28 ± 7.31, 0 (0, 40)	
	3 Well	36	1.57 ± 6.25, 0 (0, 30)	
pSRC_m	1 Poorly	122	20.18 ± 37.08, 0 (0, 180)	.0002
	2 Moderately	199	18.65 ± 37.05, 0 (0, 190)	
	3 Well	36	1.85 ± 8.49, 0 (0, 50)	
pmTOR_c	1 Poorly	121	31.07 ± 39.04, 13.33 (0, 170)	<.0001
	2 Moderately	199	34.62 ± 48.23, 10 (0, 230)	
	3 Well	36	83.89 ± 73.54, 60 (0, 230)	
pmTOR_m	1 Poorly	122	56.5 ± 65.85, 28.33 (0, 240)	.0043
	2 Moderately	199	57.97 ± 68.48, 30 (0, 240)	
	3 Well	36	93.33 ± 70.37, 67.5 (0, 230)	

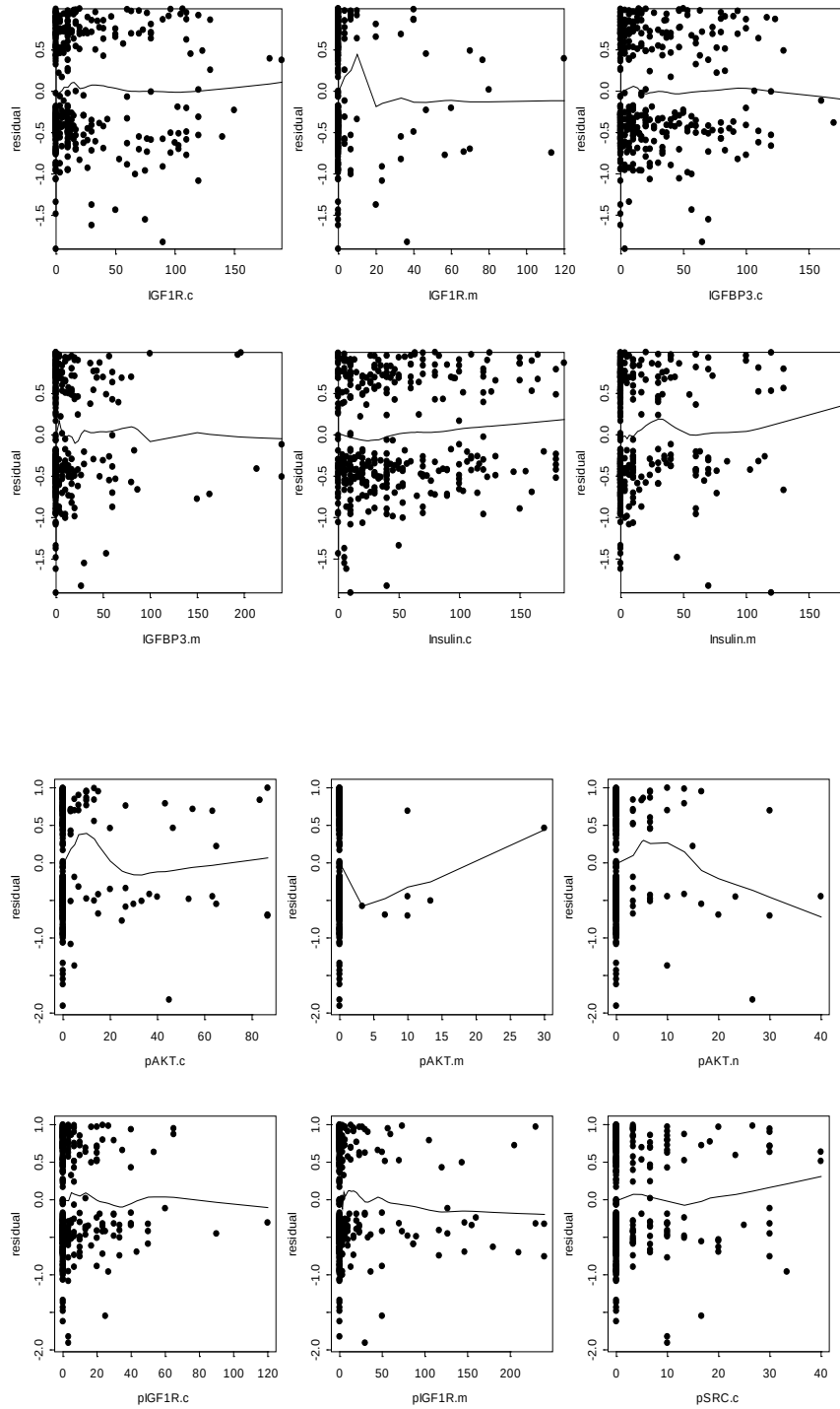
2.2 Overall survival

Table 14. Univariate Cox model assessing effect of covariates on overall survival

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
IGF1R_c	0.0026	0.0020	1.0026	0.9987	1.0066	0.1886	367	157	210
IGF1R_c_01	0.3260	0.1663	1.3855	1.0002	1.9192	0.0499	367	157	210
IGF1R_c_01m	0.2856	0.1613	1.3306	0.9699	1.8253	0.0766	367	157	210
IGF1R_m	0.0019	0.0049	1.0019	0.9923	1.0115	0.7007	367	157	210
IGF1R_m_01	0.3614	0.2149	1.4353	0.9419	2.1871	0.0926	367	157	210
IGFBP3_c	0.0003	0.0024	1.0003	0.9955	1.0050	0.9082	370	160	210
IGFBP3_c_01	0.1027	0.1729	1.1081	0.7897	1.5550	0.5525	370	160	210
IGFBP3_c_01m	0.0588	0.1589	1.0605	0.7767	1.4480	0.7115	370	160	210
IGFBP3_m	0.0005	0.0026	1.0005	0.9954	1.0057	0.8347	370	160	210
IGFBP3_m_01	0.1080	0.1648	1.1140	0.8065	1.5388	0.5122	370	160	210
Insulin_c	0.0016	0.0017	1.0016	0.9983	1.0049	0.3366	367	160	207
Insulin_c_01	-0.3352	0.2115	0.7152	0.4725	1.0826	0.1130	367	160	207
Insulin_c_01m	0.1229	0.1586	1.1307	0.8286	1.5430	0.4386	367	160	207
Insulin_m	0.0044	0.0026	1.0044	0.9993	1.0096	0.0915	366	160	206
Insulin_m_01	0.2329	0.1588	1.2622	0.9246	1.7231	0.1426	366	160	206
pAKT_c	-0.0008	0.0059	0.9992	0.9877	1.0108	0.8865	369	159	210
pAKT_c_01	0.2679	0.2101	1.3073	0.8659	1.9735	0.2023	369	159	210
pAKT_m	-0.0164	0.0420	0.9837	0.9059	1.0682	0.6958	370	160	210
pAKT_m_01	-0.8438	0.7144	0.4301	0.1060	1.7444	0.2375	370	160	210
pAKT_n	-0.0070	0.0169	0.9931	0.9607	1.0264	0.6794	370	160	210
pAKT_n_01	0.2613	0.2354	1.2986	0.8187	2.0598	0.2669	370	160	210
pAMPK_c	-0.0032	0.0017	0.9968	0.9934	1.0002	0.0627	370	160	210
pAMPK_c_01	-0.3682	0.1659	0.6920	0.4999	0.9579	0.0265	370	160	210
pAMPK_m	0.0003	0.0077	1.0003	0.9853	1.0155	0.9732	370	160	210
pAMPK_m_01	-0.1535	0.2297	0.8577	0.5468	1.3456	0.5042	370	160	210
pEGFR_c	-0.0003	0.0023	0.9997	0.9953	1.0041	0.8865	368	158	210
pEGFR_c_01	-0.0092	0.1609	0.9908	0.7228	1.3582	0.9543	368	158	210
pEGFR_m	-0.0003	0.0027	0.9997	0.9944	1.0049	0.8967	369	159	210
pEGFR_m_01	-0.0165	0.1589	0.9836	0.7204	1.3430	0.9172	369	159	210

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
pEGFR_m_01m	-0.0165	0.1589	0.9836	0.7204	1.3430	0.9172	369	159	210
pEGFR_n	0.0002	0.0021	1.0002	0.9961	1.0042	0.9394	369	160	209
pEGFR_n_01	-0.1673	0.1700	0.8460	0.6062	1.1805	0.3252	369	160	209
pEGFR_n_01m	-0.0199	0.1586	0.9803	0.7184	1.3377	0.9001	369	160	209
plGF1R_c	-0.0027	0.0064	0.9973	0.9848	1.0099	0.6690	368	160	208
plGF1R_c_01	0.0946	0.1681	1.0993	0.7906	1.5284	0.5735	368	160	208
plGF1R_m	-0.0020	0.0024	0.9980	0.9933	1.0027	0.4052	368	160	208
plGF1R_m_01	-0.0056	0.1752	0.9944	0.7055	1.4017	0.9744	368	160	208
pS6_c	-0.0003	0.0018	0.9997	0.9961	1.0033	0.8723	370	160	210
pS6_c_01	0.0114	0.2257	1.0115	0.6499	1.5742	0.9598	370	160	210
pS6_c_01m	-0.0543	0.1583	0.9472	0.6944	1.2918	0.7317	370	160	210
pS6_m	0.0020	0.0034	1.0020	0.9954	1.0086	0.5596	370	160	210
pS6_m_01	0.0711	0.1924	1.0737	0.7364	1.5654	0.7118	370	160	210
pSRC_c	0.0121	0.0102	1.0121	0.9922	1.0325	0.2356	369	160	209
pSRC_c_01	0.2116	0.1717	1.2357	0.8825	1.7301	0.2179	369	160	209
pSRC_m	0.0009	0.0022	1.0009	0.9966	1.0053	0.6719	370	160	210
pSRC_m_01	0.1656	0.1610	1.1801	0.8608	1.6179	0.3035	370	160	210
pmTOR_c	-0.0020	0.0017	0.9980	0.9946	1.0014	0.2392	369	160	209
pmTOR_c_01	-0.4985	0.1708	0.6075	0.4347	0.8489	0.0035	369	160	209
pmTOR_c_01m	-0.3043	0.1589	0.7376	0.5402	1.0071	0.0554	369	160	209
pmTOR_m	-0.0016	0.0012	0.9984	0.9960	1.0007	0.1678	370	160	210
pmTOR_m_01	-0.1991	0.1798	0.8195	0.5761	1.1657	0.2682	370	160	210
pmTOR_m_01m	-0.2709	0.1584	0.7627	0.5591	1.0405	0.0873	370	160	210

Figure 3. Martingale residual from Cox model with age, gender, histology and stage for overall survival against each marker, Dr. HY Lee



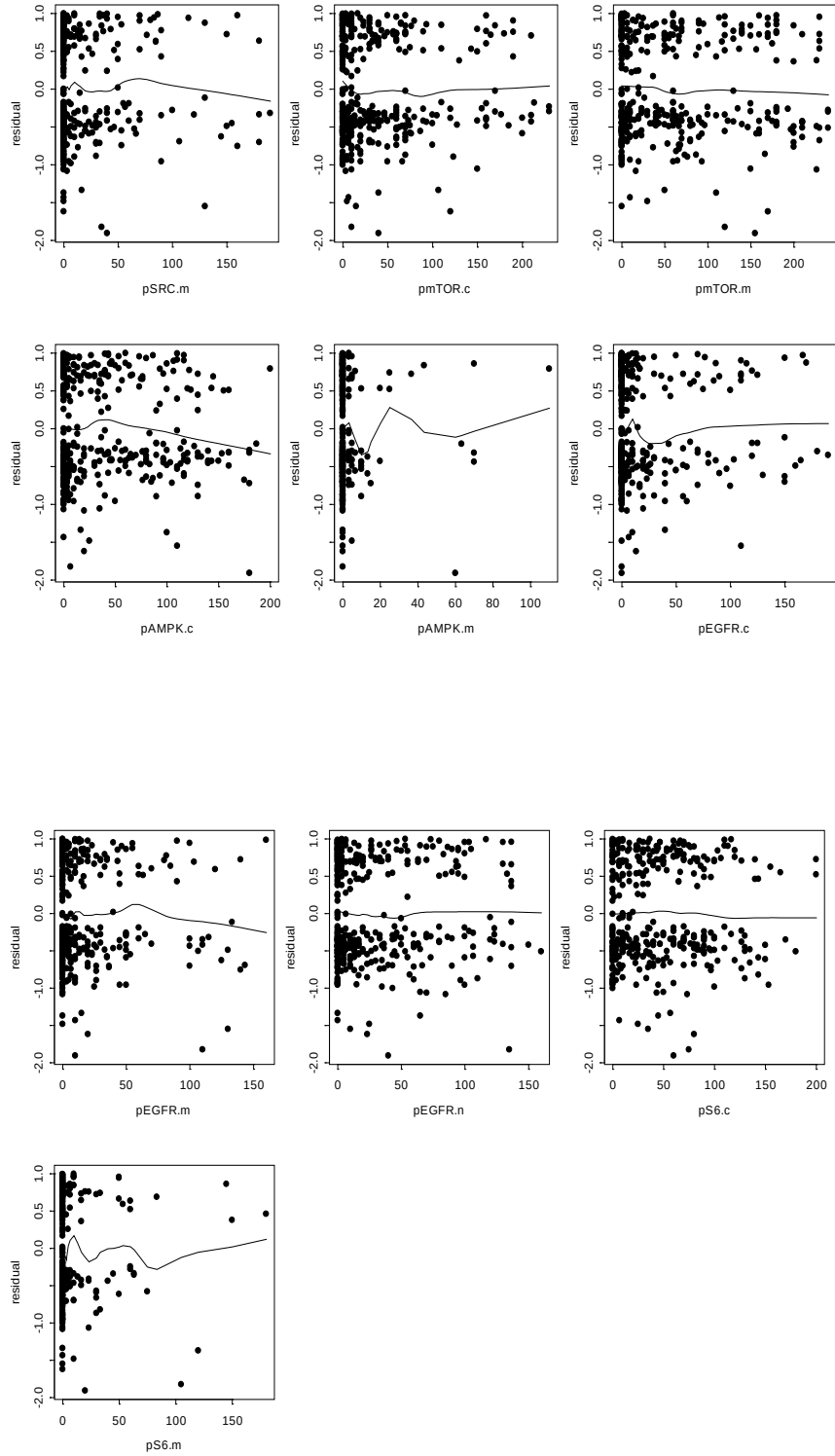


Table 15. Multicovariate Cox model assessing effect of each marker independently on overall survival, adjusting for age, gender, histology, stage and neoadjuvant treatment

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
IGF1R_c	0.0015	0.0024	1.0015	0.9968	1.0062	0.5275	364	155	209
IGF1R_c_01	0.2619	0.1866	1.2994	0.9014	1.8732	0.1604	364	155	209
IGF1R_c_01m	0.3100	0.1842	1.3635	0.9502	1.9565	0.0924	364	155	209
IGF1R_m	-0.0026	0.0054	0.9974	0.9869	1.0080	0.6295	364	155	209
IGF1R_m_01	0.1450	0.2410	1.1560	0.7208	1.8541	0.5474	364	155	209
IGFBP3_c	-0.0009	0.0025	0.9991	0.9941	1.0041	0.7175	367	158	209
IGFBP3_c_01	0.0452	0.1782	1.0463	0.7378	1.4837	0.7997	367	158	209
IGFBP3_c_01m	-0.0334	0.1652	0.9671	0.6996	1.3370	0.8396	367	158	209
IGFBP3_m	-0.0003	0.0029	0.9997	0.9941	1.0054	0.9296	367	158	209
IGFBP3_m_01	0.0030	0.1730	1.0030	0.7146	1.4078	0.9864	367	158	209
Insulin_c	0.0028	0.0018	1.0028	0.9994	1.0063	0.1083	364	158	206
Insulin_c_01	-0.3243	0.2173	0.7231	0.4723	1.1070	0.1357	364	158	206
Insulin_c_01m	0.1440	0.1638	1.1549	0.8377	1.5921	0.3794	364	158	206
Insulin_m	0.0038	0.0027	1.0038	0.9985	1.0090	0.1579	363	158	205
Insulin_m_01	0.2356	0.1623	1.2657	0.9208	1.7397	0.1466	363	158	205
pAKT_c	-0.0006	0.0060	0.9994	0.9878	1.0111	0.9200	366	157	209
pAKT_c_01	0.3100	0.2203	1.3634	0.8853	2.0996	0.1594	366	157	209
pAKT_m	-0.0037	0.0390	0.9963	0.9230	1.0753	0.9233	367	158	209
pAKT_m_01	-0.6073	0.7190	0.5448	0.1331	2.2299	0.3983	367	158	209
pAKT_n	-0.0135	0.0181	0.9866	0.9522	1.0222	0.4558	367	158	209
pAKT_n_01	0.2472	0.2400	1.2804	0.7999	2.0495	0.3031	367	158	209
pAMPK_c	-0.0028	0.0017	0.9972	0.9938	1.0006	0.1069	367	158	209
pAMPK_c_01	-0.4263	0.1694	0.6529	0.4684	0.9101	0.0119	367	158	209
pAMPK_m	0.0015	0.0077	1.0015	0.9864	1.0168	0.8453	367	158	209
pAMPK_m_01	-0.1301	0.2311	0.8780	0.5582	1.3811	0.5735	367	158	209
pEGFR_c	-0.0002	0.0024	0.9998	0.9952	1.0045	0.9482	365	156	209
pEGFR_c_01	-0.0467	0.1704	0.9544	0.6835	1.3328	0.7841	365	156	209
pEGFR_m	-0.0022	0.0028	0.9978	0.9924	1.0032	0.4280	366	157	209
pEGFR_m_01	-0.1131	0.1625	0.8930	0.6495	1.2279	0.4862	366	157	209
pEGFR_m_01m	-0.1131	0.1625	0.8930	0.6495	1.2279	0.4862	366	157	209

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
pEGFR_n	-0.0002	0.0021	0.9998	0.9957	1.0040	0.9288	366	158	208
pEGFR_n_01	-0.0957	0.1714	0.9088	0.6495	1.2715	0.5766	366	158	208
pEGFR_n_01m	-0.0261	0.1613	0.9743	0.7102	1.3365	0.8716	366	158	208
plGF1R_c	-0.0034	0.0068	0.9966	0.9835	1.0100	0.6197	365	158	207
plGF1R_c_01	0.0132	0.1756	1.0133	0.7183	1.4295	0.9400	365	158	207
plGF1R_m	-0.0027	0.0026	0.9973	0.9923	1.0023	0.2865	365	158	207
plGF1R_m_01	-0.0459	0.1810	0.9551	0.6699	1.3618	0.7997	365	158	207
pS6_c	-0.0005	0.0019	0.9995	0.9958	1.0032	0.7888	367	158	209
pS6_c_01	0.0547	0.2376	1.0563	0.6631	1.6826	0.8178	367	158	209
pS6_c_01m	-0.0825	0.1641	0.9208	0.6676	1.2701	0.6151	367	158	209
pS6_m	-0.0007	0.0035	0.9993	0.9925	1.0061	0.8331	367	158	209
pS6_m_01	0.0074	0.2004	1.0075	0.6803	1.4920	0.9704	367	158	209
pSRC_c	0.0099	0.0113	1.0100	0.9879	1.0325	0.3778	366	158	208
pSRC_c_01	0.1417	0.1848	1.1522	0.8021	1.6551	0.4433	366	158	208
pSRC_m	-0.0001	0.0024	0.9999	0.9952	1.0046	0.9526	367	158	209
pSRC_m_01	0.0665	0.1680	1.0688	0.7689	1.4856	0.6921	367	158	209
pmTOR_c	-0.0009	0.0019	0.9991	0.9954	1.0028	0.6315	366	158	208
pmTOR_c_01	-0.5680	0.1997	0.5667	0.3831	0.8381	0.0045	366	158	208
pmTOR_c_01m	-0.1809	0.1783	0.8345	0.5884	1.1837	0.3104	366	158	208
pmTOR_m	-0.0013	0.0014	0.9987	0.9959	1.0015	0.3617	367	158	209
pmTOR_m_01	-0.1737	0.2217	0.8406	0.5443	1.2981	0.4335	367	158	209
pmTOR_m_01m	-0.2320	0.1904	0.7929	0.5460	1.1516	0.2230	367	158	209

Final multivariate Cox Model assessing the following covariates on overall survival

Analysis of Maximum Likelihood Estimates						
Variable	Parameter Estimate	Standard Error	p-value	Hazard Ratio	95% Hazard Ratio Confidence Limits	
Age	0.0317	0.0090	0.0004	1.032	1.014	1.051
Gender M vs F	0.2180	0.1648	0.1859	1.244	0.900	1.718
Stage II vs I	0.3711	0.2074	0.0736	1.449	0.965	2.176
III vs I	0.9189	0.2146	<.0001	2.507	1.646	3.817
Neoadjuvant (Yes vs No)	0.4398	0.2113	0.0374	1.552	1.026	2.349
pAMPK_c_01 (Pos vs 0)	-0.3884	0.1692	0.0217	0.678	0.487	0.945
pmTOR_c_01 (Pos vs 0)	-0.4720	0.1794	0.0085	0.624	0.439	0.887

2.3 Recurrence free survival

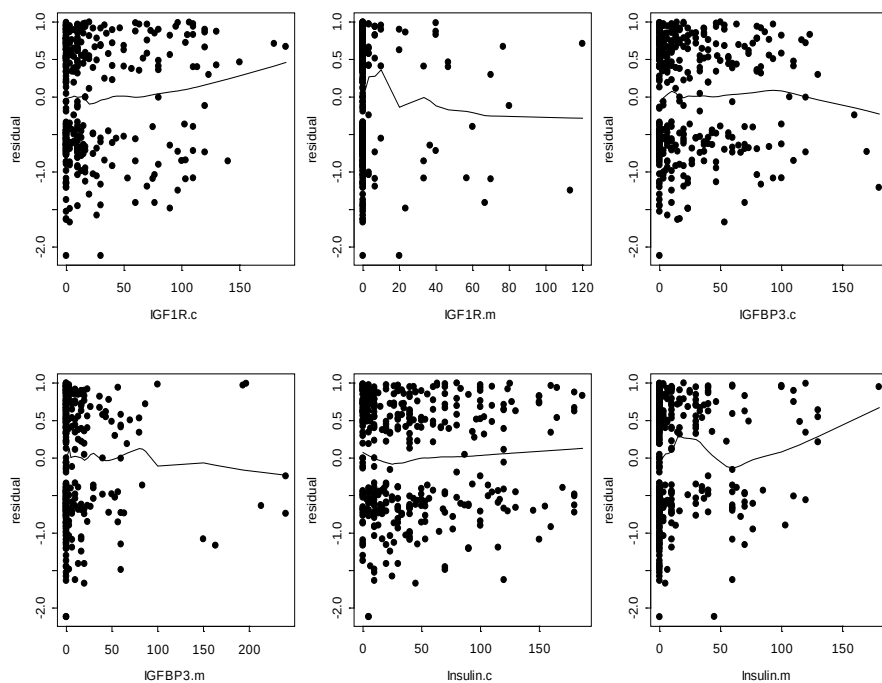
Table 16. Univariate Cox model assessing effect of covariates on recurrence free survival

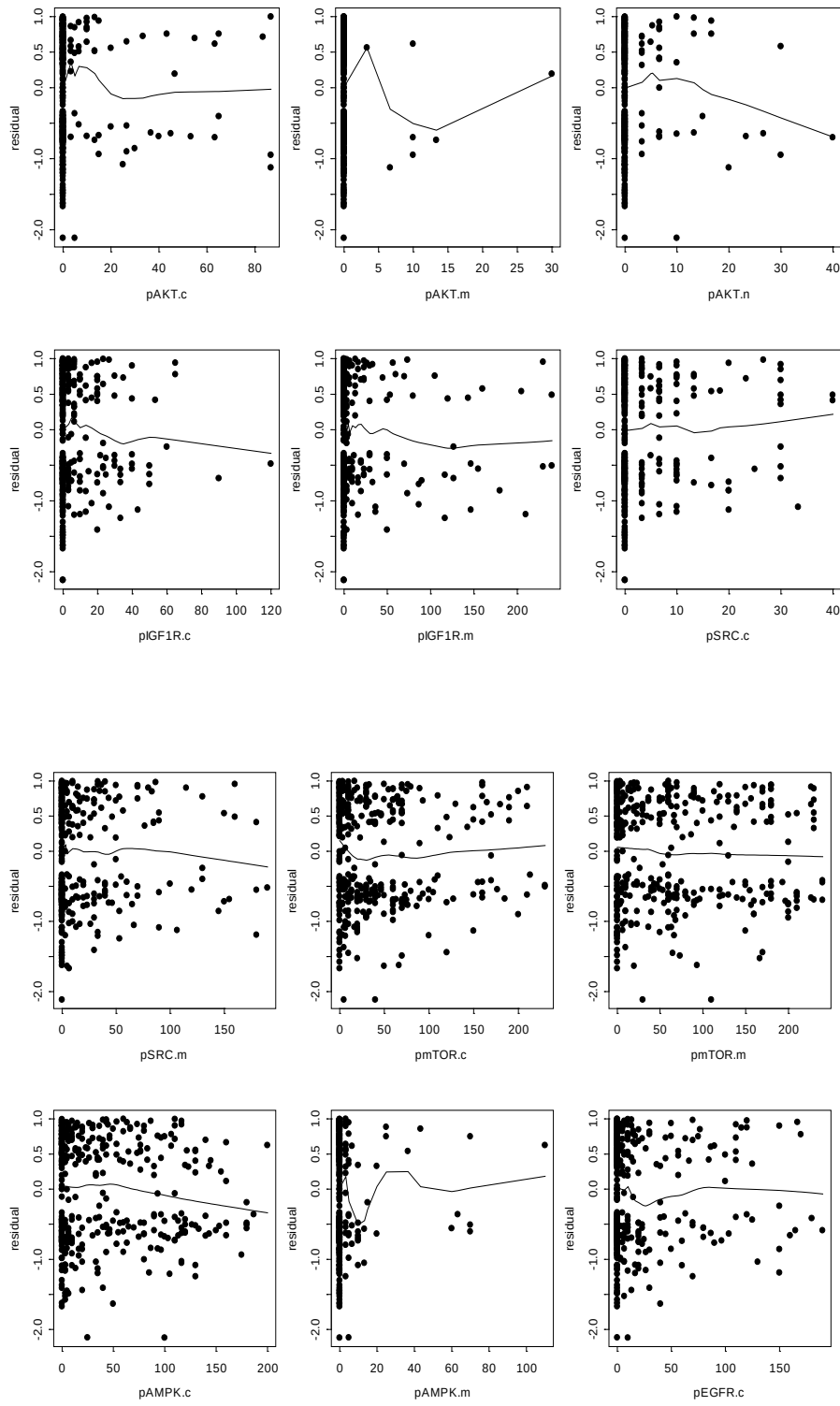
covariate	Estimate	StdErr	HazardRatio	HRLowerCL	HRUpperCL	pValue	Total	Event	Censored
IGF1R_c	0.0039	0.0018	1.0039	1.0004	1.0075	0.0300	367	206	161
IGF1R_c_01	0.1312	0.1422	1.1402	0.8629	1.5065	0.3562	367	206	161
IGF1R_c_01m	0.1178	0.1401	1.1250	0.8548	1.4805	0.4007	367	206	161
IGF1R_m	0.0011	0.0046	1.0011	0.9921	1.0102	0.8148	367	206	161
IGF1R_m_01	0.3607	0.1925	1.4344	0.9835	2.0919	0.0610	367	206	161
IGFBP3_c	0.0013	0.0020	1.0013	0.9973	1.0053	0.5231	370	209	161
IGFBP3_c_01	0.2362	0.1539	1.2664	0.9366	1.7124	0.1249	370	209	161
IGFBP3_c_01m	0.1895	0.1396	1.2087	0.9193	1.5892	0.1747	370	209	161
IGFBP3_m	-0.0006	0.0024	0.9994	0.9947	1.0041	0.7937	370	209	161
IGFBP3_m_01	0.1059	0.1444	1.1117	0.8377	1.4753	0.4632	370	209	161
Insulin_c	-0.0002	0.0015	0.9998	0.9968	1.0028	0.8965	367	209	158
Insulin_c_01	-0.2956	0.1879	0.7441	0.5148	1.0754	0.1157	367	209	158
Insulin_c_01m	-0.0065	0.1385	0.9935	0.7573	1.3033	0.9623	367	209	158
Insulin_m	0.0036	0.0024	1.0036	0.9990	1.0083	0.1283	366	209	157
Insulin_m_01	0.2649	0.1392	1.3033	0.9922	1.7120	0.0569	366	209	157

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
pAKT_c	-0.0018	0.0053	0.9982	0.9879	1.0087	0.7368	369	208	161
pAKT_c_01	0.1137	0.1937	1.1205	0.7666	1.6377	0.5570	369	208	161
pAKT_m	-0.0337	0.0411	0.9669	0.8921	1.0479	0.4121	370	209	161
pAKT_m_01	-0.7017	0.5853	0.4958	0.1574	1.5613	0.2306	370	209	161
pAKT_n	-0.0087	0.0151	0.9913	0.9625	1.0210	0.5615	370	209	161
pAKT_n_01	0.1099	0.2174	1.1162	0.7289	1.7093	0.6132	370	209	161
pAMPK_c	-0.0036	0.0015	0.9964	0.9935	0.9993	0.0154	370	209	161
pAMPK_c_01	-0.4203	0.1459	0.6568	0.4935	0.8742	0.0040	370	209	161
pAMPK_m	-0.0023	0.0071	0.9977	0.9840	1.0116	0.7465	370	209	161
pAMPK_m_01	-0.2542	0.2065	0.7755	0.5174	1.1623	0.2182	370	209	161
pEGFR_c	-0.0016	0.0020	0.9984	0.9945	1.0024	0.4375	368	207	161
pEGFR_c_01	-0.1404	0.1414	0.8690	0.6586	1.1466	0.3208	368	207	161
pEGFR_m	-0.0010	0.0024	0.9990	0.9943	1.0037	0.6795	369	208	161
pEGFR_m_01	-0.0561	0.1388	0.9454	0.7202	1.2411	0.6860	369	208	161
pEGFR_m_01m	-0.0561	0.1388	0.9454	0.7202	1.2411	0.6860	369	208	161
pEGFR_n	-0.0006	0.0018	0.9994	0.9959	1.0030	0.7605	369	209	160
pEGFR_n_01	-0.1760	0.1497	0.8386	0.6253	1.1246	0.2397	369	209	160
pEGFR_n_01m	-0.1250	0.1388	0.8825	0.6723	1.1585	0.3679	369	209	160
pIGF1R_c	-0.0083	0.0061	0.9918	0.9800	1.0037	0.1747	368	208	160
pIGF1R_c_01	0.0266	0.1480	1.0270	0.7684	1.3724	0.8574	368	208	160
pIGF1R_m	-0.0022	0.0021	0.9978	0.9938	1.0018	0.2850	368	208	160
pIGF1R_m_01	-0.0806	0.1541	0.9226	0.6820	1.2479	0.6010	368	208	160
pS6_c	-0.0006	0.0016	0.9994	0.9962	1.0025	0.6896	370	209	161
pS6_c_01	0.0294	0.1977	1.0298	0.6990	1.5172	0.8818	370	209	161
pS6_c_01m	0.0127	0.1386	1.0127	0.7718	1.3289	0.9273	370	209	161
pS6_m	0.0044	0.0029	1.0044	0.9986	1.0102	0.1363	370	209	161
pS6_m_01	0.1027	0.1704	1.1081	0.7935	1.5475	0.5467	370	209	161
pSRC_c	0.0071	0.0094	1.0071	0.9888	1.0258	0.4484	369	209	160
pSRC_c_01	0.1245	0.1524	1.1326	0.8401	1.5269	0.4139	369	209	160
pSRC_m	-0.0008	0.0020	0.9992	0.9952	1.0031	0.6761	370	209	161
pSRC_m_01	0.0526	0.1416	1.0540	0.7985	1.3912	0.7104	370	209	161
pmTOR_c	-0.0022	0.0015	0.9978	0.9948	1.0007	0.1383	369	209	160

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
pmTOR_c_01	-0.4859	0.1502	0.6152	0.4582	0.8258	0.0012	369	209	160
pmTOR_c_01m	-0.4377	0.1394	0.6455	0.4912	0.8485	0.0017	369	209	160
pmTOR_m	-0.0020	0.0011	0.9980	0.9960	1.0001	0.0625	370	209	161
pmTOR_m_01	-0.2257	0.1564	0.7979	0.5873	1.0842	0.1489	370	209	161
pmTOR_m_01m	-0.3023	0.1387	0.7391	0.5631	0.9700	0.0293	370	209	161

Figure 4. Martingale residual from Cox model with age, gender, histology and stage for recurrence free survival against each marker, Dr. HY Lee





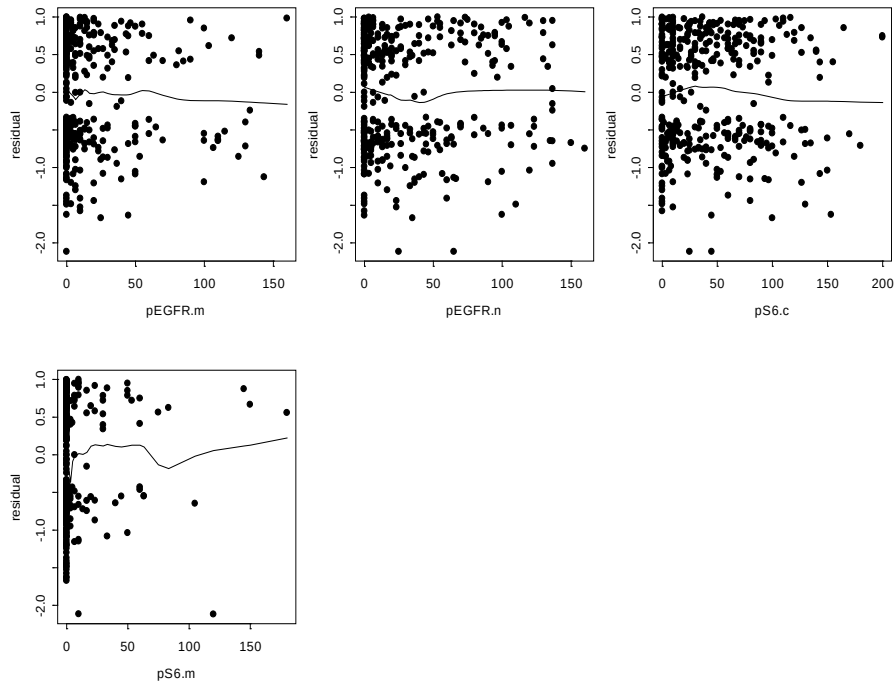


Table 17. Multivariate Cox model assessing effect of each marker independently on recurrence free survival, adjusting for age, gender, histology, stage and neoadjuvant treatment

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
IGF1R_c	0.0034	0.0022	1.0034	0.9991	1.0077	0.1257	364	204	160
IGF1R_c_01	0.0189	0.1590	1.0191	0.7462	1.3918	0.9054	364	204	160
IGF1R_c_01m	0.0753	0.1595	1.0782	0.7887	1.4738	0.6370	364	204	160
IGF1R_m	-0.0041	0.0051	0.9959	0.9860	1.0059	0.4174	364	204	160
IGF1R_m_01	0.1278	0.2152	1.1363	0.7453	1.7325	0.5525	364	204	160
IGFBP3_c	-0.0001	0.0021	0.9999	0.9958	1.0041	0.9795	367	207	160
IGFBP3_c_01	0.1661	0.1583	1.1807	0.8657	1.6104	0.2942	367	207	160
IGFBP3_c_01m	0.0760	0.1460	1.0790	0.8104	1.4365	0.6027	367	207	160
IGFBP3_m	-0.0018	0.0027	0.9982	0.9930	1.0034	0.4996	367	207	160
IGFBP3_m_01	-0.0360	0.1522	0.9647	0.7158	1.3001	0.8133	367	207	160
Insulin_c	0.0010	0.0016	1.0010	0.9979	1.0042	0.5218	364	207	157
Insulin_c_01	-0.2639	0.1935	0.7681	0.5256	1.1222	0.1726	364	207	157
Insulin_c_01m	0.0471	0.1427	1.0482	0.7924	1.3866	0.7415	364	207	157
Insulin_m	0.0036	0.0024	1.0036	0.9990	1.0083	0.1267	363	207	156

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
Insulin_m_01	0.3142	0.1426	1.3692	1.0353	1.8109	0.0276	363	207	156
pAKT_c	-0.0021	0.0053	0.9979	0.9876	1.0083	0.6877	366	206	160
pAKT_c_01	0.1373	0.2007	1.1471	0.7741	1.6999	0.4940	366	206	160
pAKT_m	-0.0222	0.0390	0.9780	0.9060	1.0558	0.5694	367	207	160
pAKT_m_01	-0.5288	0.5902	0.5893	0.1854	1.8737	0.3703	367	207	160
pAKT_n	-0.0150	0.0162	0.9851	0.9544	1.0168	0.3539	367	207	160
pAKT_n_01	0.0474	0.2212	1.0485	0.6797	1.6175	0.8305	367	207	160
pAMPK_c	-0.0030	0.0015	0.9970	0.9940	0.9999	0.0459	367	207	160
pAMPK_c_01	-0.4680	0.1485	0.6263	0.4681	0.8378	0.0016	367	207	160
pAMPK_m	0.0003	0.0070	1.0003	0.9866	1.0141	0.9712	367	207	160
pAMPK_m_01	-0.1904	0.2085	0.8267	0.5493	1.2439	0.3612	367	207	160
pEGFR_c	-0.0014	0.0021	0.9986	0.9945	1.0028	0.5250	365	205	160
pEGFR_c_01	-0.1730	0.1498	0.8411	0.6271	1.1282	0.2482	365	205	160
pEGFR_m	-0.0022	0.0025	0.9978	0.9930	1.0027	0.3767	366	206	160
pEGFR_m_01	-0.1368	0.1421	0.8721	0.6601	1.1523	0.3357	366	206	160
pEGFR_m_01m	-0.1368	0.1421	0.8721	0.6601	1.1523	0.3357	366	206	160
pEGFR_n	-0.0008	0.0018	0.9992	0.9956	1.0028	0.6577	366	207	159
pEGFR_n_01	-0.1134	0.1507	0.8928	0.6644	1.1996	0.4518	366	207	159
pEGFR_n_01m	-0.1426	0.1406	0.8671	0.6582	1.1423	0.3106	366	207	159
pIGF1R_c	-0.0080	0.0063	0.9920	0.9798	1.0044	0.2047	365	206	159
pIGF1R_c_01	0.0032	0.1547	1.0032	0.7408	1.3587	0.9833	365	206	159
pIGF1R_m	-0.0026	0.0022	0.9974	0.9932	1.0017	0.2347	365	206	159
pIGF1R_m_01	-0.0923	0.1587	0.9118	0.6681	1.2445	0.5607	365	206	159
pS6_c	-0.0008	0.0017	0.9992	0.9960	1.0025	0.6503	367	207	160
pS6_c_01	0.1173	0.2101	1.1245	0.7449	1.6976	0.5766	367	207	160
pS6_c_01m	-0.0015	0.1432	0.9985	0.7541	1.3220	0.9914	367	207	160
pS6_m	0.0017	0.0030	1.0017	0.9958	1.0077	0.5715	367	207	160
pS6_m_01	0.0577	0.1754	1.0594	0.7513	1.4940	0.7420	367	207	160
pSRC_c	0.0076	0.0102	1.0076	0.9876	1.0281	0.4582	366	207	159
pSRC_c_01	0.1231	0.1649	1.1310	0.8186	1.5626	0.4554	366	207	159
pSRC_m	-0.0014	0.0022	0.9986	0.9944	1.0029	0.5313	367	207	160
pSRC_m_01	-0.0209	0.1483	0.9793	0.7323	1.3096	0.8879	367	207	160

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
pmTOR_c	-0.0010	0.0016	0.9990	0.9957	1.0022	0.5308	366	207	159
pmTOR_c_01	-0.4535	0.1710	0.6354	0.4544	0.8885	0.0080	366	207	159
pmTOR_c_01m	-0.3559	0.1561	0.7006	0.5159	0.9513	0.0226	366	207	159
pmTOR_m	-0.0014	0.0012	0.9986	0.9962	1.0010	0.2612	367	207	160
pmTOR_m_01	-0.1325	0.1921	0.8759	0.6010	1.2764	0.4903	367	207	160
pmTOR_m_01m	-0.2206	0.1661	0.8021	0.5792	1.1106	0.1841	367	207	160

Final multivariate Cox Model assessing the following covariates on recurrence free survival

<i>Analysis of Maximum Likelihood Estimates</i>						
<i>Variable</i>	<i>Parameter Estimate</i>	<i>Standard Error</i>	<i>p-value</i>	<i>Hazard Ratio</i>	<i>95% Hazard Ratio Confidence Limits</i>	
<i>Age</i>	0.0255	0.0076	0.0008	1.026	1.011	1.041
<i>Stage II vs I</i>	0.5090	0.1808	0.0049	1.664	1.167	2.371
<i>III vs I</i>	0.9843	0.1856	<.0001	2.676	1.860	3.850
<i>Neoadjuvant (Yes vs No)</i>	0.2768	0.1884	0.1418	1.319	0.912	1.908
<i>IGF1R_c</i>	0.0032	0.0019	0.1021	1.003	0.999	1.007
<i>Insulin_m_01 (Pos vs 0)</i>	0.3957	0.1455	0.0065	1.485	1.117	1.976
<i>pAMPK_c_01 (Pos vs 0)</i>	-0.4929	0.1504	0.0010	0.611	0.455	0.820
<i>pmTOR_c_01 (Pos vs 0)</i>	-0.3956	0.1627	0.0150	0.673	0.489	0.926

3. Dr. Koo

3.1 Markers

Table 18. Descriptive summary of markers

<i>covariate</i>	<i>n</i>	<i>mean $\pm$ std, median (min, max)</i>
CASK_c	370	23.02 $\pm$ 31.18, 10 (0, 150)
CASK_m	370	23.5 $\pm$ 41.57, 0 (0, 180)
CD51_c	370	36.11 $\pm$ 42.47, 20 (0, 210)
CD51_m	370	29.2 $\pm$ 43.02, 10 (0, 240)
CXCR2_c	370	32.41 $\pm$ 30.83, 23.33 (0, 150)
CXCR2_n	367	15.29 $\pm$ 23.46, 3.33 (0, 120)
EpCAM_c	368	25.04 $\pm$ 32.99, 10 (0, 180)
EpCAM_m	370	63.27 $\pm$ 64.94, 44.17 (0, 270)
SPP1_c	370	5.3 $\pm$ 25.67, 0 (0, 180)
SPP1_n	369	56.53 $\pm$ 53.59, 45 (0, 240)

Figure 5. Distribution of markers, Dr. Koo (Red line – Mean, Green line – Median)

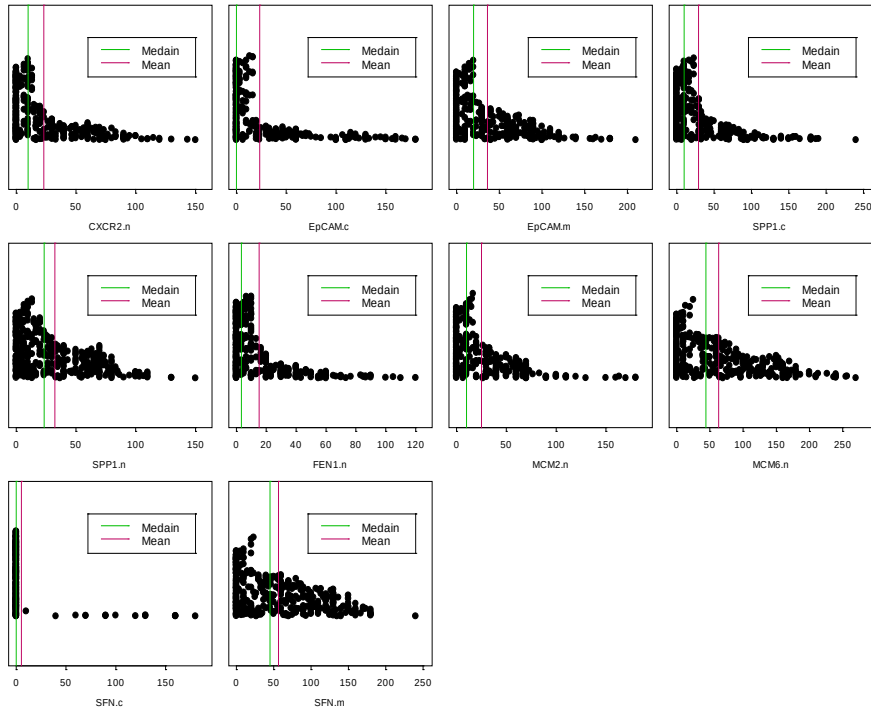


Table 19. Correlation between markers

Spearman Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations										
	CASK_c	CASK_m	CD51_c	CD51_m	CXCR2_c	CXCR2_n	EpCAM_c	EpCAM_m	SPP1_c	SPP1_n
CASK_c	1	0.20672	-0.20549	-0.25580	0.15845	-0.07132	0.00706	-0.08818	0.05163	-0.01888
	<.0001	<.0001	<.0001	<.0001	0.0022	0.1727	0.8927	0.0903	0.3220	0.7177
	370	370	370	370	370	367	368	370	370	369
CASK_m	0.20672	1	-0.19539	-0.13551	0.03810	-0.00168	-0.16675	-0.04155	-0.00873	0.13647
	<.0001	<.0001	0.0002	0.0091	0.4650	0.9745	0.0013	0.4255	0.8671	0.0087
	370	370	370	370	370	367	368	370	370	369
CD51_c	-0.20549	-0.19539	1	0.68597	0.19444	0.00772	0.18579	0.11893	0.09961	0.19700
	<.0001	0.0002	<.0001	<.0001	0.0002	0.8828	0.0003	0.0221	0.0556	0.0001
	370	370	370	370	370	367	368	370	370	369
CD51_m	-0.25580	-0.13551	0.68597	1	0.07311	-0.07714	0.04205	0.08436	0.00061	0.23997
	<.0001	0.0091	<.0001	<.0001	0.1605	0.1402	0.4213	0.1052	0.9907	<.0001
	370	370	370	370	370	367	368	370	370	369
CXCR2_c	0.15845	0.03810	0.19444	0.07311	1	-0.00031	0.24559	0.20347	0.21061	0.30982
	0.0022	0.4650	0.0002	0.1605	<.0001	0.9952	<.0001	<.0001	<.0001	<.0001
	370	370	370	370	370	367	368	370	370	369
CXCR2_n	-0.07132	-0.00168	0.00772	-0.07714	-0.00031	1	-0.03994	-0.12380	0.00190	-0.12987
	0.1727	0.9745	0.8828	0.1402	0.9952	<.0001	0.4468	0.0177	0.9711	0.0129
	367	367	367	367	367	367	365	367	367	366
EpCAM_c	0.00706	-0.16675	0.18579	0.04205	0.24559	-0.03994	1	0.47211	0.08455	0.06822
	0.8927	0.0013	0.0003	0.4213	<.0001	0.4468	<.0001	<.0001	0.1054	0.1922
	368	368	368	368	368	365	368	368	368	367
EpCAM_m	-0.08818	-0.04155	0.11893	0.08436	0.20347	-0.12380	0.47211	1	0.08976	0.17398
	0.0903	0.4255	0.0221	0.1052	<.0001	0.0177	<.0001	<.0001	0.0847	0.0008
	370	370	370	370	370	367	368	370	370	369
SPP1_c	0.05163	-0.00873	0.09961	0.00061	0.21061	0.00190	0.08455	0.08976	1	0.01999
	0.3220	0.8671	0.0556	0.9907	<.0001	0.9711	0.1054	0.0847	<.0001	0.7019
	370	370	370	370	370	367	368	370	370	369
SPP1_n	-0.01888	0.13647	0.19700	0.23997	0.30982	-0.12987	0.06822	0.17398	0.01999	1
	0.7177	0.0087	0.0001	<.0001	<.0001	0.0129	0.1922	0.0008	0.7019	<.0001
	369	369	369	369	369	366	367	369	369	369

Figure 6. Correlation between markers, Dr. Koo

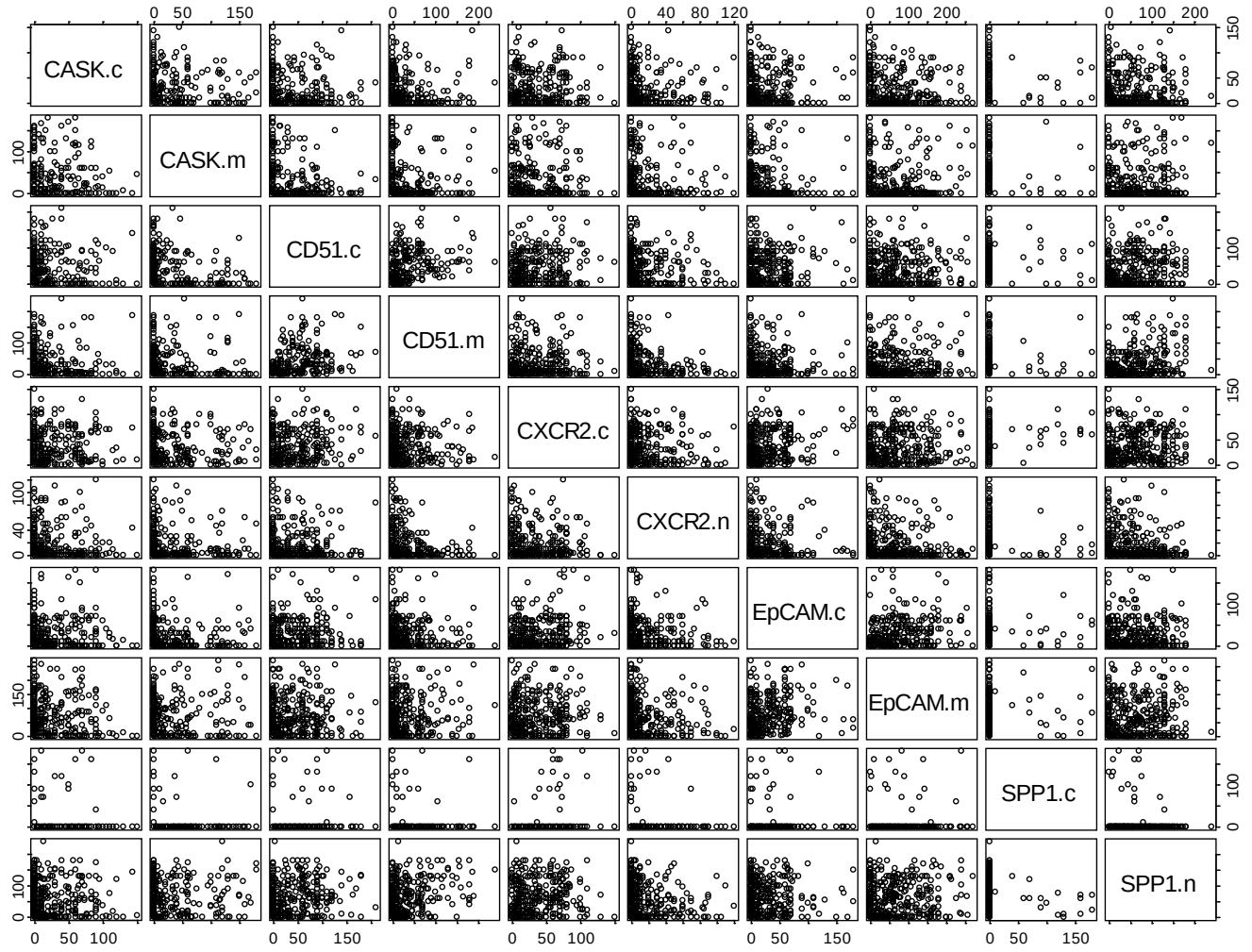


Table 20. Frequency tables for dichotomized markers

<i>covariate</i>	<i>levels</i>	<i>N (%)</i>	<i>covariate</i>	<i>levels</i>	<i>N (%)</i>
CASK_c_01	0	155(41.9%)	CASK_c_01m	0	178(48.1%)
	1	215(58.1%)		1	192(51.9%)
CASK_m_01	0	212(57.3%)	CD51_c_01m	0	184(49.7%)
	1	158(42.7%)		1	186(50.3%)
CD51_c_01	0	130(35.1%)	CD51_m_01m	0	160(43.2%)
	1	240(64.9%)		1	210(56.8%)
CD51_m_01	0	130(35.1%)	CXCR2_c_01m	0	184(49.7%)
	1	240(64.9%)		1	186(50.3%)

<i>covariate</i>	<i>levels</i>	<i>N (%)</i>	<i>covariate</i>	<i>levels</i>	<i>N (%)</i>
CXCR2_c_01	0	50(13.5%)	CXCR2_n_01m	missing	3
	1	320(86.5%)		0	158(43.1%)
CXCR2_n_01	missing	3		1	209(56.9%)
	0	156(42.5%)	EpCAM_c_01m	missing	2
	1	211(57.5%)		0	154(41.8%)
EpCAM_c_01	missing	2		1	214(58.2%)
	0	136(37%)	EpCAM_m_01m	0	185(50%)
	1	232(63%)		1	185(50%)
EpCAM_m_01	0	85(23%)	SPP1_n_01m	missing	1
	1	285(77%)		0	181(49.1%)
SPP1_c_01	0	352(95.1%)		1	188(50.9%)
	1	18(4.9%)			
SPP1_n_01	Missing	1			
	0	84(22.8%)			
	1	285(77.2%)			

Table 21. Markers by gender

<i>covariate</i>	<i>Gender</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CASK_c	F	184	19.72 ± 30.06, 6.67 (0, 143.33)	.0709
	M	186	26.28 ± 32, 10 (0, 150)	
CASK_m	F	184	24.4 ± 43.94, 0 (0, 180)	.8817
	M	186	22.61 ± 39.18, 0 (0, 180)	
CD51_c	F	184	39.15 ± 42.68, 30 (0, 180)	.0621
	M	186	33.11 ± 42.16, 11.67 (0, 210)	
CD51_m	F	184	31.9 ± 43.14, 10 (0, 190)	.0187
	M	186	26.53 ± 42.85, 10 (0, 240)	
CXCR2_c	F	184	30.33 ± 30.23, 20 (0, 150)	.1397
	M	186	34.47 ± 31.35, 25.83 (0, 130)	
CXCR2_n	F	184	15 ± 22.76, 5 (0, 105)	.4713
	M	183	15.59 ± 24.2, 3.33 (0, 120)	
EpCAM_c	F	184	25.38 ± 34.73, 10 (0, 180)	.7882

<i>covariate</i>	<i>Gender</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	M	184	24.7 ± 31.24, 10 (0, 180)	
EpCAM_m	F	184	63.13 ± 67.18, 41.67 (0, 256.67)	.7333
	M	186	63.4 ± 62.84, 46.67 (0, 270)	
SPP1_c	F	184	3.8 ± 20.34, 0 (0, 180)	.6123
	M	186	6.77 ± 30.01, 0 (0, 180)	
SPP1_n	F	184	53.48 ± 50.83, 45.83 (0, 180)	.3588
	M	185	59.56 ± 56.18, 43.33 (0, 240)	

Table 22. Markers by Race

<i>covariate</i>	<i>Race</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CASK_c	Black	21	24.29 ± 28.47, 15 (0, 90)	.8357
	Hispanic	14	27.02 ± 40.67, 8.33 (0, 143.33)	
	Oriental	5	29.33 ± 39.54, 3.33 (0, 83.33)	
	White	330	22.67 ± 30.91, 10 (0, 150)	
CASK_m	Black	21	25.32 ± 43.11, 0 (0, 133.33)	.5940
	Hispanic	14	7.38 ± 16.75, 0 (0, 60)	
	Oriental	5	23.33 ± 48.53, 0 (0, 110)	
	White	330	24.07 ± 42.12, 0 (0, 180)	
CD51_c	Black	21	61.43 ± 70.17, 40 (0, 210)	.7053
	Hispanic	14	35.24 ± 45, 20 (0, 140)	
	Oriental	5	36.33 ± 42.63, 10 (0, 90)	
	White	330	34.54 ± 39.73, 20 (0, 180)	
CD51_m	Black	21	37.38 ± 44.2, 23.33 (0, 150)	.7361
	Hispanic	14	25.71 ± 49.5, 7.5 (0, 186.67)	
	Oriental	5	45.33 ± 76.11, 10 (0, 180)	
	White	330	28.59 ± 42.2, 10 (0, 240)	
CXCR2_c	Black	21	36.67 ± 30.72, 33.33 (0, 85)	.7963
	Hispanic	14	39.52 ± 31.56, 60 (0, 76.67)	
	Oriental	5	32.67 ± 35.85, 10 (3.33, 76.67)	
	White	330	31.83 ± 30.82, 20 (0, 150)	
CXCR2_n	Black	21	13.29 ± 19.88, 5 (0, 83.33)	.3951

<i>covariate</i>	<i>Race</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	Hispanic	14	24.05 ± 26.99, 13.33 (0, 86.67)	
	Oriental	5	3.33 ± 4.71, 0 (0, 10)	
	White	327	15.23 ± 23.63, 3.33 (0, 120)	
	Black	21	19.6 ± 29.19, 6.67 (0, 110)	
EpCAM_c	Hispanic	14	43.45 ± 59.98, 5 (0, 163.33)	.7884
	Oriental	5	62 ± 80.12, 20 (0, 180)	
	White	328	24.04 ± 30.16, 10 (0, 180)	
	Black	21	73.89 ± 58.36, 83.33 (0, 166.67)	
EpCAM_m	Hispanic	14	43.45 ± 64.52, 20.83 (0, 240)	.3924
	Oriental	5	67.33 ± 68.94, 50 (0, 180)	
	White	330	63.37 ± 65.4, 44.17 (0, 270)	
	Black	21	3.33 ± 15.28, 0 (0, 70)	
SPP1_c	Hispanic	14	20.71 ± 52.98, 0 (0, 160)	.1207
	Oriental	5	32 ± 71.55, 0 (0, 160)	
	White	330	4.36 ± 23.02, 0 (0, 180)	
	Black	21	69.76 ± 60.28, 70 (0, 180)	
SPP1_n	Hispanic	14	42.08 ± 42.62, 31.25 (0, 143.33)	.7318
	Oriental	5	54 ± 61.48, 60 (0, 150)	
	White	329	56.34 ± 53.5, 45 (0, 240)	
	Black	21	69.76 ± 60.28, 70 (0, 180)	

Table 23. Markers by Smoking Status

<i>covariate</i>	<i>smoker</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CASK_c	1 Never	38	19.25 ± 32.9, 5 (0, 120)	.5451
	2 Former	170	23.76 ± 32.39, 10 (0, 150)	
	3 Current	162	23.12 ± 29.57, 10 (0, 143.33)	
CASK_m	1 Never	38	17.72 ± 33.65, 0 (0, 130)	.9028
	2 Former	170	25.5 ± 45.02, 0 (0, 180)	
	3 Current	162	22.76 ± 39.49, 0 (0, 180)	
CD51_c	1 Never	38	36.45 ± 34.69, 30 (0, 90)	.0634
	2 Former	170	30.72 ± 39.54, 14.17 (0, 210)	
	3 Current	162	41.7 ± 46.41, 25 (0, 180)	

<i>covariate</i>	<i>smoker</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CD51_m	1 Never	38	30.92 ± 33.07, 21.67 (0, 130)	.1310
	2 Former	170	25.4 ± 41.44, 10 (0, 240)	
	3 Current	162	32.79 ± 46.46, 10 (0, 186.67)	
CXCR2_c	1 Never	38	24.47 ± 27.99, 13.33 (0, 80)	.0265
	2 Former	170	31.83 ± 33.05, 16.67 (0, 150)	
	3 Current	162	34.88 ± 28.81, 30 (0, 110)	
CXCR2_n	1 Never	38	18.97 ± 25.07, 10 (0, 100)	.0899
	2 Former	170	15.18 ± 22.72, 4.17 (0, 110)	
	3 Current	159	14.54 ± 23.9, 3.33 (0, 120)	
EpCAM_c	1 Never	38	32.79 ± 52.63, 10 (0, 180)	.9193
	2 Former	169	23 ± 26.48, 13.33 (0, 110)	
	3 Current	161	25.36 ± 33.18, 10 (0, 180)	
EpCAM_m	1 Never	38	49.25 ± 53.29, 30 (0, 180)	.6047
	2 Former	170	63.38 ± 63.06, 45 (0, 255)	
	3 Current	162	66.44 ± 69.16, 46.67 (0, 270)	
SPP1_c	1 Never	38	0 ± 0, 0 (0, 0)	.2727
	2 Former	170	4.18 ± 21.88, 0 (0, 180)	
	3 Current	162	7.72 ± 31.53, 0 (0, 180)	
SPP1_n	1 Never	38	45.85 ± 49.49, 34.92 (0, 160)	.3213
	2 Former	169	56.82 ± 55.15, 43.33 (0, 180)	
	3 Current	162	58.73 ± 52.88, 51.67 (0, 240)	

Table 24. Markers by Histology

<i>covariate</i>	<i>histology0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CASK_c	ADENO	227	13.47 ± 23.91, 0 (0, 120)	<.0001
	Other	17	34.22 ± 36.93, 20 (0, 105)	
	SCC	126	38.7 ± 35.08, 30 (0, 150)	
CASK_m	ADENO	227	14.04 ± 31.6, 0 (0, 160)	<.0001
	Other	17	18.04 ± 38.28, 0 (0, 150)	
	SCC	126	41.28 ± 51.21, 16.67 (0, 180)	
CD51_c	ADENO	227	42.37 ± 40.13, 30 (0, 210)	<.0001

<i>covariate</i>	<i>histology0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	Other	17	45.69 ± 54, 10 (0, 156.67)	
	SCC	126	23.54 ± 42.38, 0 (0, 180)	
CD51_m	ADENO	227	33.19 ± 41.5, 16.67 (0, 240)	<.0001
	Other	17	32.75 ± 55.81, 10 (0, 190)	
	SCC	126	21.55 ± 43.14, 0 (0, 186.67)	
CXCR2_c	ADENO	227	31.59 ± 32.04, 20 (0, 150)	.1277
	Other	17	43.92 ± 29.18, 50 (0, 85)	
	SCC	126	32.34 ± 28.64, 23.33 (0, 110)	
CXCR2_n	ADENO	225	16.68 ± 24.38, 5 (0, 105)	.4110
	Other	17	7.57 ± 14.44, 2 (0, 56.67)	
	SCC	125	13.85 ± 22.58, 3.33 (0, 120)	
EpCAM_c	ADENO	226	31.02 ± 34.88, 20 (0, 180)	<.0001
	Other	17	36.86 ± 28.59, 33.33 (0, 90)	
	SCC	125	12.63 ± 25.84, 0 (0, 160)	
EpCAM_m	ADENO	227	77.49 ± 65.39, 60 (0, 270)	<.0001
	Other	17	105.88 ± 74.92, 100 (0, 255)	
	SCC	126	31.9 ± 48.93, 6.67 (0, 200)	
SPP1_c	ADENO	227	2.38 ± 16.12, 0 (0, 130)	.0004
	Other	17	22.35 ± 49.06, 0 (0, 180)	
	SCC	126	8.25 ± 33.23, 0 (0, 180)	
SPP1_n	ADENO	227	52.33 ± 52.04, 40 (0, 180)	.1672
	Other	17	61.67 ± 51.37, 66.67 (0, 166.67)	
	SCC	125	63.45 ± 56.24, 53.33 (0, 240)	

Table 25. Markers by Stage

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CASK_c	IA	103	18.07 ± 28.63, 6.67 (0, 150)	.0424
	IB	131	25.66 ± 32.4, 10 (0, 143.33)	
	IIA	22	39.24 ± 36.39, 35 (0, 110)	
	IIB	53	25.75 ± 31.64, 10 (0, 110)	
	IIIA	61	17.46 ± 28.09, 3.33 (0, 130)	

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CASK_m	IA	103	21.36 ± 39.14, 0 (0, 156.67)	.2911
	IB	131	23.37 ± 40.52, 0 (0, 170)	
	IIA	22	16.06 ± 39.7, 0 (0, 150)	
	IIB	53	33.18 ± 49.05, 3.33 (0, 180)	
	IIIA	61	21.67 ± 41.4, 0 (0, 180)	
CD51_c	IA	103	31.39 ± 35.77, 20 (0, 160)	.2951
	IB	131	37.53 ± 42.19, 20 (0, 210)	
	IIA	22	21.74 ± 34.24, 10 (0, 120)	
	IIB	53	34.18 ± 44.51, 15 (0, 180)	
	IIIA	61	47.9 ± 51.68, 33.33 (0, 180)	
CD51_m	IA	103	22.02 ± 33.65, 10 (0, 180)	.5473
	IB	131	29.33 ± 41.08, 10 (0, 190)	
	IIA	22	22.27 ± 41.55, 10 (0, 180)	
	IIB	53	33.55 ± 49, 10 (0, 180)	
	IIIA	61	39.78 ± 53.66, 10 (0, 240)	
CXCR2_c	IA	103	25.68 ± 28.83, 16.67 (0, 130)	.0548
	IB	131	35.94 ± 32.28, 30 (0, 130)	
	IIA	22	39.24 ± 28.98, 35 (0, 90)	
	IIB	53	30.75 ± 28.21, 16.67 (0, 100)	
	IIIA	61	35.16 ± 32.52, 26.67 (0, 150)	
CXCR2_n	IA	103	20.57 ± 28.24, 6.67 (0, 110)	.2727
	IB	129	13.29 ± 21.54, 3.33 (0, 90)	
	IIA	22	19.32 ± 30.21, 5 (0, 120)	
	IIB	52	11.83 ± 18.34, 3.33 (0, 73.33)	
	IIIA	61	12.12 ± 17.97, 5 (0, 86.67)	
EpCAM_c	IA	101	21.11 ± 24.1, 10 (0, 110)	.7732
	IB	131	26.32 ± 37.39, 10 (0, 180)	
	IIA	22	31.29 ± 41.92, 10 (0, 180)	
	IIB	53	23.52 ± 31.17, 3.33 (0, 130)	
	IIIA	61	27.88 ± 33.88, 10 (0, 160)	
EpCAM_m	IA	103	51.88 ± 52.17, 40 (0, 186.67)	.7263
	IB	131	69.52 ± 71.97, 50 (0, 256.67)	

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
SPP1_c	IIA	22	65.83 ± 67.43, 43.33 (0, 240)	.1738
	IIB	53	66.5 ± 70.5, 51 (0, 240)	
	IIIA	61	65.36 ± 62.18, 50 (0, 270)	
	IA	103	1.65 ± 11.21, 0 (0, 90)	
	IB	131	4.05 ± 21.4, 0 (0, 160)	
	IIA	22	0 ± 0, 0 (0, 0)	
	IIB	53	11.13 ± 38.11, 0 (0, 180)	
	IIIA	61	10.98 ± 38.59, 0 (0, 180)	
	IA	103	50.96 ± 55.14, 32.5 (0, 240)	
SPP1_n	IB	130	58.46 ± 51.08, 50 (0, 180)	.1666
	IIA	22	37.95 ± 43.13, 29.17 (0, 180)	
	IIB	53	66.19 ± 56.2, 63.33 (0, 180)	
	IIIA	61	60.14 ± 56.26, 46.67 (0, 180)	
	IA	103	50.96 ± 55.14, 32.5 (0, 240)	

<i>covariate</i>	<i>stage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CASK_c	I	234	22.32 ± 30.96, 10 (0, 150)	.1094
	II	75	29.71 ± 33.43, 15 (0, 110)	
	IIIA	61	17.46 ± 28.09, 3.33 (0, 130)	
CASK_m	I	234	22.49 ± 39.85, 0 (0, 170)	.5254
	II	75	28.16 ± 46.9, 0 (0, 180)	
	IIIA	61	21.67 ± 41.4, 0 (0, 180)	
CD51_c	I	234	34.83 ± 39.53, 20 (0, 210)	.1404
	II	75	30.53 ± 41.92, 10 (0, 180)	
	IIIA	61	47.9 ± 51.68, 33.33 (0, 180)	
CD51_m	I	234	26.11 ± 38.08, 10 (0, 190)	.2465
	II	75	30.24 ± 46.94, 10 (0, 180)	
	IIIA	61	39.78 ± 53.66, 10 (0, 240)	
CXCR2_c	I	234	31.42 ± 31.16, 20 (0, 130)	.3639
	II	75	33.24 ± 28.51, 20 (0, 100)	
	IIIA	61	35.16 ± 32.52, 26.67 (0, 150)	
CXCR2_n	I	232	16.52 ± 24.95, 5 (0, 110)	.6085
	II	74	14.05 ± 22.57, 3.33 (0, 120)	

<i>covariate</i>	<i>stage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
EpCAM_c	IIIA	61	12.12 ± 17.97, 5 (0, 86.67)	.6398
	I	232	24.05 ± 32.33, 10 (0, 180)	
	II	75	25.8 ± 34.55, 10 (0, 180)	
EpCAM_m	IIIA	61	27.88 ± 33.88, 10 (0, 160)	.7502
	I	234	61.75 ± 64.49, 41.67 (0, 256.67)	
	II	75	66.3 ± 69.16, 50 (0, 240)	
SPP1_c	IIIA	61	65.36 ± 62.18, 50 (0, 270)	.1985
	I	234	2.99 ± 17.66, 0 (0, 160)	
	II	75	7.87 ± 32.35, 0 (0, 180)	
SPP1_n	IIIA	61	10.98 ± 38.59, 0 (0, 180)	.8341
	I	233	55.14 ± 52.93, 40 (0, 240)	
	II	75	57.91 ± 53.99, 53.33 (0, 180)	
	IIIA	61	60.14 ± 56.26, 46.67 (0, 180)	

Table 26. Markers by Grade

<i>covariate</i>	<i>grade0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CASK_c	1 Poorly	122	23.57 ± 30.42, 10 (0, 115)	.4213
	2 Moderately	199	23.64 ± 32.97, 10 (0, 150)	
	3 Well	36	13.7 ± 20.36, 5 (0, 83.33)	
CASK_m	1 Poorly	122	23.16 ± 43.91, 0 (0, 180)	.4954
	2 Moderately	199	24.45 ± 41.34, 0 (0, 180)	
	3 Well	36	17.73 ± 34.64, 0 (0, 156.67)	
CD51_c	1 Poorly	122	39.56 ± 45.46, 21.67 (0, 180)	.4071
	2 Moderately	199	33.31 ± 41.12, 16.67 (0, 210)	
	3 Well	36	38.24 ± 37.12, 30 (0, 115)	
CD51_m	1 Poorly	122	28.91 ± 44.01, 10 (0, 240)	.3402
	2 Moderately	199	29.43 ± 43.45, 10 (0, 186.67)	
	3 Well	36	26.94 ± 27.15, 18.33 (0, 90)	
CXCR2_c	1 Poorly	122	40.25 ± 30.45, 31.67 (0, 130)	<.0001
	2 Moderately	199	30.25 ± 31, 16.67 (0, 150)	
	3 Well	36	16.39 ± 23.48, 3.33 (0, 80)	

<i>covariate</i>	<i>grade0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
CXCR2_n	1 Poorly	120	9.67 ± 17.82, 0 (0, 105)	.0024
	2 Moderately	198	18.43 ± 26.11, 5.83 (0, 120)	
	3 Well	36	18.94 ± 23.56, 10 (0, 100)	
EpCAM_c	1 Poorly	121	29.8 ± 33.2, 20 (0, 160)	.0129
	2 Moderately	198	21.61 ± 32.58, 10 (0, 180)	
	3 Well	36	23.29 ± 34.13, 10 (0, 163.33)	
EpCAM_m	1 Poorly	122	65.98 ± 57.31, 60 (0, 240)	.1450
	2 Moderately	199	60.53 ± 69.12, 33.33 (0, 270)	
	3 Well	36	44.68 ± 44.28, 30.83 (0, 170)	
SPP1_c	1 Poorly	122	9.34 ± 34.52, 0 (0, 180)	.0959
	2 Moderately	199	2.76 ± 16.93, 0 (0, 160)	
	3 Well	36	0 ± 0, 0 (0, 0)	
SPP1_n	1 Poorly	122	68.3 ± 55.12, 66.67 (0, 180)	.0002
	2 Moderately	198	53.06 ± 53.08, 41.67 (0, 240)	
	3 Well	36	32.75 ± 42.41, 10 (0, 150)	

3.2 Overall survival

Table 27. Univariate Cox model assessing effect of covariates on overall survival

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
CASK_c	0.0002	0.0025	1.0002	0.9953	1.0051	0.9406	370	160	210
CASK_c_01	0.0926	0.1615	1.0970	0.7994	1.5054	0.5663	370	160	210
CASK_c_01m	0.1126	0.1588	1.1192	0.8198	1.5279	0.4783	370	160	210
CASK_m	-0.0019	0.0021	0.9981	0.9941	1.0021	0.3497	370	160	210
CASK_m_01	-0.1135	0.1612	0.8927	0.6509	1.2243	0.4813	370	160	210
CD51_c	0.0014	0.0018	1.0014	0.9978	1.0049	0.4502	370	160	210
CD51_c_01	0.0805	0.1673	1.0838	0.7808	1.5044	0.6304	370	160	210
CD51_c_01m	0.1594	0.1585	1.1728	0.8596	1.5999	0.3146	370	160	210
CD51_m	0.0009	0.0017	1.0009	0.9976	1.0041	0.5970	370	160	210
CD51_m_01	0.1351	0.1684	1.1446	0.8228	1.5922	0.4225	370	160	210
CD51_m_01m	0.2230	0.1620	1.2498	0.9097	1.7169	0.1688	370	160	210

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
CXCR2_c	0.0031	0.0024	1.0031	0.9985	1.0079	0.1890	370	160	210
CXCR2_c_01	0.4075	0.2566	1.5030	0.9090	2.4853	0.1123	370	160	210
CXCR2_c_01m	0.3306	0.1596	1.3917	1.0180	1.9028	0.0383	370	160	210
CXCR2_n	0.0010	0.0033	1.0010	0.9945	1.0074	0.7723	367	158	209
CXCR2_n_01	-0.0782	0.1602	0.9248	0.6756	1.2659	0.6255	367	158	209
CXCR2_n_01m	-0.0476	0.1602	0.9535	0.6966	1.3052	0.7664	367	158	209
EpCAM_c	-0.0022	0.0026	0.9978	0.9926	1.0029	0.3942	368	159	209
EpCAM_c_01	-0.3790	0.1602	0.6845	0.5001	0.9370	0.0180	368	159	209
EpCAM_c_01m	-0.3786	0.1589	0.6848	0.5016	0.9350	0.0172	368	159	209
EpCAM_m	0.0001	0.0012	1.0001	0.9977	1.0025	0.9551	370	160	210
EpCAM_m_01	0.0339	0.1933	1.0344	0.7082	1.5110	0.8610	370	160	210
EpCAM_m_01m	-0.0802	0.1584	0.9229	0.6766	1.2588	0.6124	370	160	210
SPP1_c	0.0041	0.0027	1.0041	0.9988	1.0094	0.1266	370	160	210
SPP1_c_01	0.4536	0.3273	1.5739	0.8287	2.9892	0.1658	370	160	210
SPP1_n	0.0006	0.0015	1.0006	0.9977	1.0035	0.6815	369	159	210
SPP1_n_01	0.1042	0.1897	1.1098	0.7653	1.6094	0.5828	369	159	210
SPP1_n_01m	0.0779	0.1588	1.0810	0.7918	1.4758	0.6239	369	159	210

Figure 7. Martingale residual from Cox model with age, gender, histology and stage for overall survival against each marker, Dr. Koo

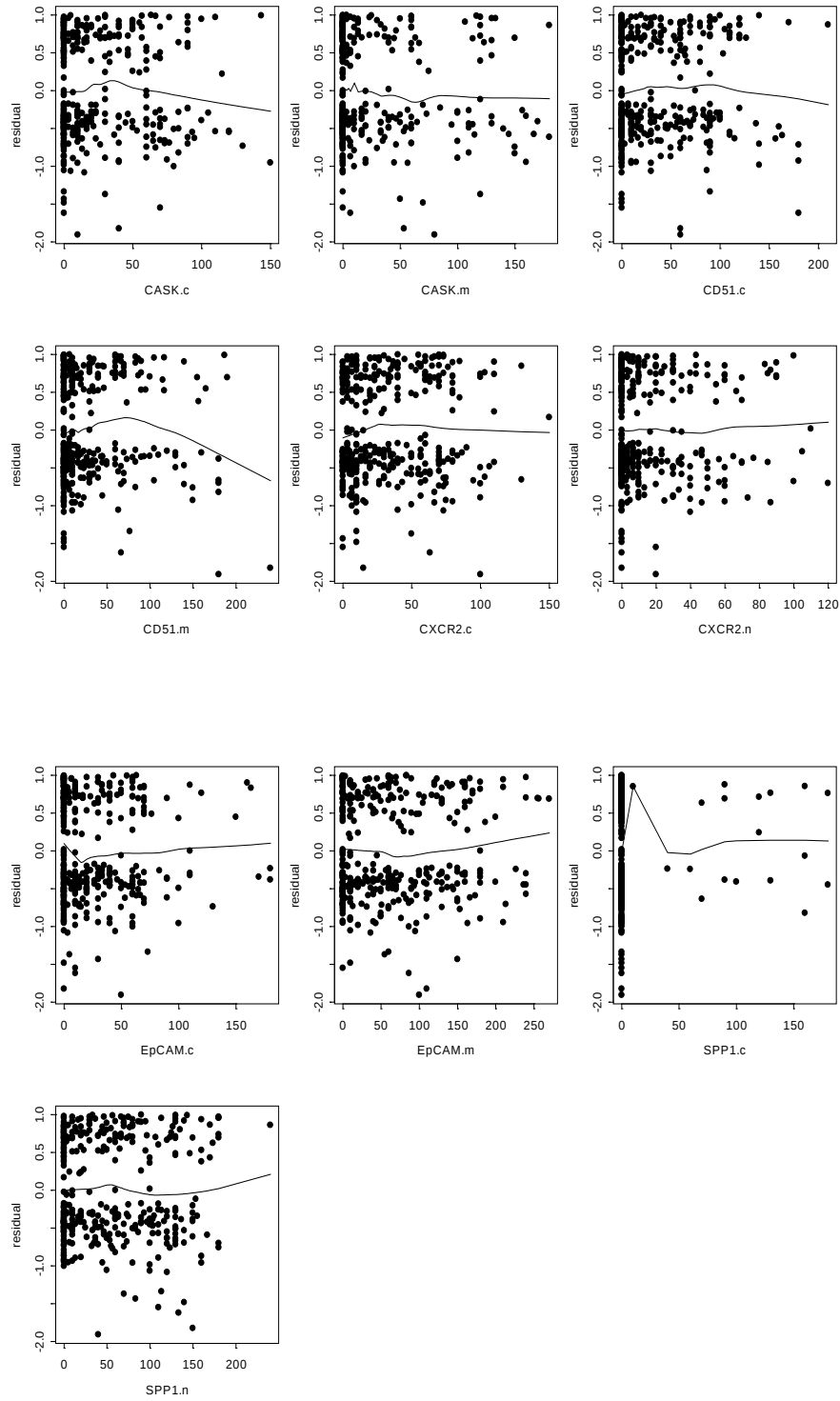


Table 28. Multicovariate Cox model assessing effect of each marker independently on overall survival, adjusting for age, gender, histology, stage and neoadjuvant treatment

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
CASK_c	-0.0023	0.0029	0.9977	0.9922	1.0034	0.4298	367	158	209
CASK_c_01	0.0482	0.1796	1.0494	0.7379	1.4922	0.7886	367	158	209
CASK_c_01m	0.0458	0.1774	1.0469	0.7394	1.4822	0.7961	367	158	209
CASK_m	-0.0033	0.0022	0.9967	0.9924	1.0011	0.1397	367	158	209
CASK_m_01	-0.2213	0.1696	0.8015	0.5748	1.1177	0.1921	367	158	209
CD51_c	0.0012	0.0018	1.0012	0.9976	1.0048	0.5162	367	158	209
CD51_c_01	0.2402	0.1816	1.2714	0.8907	1.8150	0.1860	367	158	209
CD51_c_01m	0.2333	0.1700	1.2627	0.9050	1.7619	0.1699	367	158	209
CD51_m	-0.0001	0.0016	0.9999	0.9967	1.0031	0.9595	367	158	209
CD51_m_01	0.3009	0.1853	1.3511	0.9397	1.9427	0.1043	367	158	209
CD51_m_01m	0.3323	0.1754	1.3941	0.9885	1.9661	0.0582	367	158	209
CXCR2_c	0.0028	0.0024	1.0028	0.9981	1.0075	0.2396	367	158	209
CXCR2_c_01	0.3365	0.2614	1.4001	0.8389	2.3368	0.1979	367	158	209
CXCR2_c_01m	0.3424	0.1621	1.4083	1.0250	1.9351	0.0347	367	158	209
CXCR2_n	0.0009	0.0033	1.0009	0.9944	1.0073	0.7952	364	156	208
CXCR2_n_01	-0.0189	0.1619	0.9813	0.7145	1.3479	0.9073	364	156	208
CXCR2_n_01m	0.0127	0.1621	1.0127	0.7372	1.3914	0.9377	364	156	208
EpCAM_c	-0.0020	0.0028	0.9980	0.9926	1.0035	0.4805	365	157	208
EpCAM_c_01	-0.4299	0.1710	0.6506	0.4654	0.9096	0.0119	365	157	208
EpCAM_c_01m	-0.4346	0.1720	0.6476	0.4623	0.9071	0.0115	365	157	208
EpCAM_m	0.0004	0.0014	1.0004	0.9977	1.0031	0.7734	367	158	209
EpCAM_m_01	-0.0037	0.2144	0.9963	0.6544	1.5167	0.9862	367	158	209
EpCAM_m_01m	-0.1142	0.1728	0.8921	0.6358	1.2517	0.5088	367	158	209
SPP1_c	0.0024	0.0027	1.0024	0.9971	1.0078	0.3697	367	158	209
SPP1_c_01	0.3789	0.3380	1.4607	0.7532	2.8330	0.2622	367	158	209
SPP1_n	-0.0005	0.0015	0.9995	0.9965	1.0025	0.7297	366	157	209
SPP1_n_01	0.0204	0.1921	1.0206	0.7003	1.4873	0.9154	366	157	209
SPP1_n_01m	-0.0180	0.1639	0.9822	0.7123	1.3543	0.9126	366	157	209

Final multivariate Cox Model assessing the following covariates on overall survival

Analysis of Maximum Likelihood Estimates						
Variable	Parameter Estimate	Standard Error	p-value	Hazard Ratio	95% Hazard Ratio Confidence Limits	
Age	0.0283	0.0087	0.0012	1.029	1.011	1.046
Gender M vs F	0.2622	0.1628	0.1073	1.300	0.945	1.788
Stage II vs I	0.4139	0.2051	0.0436	1.513	1.012	2.261
III vs I	0.8970	0.2148	<.0001	2.452	1.610	3.736
Neoadjuvant (Yes vs No)	0.4550	0.2101	0.0303	1.576	1.044	2.379
CXCR2_c_01m (≥ 23.3 vs < 23.3)	0.4406	0.1651	0.0076	1.554	1.124	2.147
EpCAM_c_01 (Pos vs 0)	-0.4981	0.1673	0.0029	0.608	0.438	0.844

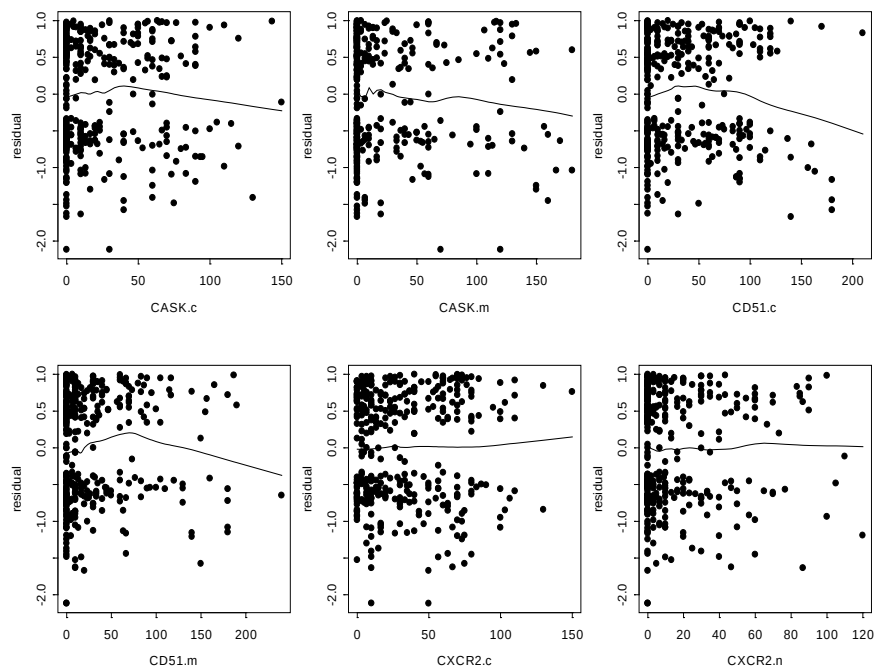
3.3 Recurrence free survival

Table 29. Univariate Cox model assessing effect of covariates on recurrence free survival

covariate	Estimate	StdErr	HazardRatio	HRLowerCL	HRUpperCL	pValue	Total	Event	Censored
CASK_c	0.0016	0.0021	1.0016	0.9975	1.0057	0.4419	370	209	161
CASK_c_01	0.1378	0.1415	1.1477	0.8698	1.5145	0.3301	370	209	161
CASK_c_01m	0.1384	0.1390	1.1485	0.8745	1.5082	0.3194	370	209	161
CASK_m	-0.0018	0.0018	0.9982	0.9948	1.0017	0.3082	370	209	161
CASK_m_01	-0.0642	0.1402	0.9378	0.7124	1.2344	0.6468	370	209	161
CD51_c	-0.0003	0.0016	0.9997	0.9965	1.0028	0.8341	370	209	161
CD51_c_01	0.0360	0.1447	1.0366	0.7807	1.3765	0.8038	370	209	161
CD51_c_01m	0.0986	0.1384	1.1037	0.8414	1.4477	0.4762	370	209	161
CD51_m	0.0022	0.0014	1.0022	0.9994	1.0050	0.1312	370	209	161
CD51_m_01	0.1429	0.1469	1.1536	0.8649	1.5386	0.3308	370	209	161
CD51_m_01m	0.2296	0.1415	1.2581	0.9533	1.6603	0.1048	370	209	161
CXCR2_c	0.0017	0.0022	1.0017	0.9974	1.0061	0.4354	370	209	161
CXCR2_c_01	0.2031	0.2096	1.2252	0.8124	1.8478	0.3327	370	209	161
CXCR2_c_01m	0.1833	0.1390	1.2011	0.9148	1.5771	0.1872	370	209	161
CXCR2_n	0.0004	0.0029	1.0004	0.9947	1.0061	0.8987	367	207	160

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
CXCR2_n_01	-0.0940	0.1400	0.9103	0.6918	1.1978	0.5020	367	207	160
CXCR2_n_01m	-0.0853	0.1399	0.9183	0.6980	1.2080	0.5422	367	207	160
EpCAM_c	-0.0009	0.0022	0.9991	0.9947	1.0035	0.6857	368	208	160
EpCAM_c_01	-0.2921	0.1409	0.7467	0.5665	0.9842	0.0382	368	208	160
EpCAM_c_01m	-0.2572	0.1393	0.7732	0.5885	1.0160	0.0649	368	208	160
EpCAM_m	-0.0007	0.0011	0.9993	0.9972	1.0014	0.5094	370	209	161
EpCAM_m_01	-0.0447	0.1657	0.9563	0.6911	1.3232	0.7873	370	209	161
EpCAM_m_01m	-0.1261	0.1385	0.8815	0.6719	1.1565	0.3625	370	209	161
SPP1_c	0.0026	0.0025	1.0026	0.9977	1.0075	0.3010	370	209	161
SPP1_c_01	0.2451	0.2978	1.2777	0.7127	2.2906	0.4106	370	209	161
SPP1_n	0.0014	0.0013	1.0014	0.9989	1.0040	0.2680	369	208	161
SPP1_n_01	0.1060	0.1673	1.1119	0.8009	1.5435	0.5263	369	208	161
SPP1_n_01m	0.1294	0.1390	1.1381	0.8667	1.4946	0.3520	369	208	161

Figure 8. Martingale residual from Cox model with age, gender, histology and stage for recurrence free survival against each marker, Dr. Koo



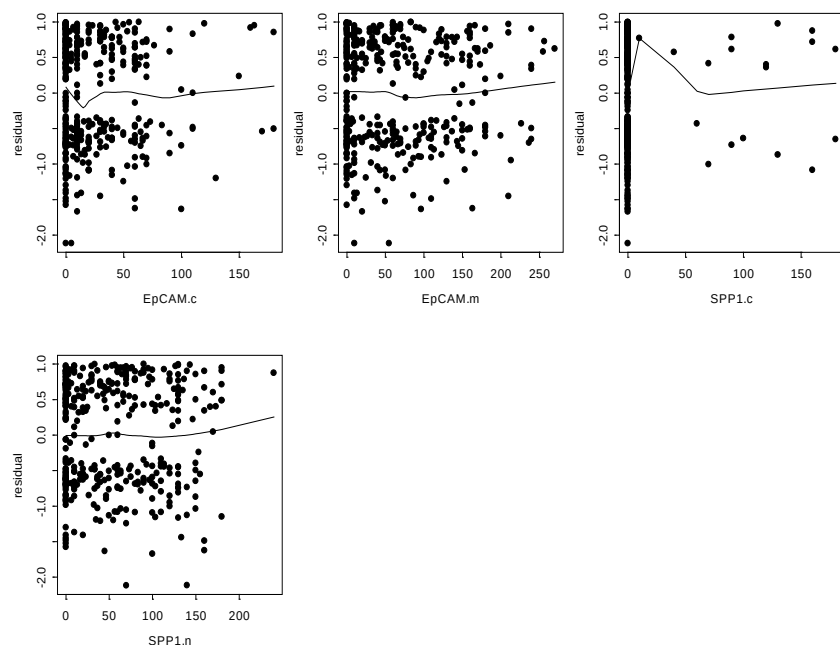


Table 30. Multivariate Cox model assessing effect of each marker independently on recurrence free survival, adjusting for age, gender, histology, stage and neoadjuvant treatment

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
CASK_c	-0.0004	0.0024	0.9996	0.9950	1.0042	0.8583	367	207	160
CASK_c_01	0.1208	0.1573	1.1284	0.8291	1.5359	0.4424	367	207	160
CASK_c_01m	0.0717	0.1545	1.0743	0.7936	1.4543	0.6427	367	207	160
CASK_m	-0.0033	0.0019	0.9967	0.9930	1.0005	0.0876	367	207	160
CASK_m_01	-0.1879	0.1479	0.8287	0.6201	1.1075	0.2041	367	207	160
CD51_c	-0.0008	0.0016	0.9992	0.9961	1.0024	0.6355	367	207	160
CD51_c_01	0.1817	0.1564	1.1993	0.8827	1.6294	0.2452	367	207	160
CD51_c_01m	0.1740	0.1479	1.1900	0.8906	1.5901	0.2395	367	207	160
CD51_m	0.0014	0.0014	1.0014	0.9986	1.0041	0.3254	367	207	160
CD51_m_01	0.3010	0.1620	1.3512	0.9835	1.8563	0.0633	367	207	160
CD51_m_01m	0.3275	0.1527	1.3875	1.0285	1.8717	0.0320	367	207	160
CXCR2_c	0.0017	0.0023	1.0017	0.9973	1.0062	0.4442	367	207	160
CXCR2_c_01	0.0934	0.2152	1.0979	0.7200	1.6739	0.6645	367	207	160
CXCR2_c_01m	0.1735	0.1427	1.1894	0.8993	1.5731	0.2240	367	207	160
CXCR2_n	0.0002	0.0029	1.0002	0.9944	1.0060	0.9528	364	205	159

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
CXCR2_n_01	-0.0670	0.1414	0.9352	0.7089	1.2339	0.6358	364	205	159
CXCR2_n_01m	-0.0622	0.1414	0.9397	0.7123	1.2398	0.6600	364	205	159
EpCAM_c	-0.0003	0.0024	0.9997	0.9951	1.0043	0.8926	365	206	159
EpCAM_c_01	-0.2922	0.1494	0.7466	0.5571	1.0007	0.0505	365	206	159
EpCAM_c_01m	-0.2479	0.1496	0.7804	0.5821	1.0462	0.0974	365	206	159
EpCAM_m	-0.0004	0.0012	0.9996	0.9972	1.0019	0.7085	367	207	160
EpCAM_m_01	-0.0778	0.1836	0.9252	0.6456	1.3258	0.6718	367	207	160
EpCAM_m_01m	-0.1515	0.1504	0.8595	0.6401	1.1541	0.3138	367	207	160
SPP1_c	0.0008	0.0025	1.0008	0.9959	1.0058	0.7373	367	207	160
SPP1_c_01	0.1544	0.3080	1.1670	0.6381	2.1342	0.6161	367	207	160
SPP1_n	0.0004	0.0013	1.0004	0.9978	1.0030	0.7576	366	206	160
SPP1_n_01	0.0637	0.1689	1.0658	0.7653	1.4841	0.7062	366	206	160
SPP1_n_01m	0.0293	0.1431	1.0297	0.7779	1.3631	0.8378	366	206	160

Final multivariate Cox Model assessing the following covariates on recurrence free survival

<i>Analysis of Maximum Likelihood Estimates</i>						
<i>Variable</i>	<i>Parameter Estimate</i>	<i>Standard Error</i>	<i>p-value</i>	<i>Hazard Ratio</i>	<i>95% Hazard Ratio Confidence Limits</i>	
<i>Age</i>	0.0227	0.0074	0.0022	1.023	1.008	1.038
<i>Stage II vs I</i>	0.5978	0.1789	0.0008	1.818	1.280	2.581
<i>III vs I</i>	1.0104	0.1871	<.0001	2.747	1.903	3.964
<i>Neoadjuvant (Yes vs No)</i>	0.3734	0.1865	0.0452	1.453	1.008	2.094
<i>CXCR2_c_01m (>=23.3 vs < 23.3)</i>	0.2999	0.1463	0.0403	1.350	1.013	1.798
<i>EpCAM_c_01 (Pos vs 0)</i>	-0.3722	0.1463	0.0110	0.689	0.517	0.918
<i>CASK_m</i>	-0.0036	0.0018	0.0525	0.996	0.993	1.000

4. Dr. Lotan

4.1 Markers

Table 31. Descriptive summary of markers

<i>covariate</i>	<i>n</i>	<i>mean $\pm$ std, median (min, max)</i>
FEN1_n	366	78.52 $\pm$ 61.68, 70 (0, 220)
MCM2_n	369	109.31 $\pm$ 84.41, 100 (0, 270)
MCM6_n	370	76.7 $\pm$ 75.25, 58.75 (0, 270)
SFN_c	370	32.18 $\pm$ 50.29, 6.67 (0, 250)
SFN_m	370	5.03 $\pm$ 16.98, 0 (0, 130)
TPX2_c	370	50.66 $\pm$ 50.71, 40 (0, 183.33)
TPX2_n	370	6.52 $\pm$ 13.73, 0 (0, 106.67)
UBE2C_c	366	26.28 $\pm$ 24.71, 20 (0, 130)
UBE2C_n	366	25.15 $\pm$ 22.18, 20 (0, 90)

Figure 9. Distribution of markers, Dr. Lotan (Red line – Mean, Green line – Median)

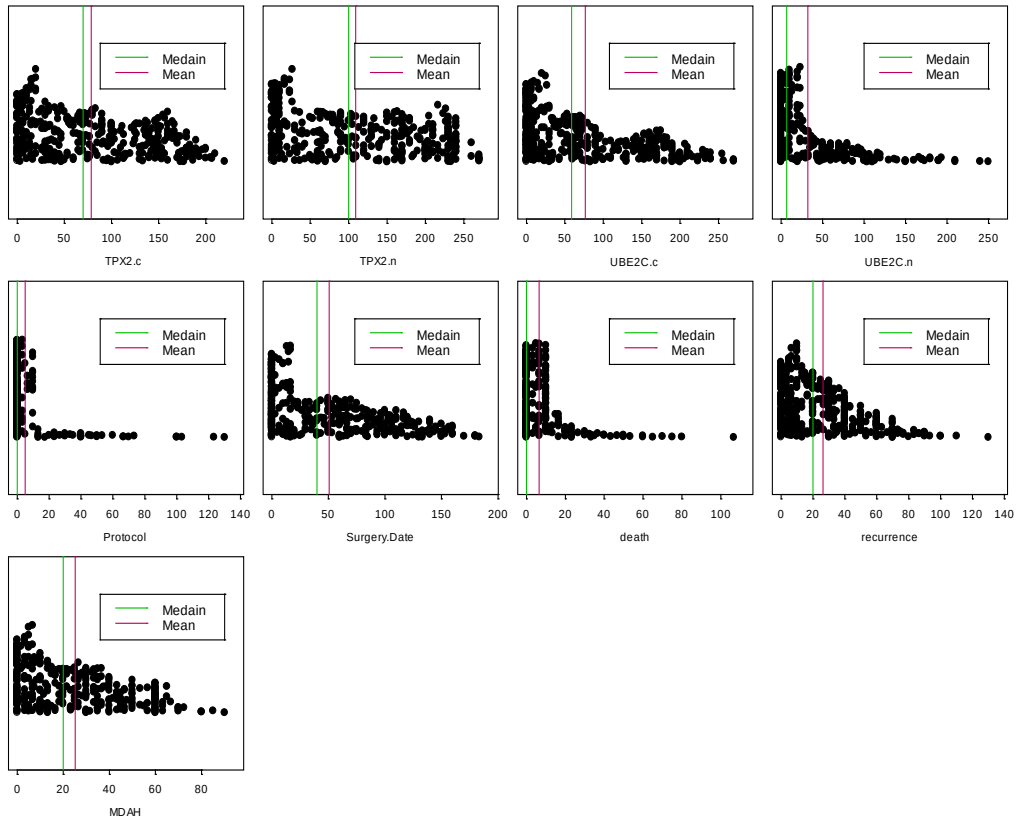


Table 32. Correlation between markers

Spearman Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations									
	FEN1_n	MCM2_n	MCM6_n	SFN_c	SFN_m	TPX2_c	TPX2_n	UBE2C_c	UBE2C_n
FEN1_n	1	0.66287	0.71338	0.30733	0.14547	-0.27785	0.56895	0.64496	0.70561
	<.0001	<.0001	<.0001	<.0001	0.0053	<.0001	<.0001	<.0001	<.0001
	366	365	366	366	366	366	366	362	362
MCM2_n	0.66287	1	0.89878	0.32588	0.25425	-0.41012	0.58737	0.65079	0.65122
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	365	369	369	369	369	369	369	366	366
MCM6_n	0.71338	0.89878	1	0.31998	0.23157	-0.40727	0.59637	0.70428	0.71228
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	366	369	370	370	370	370	370	366	366
SFN_c	0.30733	0.32588	0.31998	1	0.41136	-0.30627	0.36697	0.30834	0.32839
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	366	369	370	370	370	370	370	366	366
SFN_m	0.14547	0.25425	0.23157	0.41136	1	-0.21483	0.20614	0.18474	0.17561
	0.0053	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	0.0004	0.0007
	366	369	370	370	370	370	370	366	366
TPX2_c	-0.27785	-0.41012	-0.40727	-0.30627	-0.21483	1	-0.56411	-0.24807	-0.27633
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	366	369	370	370	370	370	370	366	366
TPX2_n	0.56895	0.58737	0.59637	0.36697	0.20614	-0.56411	1	0.50280	0.54006
	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
	366	369	370	370	370	370	370	366	366
UBE2C_c	0.64496	0.65079	0.70428	0.30834	0.18474	-0.24807	0.50280	1	0.88618
	<.0001	<.0001	<.0001	<.0001	0.0004	<.0001	<.0001	<.0001	<.0001
	362	366	366	366	366	366	366	366	366
UBE2C_n	0.70561	0.65122	0.71228	0.32839	0.17561	-0.27633	0.54006	0.88618	1
	<.0001	<.0001	<.0001	<.0001	0.0007	<.0001	<.0001	<.0001	<.0001
	362	366	366	366	366	366	366	366	366

Figure 10. Correlation between markers, Dr. Lotan

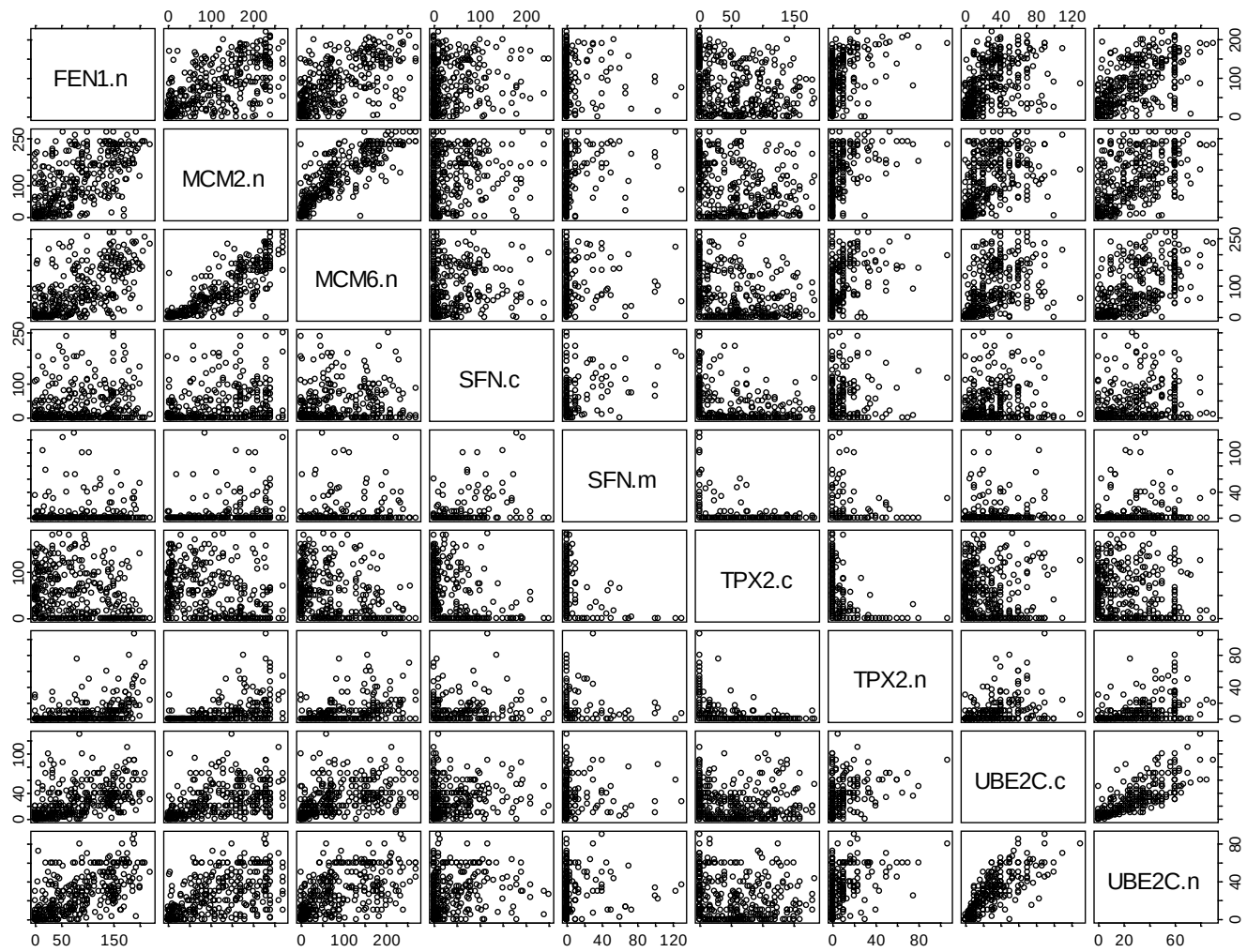


Table 33. Frequency table for dichotomized markers

<i>covariate</i>	<i>levels</i>	<i>N (%)</i>	<i>covariate</i>	<i>levels</i>	<i>N (%)</i>
FEN1_n_01	missing	4	FEN1_n_01m	missing	4
	0	30(8.2%)		0	178(48.6%)
	1	336(91.8%)		1	188(51.4%)
MCM2_n_01	missing	1	MCM2_n_01m	missing	1
	0	27(7.3%)		0	181(49.1%)
	1	342(92.7%)		1	188(50.9%)
MCM6_n_01	0	60(16.2%)	MCM6_n_01m	0	185(50%)
	1	310(83.8%)		1	185(50%)
SFN_c_01	0	162(43.8%)	SFN_c_01m	0	176(47.6%)
	1	208(56.2%)		1	194(52.4%)
SFN_m_01	0	308(83.2%)	TPX2_c_01m	0	184(49.7%)
	1	62(16.8%)		1	186(50.3%)
TPX2_c_01	0	108(29.2%)	UBE2C_c_01m	missing	4
	1	262(70.8%)		0	169(46.2%)
TPX2_n_01	0	238(64.3%)		1	197(53.8%)
	1	132(35.7%)	UBE2C_n_01m	missing	4
UBE2C_c_01	missing	4		0	172(47%)
	0	59(16.1%)		1	194(53%)
	1	307(83.9%)			
UBE2C_n_01	missing	4			
	0	68(18.6%)			
	1	298(81.4%)			

Table 34. Markers by gender

<i>covariate</i>	<i>Gender</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
FEN1_n	F	181	72.4 ± 60.86, 60 (0, 206.67)	.0560
	M	185	84.51 ± 62.05, 77.5 (0, 220)	
MCM2_n	F	184	95.99 ± 83.46, 78.33 (0, 270)	.0025
	M	185	122.56 ± 83.47, 120 (0, 270)	
MCM6_n	F	184	68.01 ± 74.01, 35.83 (0, 270)	.0132

<i>covariate</i>	<i>Gender</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
SFN_c	M	186	85.3 ± 75.68, 66.67 (0, 270)	.3176
	F	184	33.32 ± 55.9, 3.33 (0, 250)	
SFN_m	M	186	31.05 ± 44.18, 8.75 (0, 193.33)	.0677
	F	184	4.35 ± 16.43, 0 (0, 130)	
TPX2_c	M	186	5.7 ± 17.54, 0 (0, 123.33)	.4051
	F	184	52.42 ± 50.83, 40 (0, 183.33)	
TPX2_n	M	186	48.91 ± 50.66, 31.67 (0, 160)	.0365
	F	184	5.99 ± 14.38, 0 (0, 106.67)	
UBE2C_c	M	186	7.06 ± 13.07, 0 (0, 75)	.0051
	F	181	23.18 ± 24, 16.67 (0, 110)	
UBE2C_n	M	185	29.32 ± 25.07, 23.33 (0, 130)	.0234
	F	181	22.51 ± 21.25, 16.67 (0, 80)	
	M	185	27.74 ± 22.82, 25 (0, 90)	

Table 35. Markers by race

<i>covariate</i>	<i>Race</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
FEN1_n	Black	21	86.23 ± 70.82, 80 (0, 190)	.0336
	Hispanic	14	66.79 ± 56.39, 52.5 (1.67, 170)	
	Oriental	5	7.67 ± 7.23, 10 (0, 15)	
	White	326	79.62 ± 61.22, 72.5 (0, 220)	
MCM2_n	Black	21	141.27 ± 91.86, 170 (0, 260)	.1017
	Hispanic	14	73.33 ± 66.42, 60 (5, 180)	
	Oriental	5	61.33 ± 61.67, 63.33 (0, 160)	
	White	329	109.53 ± 84.27, 100 (0, 270)	
MCM6_n	Black	21	99.56 ± 84.17, 90 (0, 230)	.1870
	Hispanic	14	42.74 ± 60.6, 11.67 (0, 183.33)	
	Oriental	5	33.33 ± 39.3, 16.67 (3.33, 100)	
	White	330	77.34 ± 75.14, 60 (0, 270)	
SFN_c	Black	21	26.87 ± 52.68, 6.67 (0, 180)	.7783
	Hispanic	14	43.69 ± 81.51, 2.5 (0, 240)	
	Oriental	5	30.67 ± 66.72, 0 (0, 150)	

<i>covariate</i>	<i>Race</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
SFN_m	White	330	32.05 ± 48.42, 6.67 (0, 250)	.7506
	Black	21	3.53 ± 7.92, 0 (0, 26.67)	
	Hispanic	14	0.71 ± 2.67, 0 (0, 10)	
	Oriental	5	20.67 ± 46.21, 0 (0, 103.33)	
	White	330	5.07 ± 17, 0 (0, 130)	
TPX2_c	Black	21	51.11 ± 56.91, 30 (0, 153.33)	.2146
	Hispanic	14	78.57 ± 55.51, 84.17 (0, 160)	
	Oriental	5	35.33 ± 60.17, 3.33 (0, 140)	
	White	330	49.67 ± 49.83, 36.67 (0, 183.33)	
	Black	21	7.62 ± 12.7, 0 (0, 50)	.3958
TPX2_n	Hispanic	14	4.29 ± 13.42, 0 (0, 50)	
	Oriental	5	2.67 ± 5.96, 0 (0, 13.33)	
	White	330	6.61 ± 13.92, 0 (0, 106.67)	
	Black	21	33.65 ± 30.33, 20 (0, 110)	.2798
UBE2C_c	Hispanic	14	16.43 ± 18.36, 10 (0, 70)	
	Oriental	5	22.67 ± 34.19, 10 (0, 83.33)	
	White	326	26.29 ± 24.35, 20 (0, 130)	
	Black	21	29.05 ± 25.67, 20 (0, 80)	.0917
UBE2C_n	Hispanic	14	13.69 ± 19.03, 10 (0, 70)	
	Oriental	5	9.67 ± 7.85, 10 (0, 21.67)	
	White	326	25.63 ± 22.05, 23.33 (0, 90)	

Table 36. Markers by Smoking Status

<i>covariate</i>	<i>smoker</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
FEN1_n	1 Never	38	34.71 ± 41.77, 15 (0, 165)	<.0001
	2 Former	168	81.88 ± 62.39, 71.67 (0, 220)	
	3 Current	160	85.4 ± 61, 80 (0, 210)	
MCM2_n	1 Never	38	39.04 ± 54.17, 13.33 (0, 240)	<.0001
	2 Former	170	111.72 ± 85.56, 98.33 (0, 270)	
	3 Current	161	123.36 ± 81.22, 126.67 (0, 270)	

<i>covariate</i>	<i>smoker</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
MCM6_n	1 Never	38	22.5 ± 39.43, 5.83 (0, 190)	<.0001
	2 Former	170	78.58 ± 78.6, 55.83 (0, 270)	
	3 Current	162	87.44 ± 72.96, 73.33 (0, 270)	
SFN_c	1 Never	38	6.54 ± 12.49, 0 (0, 53.33)	.0025
	2 Former	170	33.24 ± 52.15, 6.67 (0, 240)	
	3 Current	162	37.08 ± 52.21, 10 (0, 250)	
SFN_m	1 Never	38	1.32 ± 8.11, 0 (0, 50)	.0264
	2 Former	170	4.72 ± 17.44, 0 (0, 130)	
	3 Current	162	6.23 ± 17.94, 0 (0, 103.33)	
TPX2_c	1 Never	38	68.05 ± 46.74, 60 (0, 180)	.0200
	2 Former	170	50.98 ± 51.38, 38.33 (0, 180)	
	3 Current	162	46.23 ± 50.28, 30 (0, 183.33)	
TPX2_n	1 Never	38	1.23 ± 5.17, 0 (0, 30)	.0005
	2 Former	170	6.09 ± 12.3, 0 (0, 80)	
	3 Current	162	8.23 ± 16.05, 0 (0, 106.67)	
UBE2C_c	1 Never	38	10.5 ± 19.71, 3.33 (0, 100)	<.0001
	2 Former	169	26.5 ± 23.54, 23.33 (0, 110)	
	3 Current	159	29.83 ± 25.64, 20 (0, 130)	
UBE2C_n	1 Never	38	7.52 ± 12.04, 1.67 (0, 60)	<.0001
	2 Former	169	25.89 ± 22.08, 25 (0, 90)	
	3 Current	159	28.59 ± 22.3, 25 (0, 85)	

Table 37. Markers by histology

<i>covariate</i>	<i>histology0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
FEN1_n	ADENO	226	57.03 ± 53.67, 40 (0, 206.67)	<.0001
	Other	16	95.1 ± 71.14, 78.33 (0, 210)	
	SCC	124	115.56 ± 55.99, 130 (0, 220)	
MCM2_n	ADENO	226	71.45 ± 73.27, 50 (0, 270)	<.0001
	Other	17	147.84 ± 76.24, 150 (10, 270)	
	SCC	126	172.03 ± 61.7, 170 (0, 270)	

<i>covariate</i>	<i>histology0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
MCM6_n	ADENO	227	45.1 ± 59.56, 15 (0, 256.67)	<.0001
	Other	17	107.55 ± 82.29, 90 (0, 270)	
	SCC	126	129.46 ± 68.47, 148.33 (0, 270)	
SFN_c	ADENO	227	10.42 ± 22.83, 0 (0, 173.33)	<.0001
	Other	17	14.51 ± 28.11, 3.33 (0, 106.67)	
	SCC	126	73.76 ± 61.47, 68.33 (0, 250)	
SFN_m	ADENO	227	0.99 ± 6.16, 0 (0, 66.67)	<.0001
	Other	17	1.76 ± 5.67, 0 (0, 23.33)	
	SCC	126	12.74 ± 26.23, 0 (0, 130)	
TPX2_c	ADENO	227	70.28 ± 48.08, 65 (0, 180)	<.0001
	Other	17	68.43 ± 60.74, 65 (0, 150)	
	SCC	126	12.9 ± 27.53, 0 (0, 183.33)	
TPX2_n	ADENO	227	1.66 ± 5.65, 0 (0, 46.67)	<.0001
	Other	17	11.57 ± 23.55, 0 (0, 75)	
	SCC	126	14.61 ± 17.77, 10 (0, 106.67)	
UBE2C_c	ADENO	225	19.49 ± 23.97, 10 (0, 130)	<.0001
	Other	16	39.9 ± 22.33, 40 (6.67, 83.33)	
	SCC	125	36.78 ± 21.9, 33.33 (0, 110)	
UBE2C_n	ADENO	225	17.33 ± 19.78, 10 (0, 90)	<.0001
	Other	16	37.86 ± 19.5, 36.67 (3.33, 63.33)	
	SCC	125	37.61 ± 20.18, 36.67 (0, 85)	

Table 38. Markers by stage

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
FEN1_n	IA	102	66.34 ± 59.11, 58.33 (0, 200)	.0208
	IB	129	75.5 ± 63.02, 65 (0, 210)	
	IIA	22	86.97 ± 66.3, 85 (0, 186.67)	
	IIB	52	95.14 ± 58.5, 90 (0, 220)	
	IIIA	61	88.09 ± 60.92, 80 (0, 205)	
MCM2_n	IA	103	90.29 ± 82.81, 70 (0, 270)	.0058
	IB	131	104.75 ± 85.62, 90 (0, 270)	

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
MCM6_n	IIA	21	122.46 ± 89.83, 120 (0, 240)	.0032
	IIB	53	133.9 ± 82.88, 140 (0, 270)	
	IIIA	61	125.34 ± 77.4, 130 (3.33, 260)	
	IA	103	57.38 ± 66, 35 (0, 240)	
	IB	131	75.46 ± 78.54, 43.33 (0, 270)	
	IIA	22	91.89 ± 81.72, 68.33 (0, 240)	
	IIB	53	101.89 ± 82.82, 73.33 (0, 270)	
	IIIA	61	84.6 ± 66.49, 70 (0, 213.33)	
SFN_c	IA	103	33.37 ± 51.43, 10 (0, 250)	.8564
	IB	131	30.44 ± 48.49, 5 (0, 210)	
	IIA	22	26.29 ± 30.98, 16.67 (0, 96.67)	
	IIB	53	42.52 ± 65.63, 6.67 (0, 240)	
	IIIA	61	27.05 ± 41.86, 6.67 (0, 180)	
SFN_m	IA	103	4.56 ± 15.7, 0 (0, 100)	.5230
	IB	131	4.97 ± 15.9, 0 (0, 130)	
	IIA	22	0.76 ± 2.51, 0 (0, 10)	
	IIB	53	7.83 ± 21.96, 0 (0, 123.33)	
	IIIA	61	5.05 ± 19.32, 0 (0, 103.33)	
TPX2_c	IA	103	57.9 ± 52.24, 45 (0, 180)	.1717
	IB	131	51.83 ± 50.67, 40 (0, 180)	
	IIA	22	48.64 ± 54.18, 25 (0, 150)	
	IIB	53	44.53 ± 48.24, 33.33 (0, 160)	
	IIIA	61	41.97 ± 48.59, 16.67 (0, 183.33)	
TPX2_n	IA	103	4.42 ± 10.95, 0 (0, 75)	.0782
	IB	131	6.83 ± 16.39, 0 (0, 106.67)	
	IIA	22	9.09 ± 15.47, 0 (0, 60)	
	IIB	53	8.08 ± 12.21, 0 (0, 50)	
	IIIA	61	7.14 ± 12.2, 0 (0, 65)	
UBE2C_c	IA	102	19.84 ± 20.09, 15 (0, 86.67)	.0026
	IB	129	25.67 ± 25.82, 16.67 (0, 100)	
	IIA	21	25.24 ± 19.65, 20 (0, 60)	
	IIB	53	33.74 ± 28.13, 30 (0, 130)	

<i>covariate</i>	<i>PathStage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	IIIA	61	32.24 ± 25.3, 30 (0, 110)	
UBE2C_n	IA	102	20.85 ± 20.14, 15.83 (0, 72.5)	.0190
	IB	129	23.71 ± 22.46, 15 (0, 80)	
	IIA	21	27.82 ± 24.03, 25 (0, 85)	
	IIB	53	31.76 ± 24.16, 26.67 (0, 90)	
	IIIA	61	28.76 ± 21.12, 25 (0, 70)	

<i>covariate</i>	<i>stage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
FEN1_n	I	231	71.45 ± 61.36, 60 (0, 210)	.0076
	II	74	92.71 ± 60.58, 90 (0, 220)	
	IIIA	61	88.09 ± 60.92, 80 (0, 205)	
MCM2_n	I	234	98.38 ± 84.52, 80 (0, 270)	.0024
	II	74	130.65 ± 84.45, 135 (0, 270)	
	IIIA	61	125.34 ± 77.4, 130 (3.33, 260)	
MCM6_n	I	234	67.5 ± 73.69, 38.33 (0, 270)	.0021
	II	75	98.96 ± 82.07, 73.33 (0, 270)	
	IIIA	61	84.6 ± 66.49, 70 (0, 213.33)	
SFN_c	I	234	31.73 ± 49.72, 6.67 (0, 250)	.5994
	II	75	37.76 ± 57.92, 6.67 (0, 240)	
	IIIA	61	27.05 ± 41.86, 6.67 (0, 180)	
SFN_m	I	234	4.79 ± 15.78, 0 (0, 130)	.8210
	II	75	5.76 ± 18.74, 0 (0, 123.33)	
	IIIA	61	5.05 ± 19.32, 0 (0, 103.33)	
TPX2_c	I	234	54.5 ± 51.35, 41.67 (0, 180)	.0680
	II	75	45.73 ± 49.72, 33.33 (0, 160)	
	IIIA	61	41.97 ± 48.59, 16.67 (0, 183.33)	
TPX2_n	I	234	5.77 ± 14.27, 0 (0, 106.67)	.0202
	II	75	8.38 ± 13.15, 0 (0, 60)	
	IIIA	61	7.14 ± 12.2, 0 (0, 65)	
UBE2C_c	I	231	23.1 ± 23.6, 16.67 (0, 100)	.0013
	II	74	31.33 ± 26.16, 25 (0, 130)	

<i>covariate</i>	<i>stage</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	IIIA	61	32.24 ± 25.3, 30 (0, 110)	
UBE2C_n	I	231	22.45 ± 21.47, 15 (0, 80)	.0050
	II	74	30.64 ± 24.03, 26.67 (0, 90)	
	IIIA	61	28.76 ± 21.12, 25 (0, 70)	

Table 39. Markers by Grade

<i>covariate</i>	<i>grade0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
FEN1_n	1 Poorly	122	101.28 ± 63.58, 100 (0, 220)	<.0001
	2 Moderately	196	74.77 ± 56.97, 66.67 (0, 196.67)	
	3 Well	36	23.19 ± 36.48, 6.67 (0, 155)	
MCM2_n	1 Poorly	122	146.05 ± 77.26, 158.33 (0, 270)	<.0001
	2 Moderately	198	101.06 ± 82.37, 90 (0, 270)	
	3 Well	36	24.12 ± 40.09, 8.33 (0, 160)	
MCM6_n	1 Poorly	122	104.62 ± 74.16, 92.5 (0, 270)	<.0001
	2 Moderately	199	70.86 ± 74.05, 50 (0, 270)	
	3 Well	36	11.48 ± 25.94, 0 (0, 126.67)	
SFN_c	1 Poorly	122	28.19 ± 47.01, 5.83 (0, 250)	<.0001
	2 Moderately	199	40.83 ± 55.23, 10 (0, 240)	
	3 Well	36	7.27 ± 18.96, 0 (0, 96.67)	
SFN_m	1 Poorly	122	4.06 ± 13.38, 0 (0, 100)	.0456
	2 Moderately	199	6.81 ± 20.46, 0 (0, 130)	
	3 Well	36	0.09 ± 0.56, 0 (0, 3.33)	
TPX2_c	1 Poorly	122	48.33 ± 50.81, 34.17 (0, 183.33)	.0517
	2 Moderately	199	47.76 ± 49.6, 30 (0, 180)	
	3 Well	36	66.25 ± 49.88, 60 (0, 180)	
TPX2_n	1 Poorly	122	10.3 ± 17.33, 0 (0, 106.67)	<.0001
	2 Moderately	199	5.13 ± 11.02, 0 (0, 80)	
	3 Well	36	0.83 ± 5, 0 (0, 30)	
UBE2C_c	1 Poorly	122	36.3 ± 27.76, 30.83 (0, 130)	<.0001
	2 Moderately	196	23.34 ± 21.29, 20 (0, 86.67)	

<i>covariate</i>	<i>grade0</i>	<i>n</i>	<i>mean ± std, median (min, max)</i>	<i>pValue</i>
	3 Well	36	4.21 ± 7.85, 0 (0, 36.67)	
UBE2C_n	1 Poorly	122	33.22 ± 22.42, 31.67 (0, 90)	<.0001
	2 Moderately	196	23.51 ± 21.25, 17.5 (0, 85)	
	3 Well	36	4.12 ± 7.35, 0 (0, 26.67)	

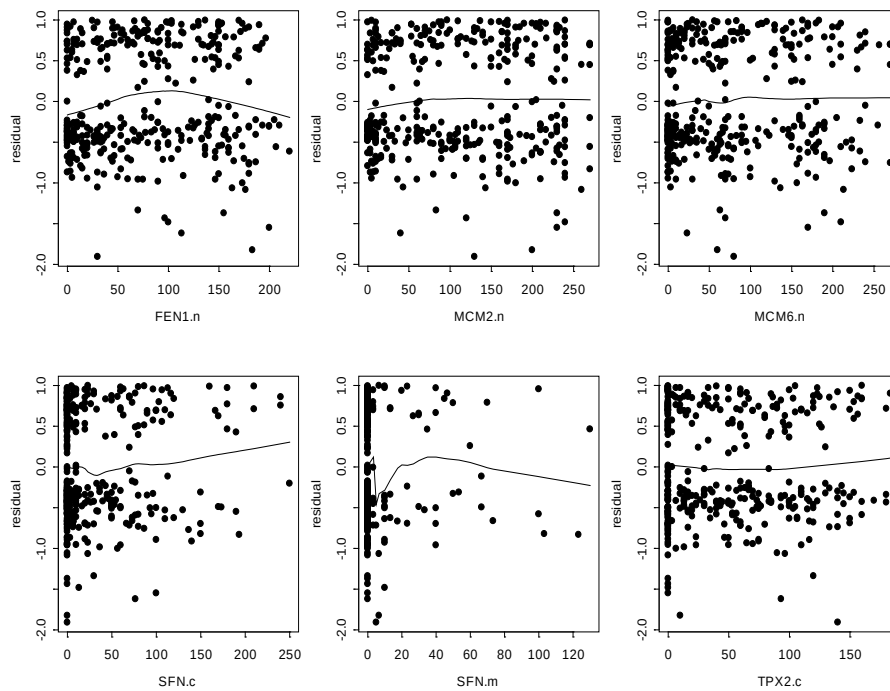
4.2 Overall survival

Table 40. Univariate Cox model assessing effect of covariates on overall survival

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
FEN1_n	0.0028	0.0012	1.0028	1.0003	1.0052	0.0258	366	158	208
FEN1_n_01	0.2005	0.3128	1.2220	0.6620	2.2559	0.5215	366	158	208
FEN1_n_01m	0.5043	0.1624	1.6558	1.2043	2.2766	0.0019	366	158	208
MCM2_n	0.0019	0.0009	1.0019	1.0001	1.0037	0.0425	369	159	210
MCM2_n_01	0.4957	0.3632	1.6416	0.8056	3.3449	0.1723	369	159	210
MCM2_n_01m	0.1819	0.1596	1.1995	0.8773	1.6398	0.2544	369	159	210
MCM6_n	0.0018	0.0010	1.0018	0.9999	1.0038	0.0694	370	160	210
MCM6_n_01	0.6533	0.2637	1.9218	1.1463	3.2222	0.0132	370	160	210
MCM6_n_01m	0.1858	0.1588	1.2042	0.8822	1.6438	0.2418	370	160	210
SFN_c	0.0022	0.0014	1.0022	0.9993	1.0050	0.1360	370	160	210
SFN_c_01	0.0021	0.1590	1.0021	0.7338	1.3685	0.9895	370	160	210
SFN_c_01m	0.0609	0.1584	1.0628	0.7792	1.4497	0.7005	370	160	210
SFN_m	0.0017	0.0041	1.0017	0.9936	1.0099	0.6806	370	160	210
SFN_m_01	-0.0772	0.2150	0.9257	0.6074	1.4108	0.7196	370	160	210
TPX2_c	-0.0012	0.0016	0.9988	0.9956	1.0019	0.4455	370	160	210
TPX2_c_01	-0.1738	0.1710	0.8405	0.6012	1.1750	0.3093	370	160	210
TPX2_c_01m	-0.1873	0.1585	0.8292	0.6077	1.1314	0.2375	370	160	210
TPX2_n	0.0082	0.0050	1.0082	0.9983	1.0182	0.1035	370	160	210
TPX2_n_01	0.2700	0.1620	1.3099	0.9536	1.7993	0.0956	370	160	210
UBE2C_c	0.0069	0.0029	1.0069	1.0012	1.0127	0.0179	366	159	207
UBE2C_c_01	0.5470	0.2569	1.7280	1.0444	2.8592	0.0333	366	159	207
UBE2C_c_01m	0.3326	0.1620	1.3946	1.0152	1.9158	0.0401	366	159	207

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
UBE2C_n	0.0065	0.0035	1.0065	0.9997	1.0134	0.0621	366	159	207
UBE2C_n_01	0.5520	0.2392	1.7367	1.0866	2.7757	0.0210	366	159	207
UBE2C_n_01m	0.2177	0.1604	1.2432	0.9078	1.7025	0.1747	366	159	207

Figure 11. Martingale residual from Cox model with age, gender, histology and stage for overall survival against each marker, Dr. Lotan



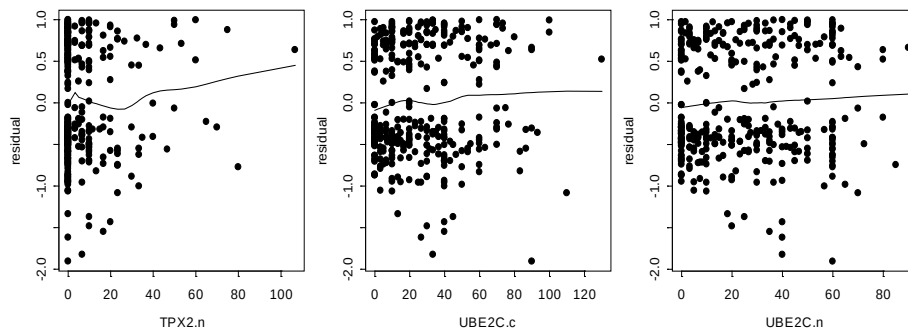


Table 41. Multicovariate Cox model assessing effect of each marker independently on overall survival , adjusting for age, gender, histology, stage and neoadjuvant treatment

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
FEN1_n	0.0018	0.0014	1.0018	0.9990	1.0046	0.2090	363	156	207
FEN1_n_01	-0.0004	0.3205	0.9996	0.5333	1.8735	0.9990	363	156	207
FEN1_n_01m	0.3721	0.1772	1.4508	1.0251	2.0531	0.0357	363	156	207
MCM2_n	0.0015	0.0012	1.0015	0.9992	1.0038	0.1963	366	157	209
MCM2_n_01	0.3662	0.3747	1.4423	0.6921	3.0058	0.3283	366	157	209
MCM2_n_01m	0.0465	0.2021	1.0476	0.7049	1.5567	0.8182	366	157	209
MCM6_n	0.0017	0.0013	1.0017	0.9992	1.0042	0.1788	367	158	209
MCM6_n_01	0.5623	0.2759	1.7546	1.0218	3.0131	0.0415	367	158	209
MCM6_n_01m	0.0371	0.1913	1.0378	0.7133	1.5099	0.8462	367	158	209
SFN_c	0.0028	0.0018	1.0028	0.9992	1.0064	0.1302	367	158	209
SFN_c_01	0.0090	0.1855	1.0090	0.7015	1.4515	0.9613	367	158	209

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
SFN_c_01m	0.0720	0.1901	1.0747	0.7404	1.5598	0.7048	367	158	209
SFN_m	-0.0001	0.0044	0.9999	0.9912	1.0086	0.9746	367	158	209
SFN_m_01	-0.2200	0.2284	0.8025	0.5129	1.2558	0.3356	367	158	209
TPX2_c	-0.0005	0.0020	0.9995	0.9956	1.0035	0.8203	367	158	209
TPX2_c_01	-0.1445	0.2249	0.8654	0.5570	1.3447	0.5204	367	158	209
TPX2_c_01m	-0.1643	0.1950	0.8485	0.5790	1.2434	0.3993	367	158	209
TPX2_n	0.0090	0.0059	1.0090	0.9974	1.0208	0.1283	367	158	209
TPX2_n_01	0.1707	0.2001	1.1862	0.8014	1.7557	0.3935	367	158	209
UBE2C_c	0.0053	0.0033	1.0053	0.9989	1.0118	0.1040	363	157	206
UBE2C_c_01	0.4351	0.2715	1.5452	0.9075	2.6310	0.1090	363	157	206
UBE2C_c_01m	0.2086	0.1824	1.2319	0.8615	1.7615	0.2530	363	157	206
UBE2C_n	0.0050	0.0039	1.0050	0.9973	1.0128	0.2024	363	157	206
UBE2C_n_01	0.4580	0.2532	1.5809	0.9625	2.5966	0.0705	363	157	206
UBE2C_n_01m	0.1223	0.1843	1.1301	0.7875	1.6217	0.5069	363	157	206

Final multivariate Cox Model assessing the following covariates on overall survival

<i>Analysis of Maximum Likelihood Estimates</i>						
<i>Variable</i>	<i>Parameter Estimate</i>	<i>Standard Error</i>	<i>p-value</i>	<i>Hazard Ratio</i>	<i>95% Hazard Ratio Confidence Limits</i>	
<i>Age</i>	0.0276	0.0088	0.0017	1.028	1.010	1.046
<i>Stage II vs I</i>	0.3833	0.2066	0.0635	1.467	0.979	2.199
<i>III vs I</i>	0.7788	0.2141	0.0003	2.179	1.432	3.315
<i>Neoadjuvant (Yes vs No)</i>	0.3947	0.2141	0.0652	1.484	0.975	2.258
<i>FEN1_n_01m (>=70 vs <70)</i>	0.2401	0.1775	0.1761	1.271	0.898	1.800
<i>MCM6_n_01 (Pos vs 0)</i>	0.4839	0.2925	0.0981	1.622	0.914	2.878

4.3 Recurrence free survival

Table 42. Univariate Cox model assessing effect of covariates on recurrence free survival

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
FEN1_n	0.0017	0.0011	1.0017	0.9996	1.0039	0.1127	366	206	160
FEN1_n_01	-0.1908	0.2472	0.8263	0.5090	1.3413	0.4402	366	206	160
FEN1_n_01m	0.3148	0.1405	1.3700	1.0402	1.8043	0.0251	366	206	160
MCM2_n	0.0015	0.0008	1.0015	0.9999	1.0031	0.0643	369	208	161
MCM2_n_01	0.4504	0.2978	1.5690	0.8753	2.8123	0.1304	369	208	161
MCM2_n_01m	0.1715	0.1393	1.1871	0.9034	1.5599	0.2183	369	208	161
MCM6_n	0.0013	0.0009	1.0013	0.9995	1.0030	0.1522	370	209	161
MCM6_n_01	0.4432	0.2097	1.5577	1.0327	2.3497	0.0346	370	209	161
MCM6_n_01m	0.1271	0.1386	1.1355	0.8655	1.4899	0.3590	370	209	161
SFN_c	0.0013	0.0013	1.0013	0.9987	1.0039	0.3183	370	209	161
SFN_c_01	0.0444	0.1394	1.0454	0.7954	1.3739	0.7503	370	209	161
SFN_c_01m	0.0646	0.1386	1.0667	0.8129	1.3997	0.6414	370	209	161
SFN_m	0.0029	0.0035	1.0029	0.9960	1.0099	0.4045	370	209	161
SFN_m_01	0.0384	0.1833	1.0392	0.7256	1.4883	0.8340	370	209	161
TPX2_c	-0.0013	0.0014	0.9987	0.9959	1.0014	0.3503	370	209	161
TPX2_c_01	-0.1557	0.1496	0.8558	0.6384	1.1473	0.2979	370	209	161
TPX2_c_01m	-0.0882	0.1386	0.9155	0.6978	1.2013	0.5243	370	209	161
TPX2_n	0.0060	0.0044	1.0060	0.9973	1.0147	0.1761	370	209	161
TPX2_n_01	0.1739	0.1432	1.1899	0.8988	1.5754	0.2245	370	209	161
UBE2C_c	0.0059	0.0027	1.0059	1.0006	1.0113	0.0301	366	208	158
UBE2C_c_01	0.3728	0.2065	1.4518	0.9686	2.1761	0.0710	366	208	158
UBE2C_c_01m	0.1966	0.1401	1.2172	0.9249	1.6019	0.1606	366	208	158
UBE2C_n	0.0047	0.0031	1.0047	0.9987	1.0108	0.1242	366	208	158
UBE2C_n_01	0.4114	0.1948	1.5089	1.0300	2.2105	0.0347	366	208	158
UBE2C_n_01m	0.1077	0.1396	1.1137	0.8471	1.4643	0.4404	366	208	158

Figure 12. Martingale residual from Cox model with age, gender, histology and stage for recurrence free survival against each marker, Dr. Lotan

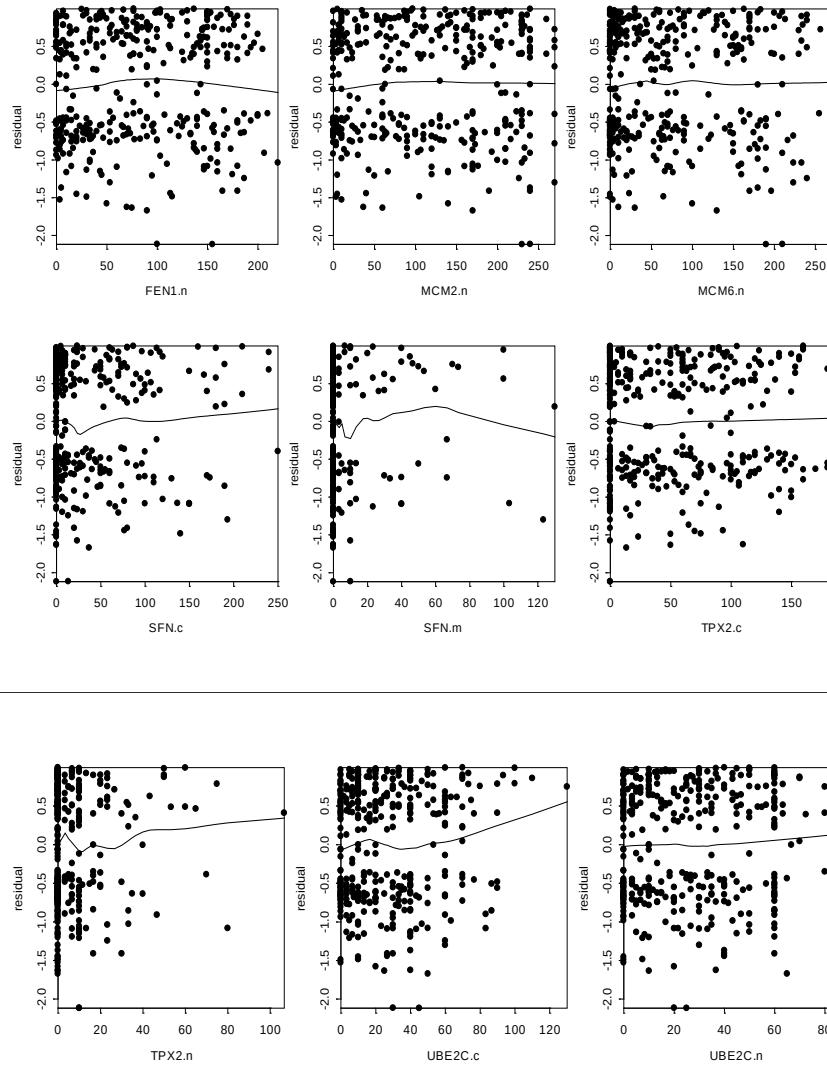


Table 43. Multicovariate Cox model assessing effect of each marker independently on recurrence free survival, adjusting for age, gender, histology, stage and neoadjuvant treatment

<i>covariate</i>	<i>Estimate</i>	<i>StdErr</i>	<i>HazardRatio</i>	<i>HRLowerCL</i>	<i>HRUpperCL</i>	<i>pValue</i>	<i>Total</i>	<i>Event</i>	<i>Censored</i>
FEN1_n	0.0009	0.0013	1.0009	0.9984	1.0034	0.4787	363	204	159
FEN1_n_01	-0.3987	0.2550	0.6712	0.4072	1.1063	0.1179	363	204	159
FEN1_n_01m	0.1649	0.1521	1.1793	0.8753	1.5889	0.2782	363	204	159
MCM2_n	0.0010	0.0010	1.0010	0.9990	1.0030	0.3229	366	206	160
MCM2_n_01	0.3055	0.3087	1.3573	0.7411	2.4858	0.3223	366	206	160
MCM2_n_01m	0.0470	0.1753	1.0482	0.7433	1.4780	0.7885	366	206	160
MCM6_n	0.0008	0.0011	1.0008	0.9986	1.0030	0.4578	367	207	160
MCM6_n_01	0.3322	0.2217	1.3941	0.9027	2.1529	0.1340	367	207	160
MCM6_n_01m	-0.0318	0.1636	0.9687	0.7029	1.3349	0.8458	367	207	160
SFN_c	0.0014	0.0016	1.0014	0.9983	1.0045	0.3689	367	207	160
SFN_c_01	0.0425	0.1609	1.0434	0.7611	1.4303	0.7919	367	207	160
SFN_c_01m	0.0550	0.1637	1.0566	0.7666	1.4563	0.7368	367	207	160
SFN_m	0.0018	0.0037	1.0018	0.9946	1.0090	0.6284	367	207	160
SFN_m_01	-0.0477	0.1966	0.9534	0.6486	1.4015	0.8082	367	207	160
TPX2_c	-0.0008	0.0016	0.9992	0.9960	1.0024	0.6164	367	207	160
TPX2_c_01	-0.0979	0.1901	0.9067	0.6247	1.3161	0.6065	367	207	160
TPX2_c_01m	-0.0733	0.1634	0.9293	0.6746	1.2801	0.6536	367	207	160
TPX2_n	0.0063	0.0051	1.0063	0.9962	1.0165	0.2202	367	207	160
TPX2_n_01	0.0646	0.1728	1.0667	0.7602	1.4968	0.7086	367	207	160
UBE2C_c	0.0041	0.0031	1.0041	0.9981	1.0101	0.1826	363	206	157
UBE2C_c_01	0.2553	0.2202	1.2908	0.8383	1.9875	0.2464	363	206	157
UBE2C_c_01m	0.0329	0.1568	1.0335	0.7601	1.4053	0.8335	363	206	157
UBE2C_n	0.0027	0.0035	1.0027	0.9960	1.0095	0.4296	363	206	157
UBE2C_n_01	0.3079	0.2082	1.3606	0.9048	2.0461	0.1391	363	206	157
UBE2C_n_01m	-0.0249	0.1602	0.9754	0.7126	1.3351	0.8763	363	206	157

5. ALL Markers

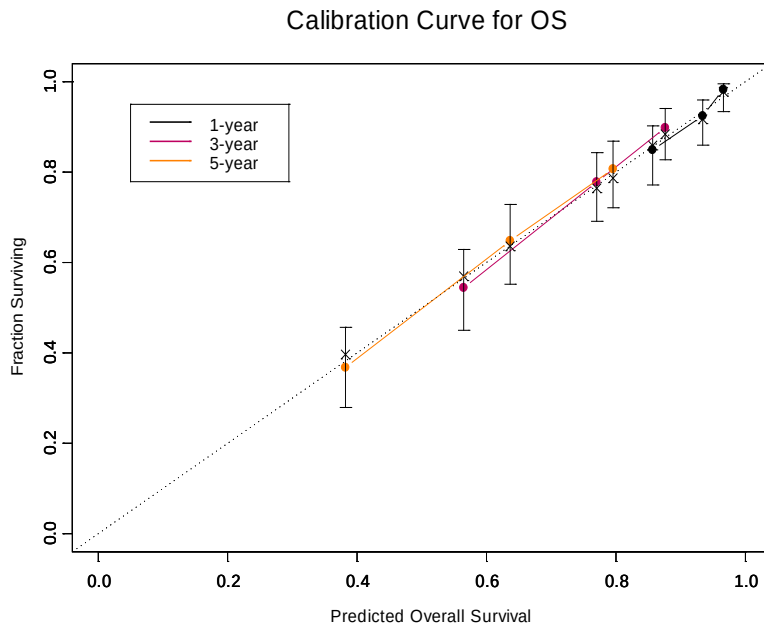
Multivariate Cox Model assessing the following covariates on Overall Survival

Analysis of Maximum Likelihood Estimates						
Variable	Parameter Estimate	Standard Error	p-value	HR	95% HR CL	
Age	0.0327	0.0089	0.0002	1.033	1.015 1.051	
Histology Adeno vs SCC	0.3998	0.2054	0.0516	1.491	0.997 2.231	
Other vs SCC	-0.1297	0.4799	0.7869	0.878	0.343 2.250	
Stage II vs I	0.3586	0.2081	0.0849	1.431	0.952 2.152	
III vs I	0.9228	0.2168	<.0001	2.516	1.645 3.849	
Neoadjuvant Yes vs No	0.4062	0.2114	0.0547	1.501	0.992 2.272	
pAMPK_c_01 (Pos vs 0)	-0.4341	0.1714	0.0113	0.648	0.463 0.907	
pmTOR_c_01 (Pos vs 0)	-0.5364	0.2032	0.0083	0.585	0.393 0.871	
CXCR2_c_01m (>=23.3 vs <23.3)	0.5065	0.1681	0.0026	1.659	1.194 2.307	
EpCAM_c_01 (Pos vs 0)	-0.5211	0.1797	0.0037	0.594	0.418 0.845	
FEN1_n_01m (>=70 vs <70)	0.4246	0.1757	0.0156	1.529	1.084 2.158	

Internal validation

$$\text{C-index} = 0.34/2 + 0.5 = 0.67$$

Calibration



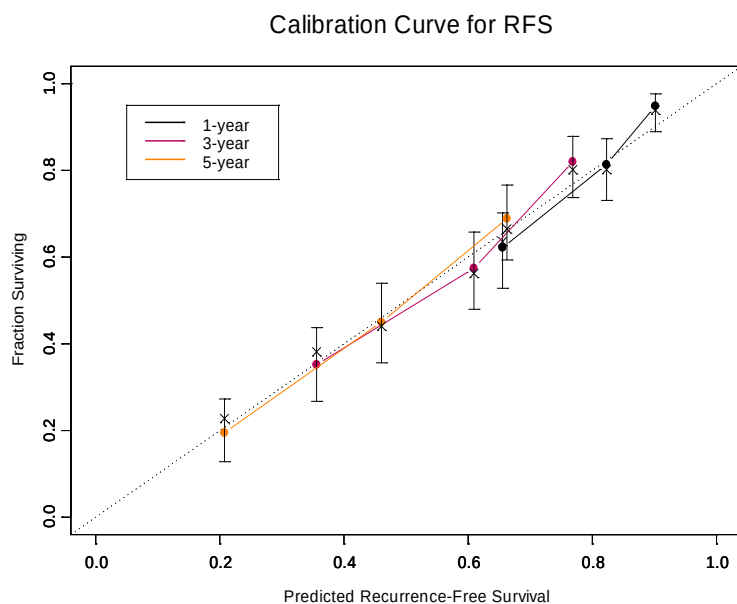
Multivariate Cox Model assessing the following covariates on Recurrence Free Survival

Analysis of Maximum Likelihood Estimates						
Variable	Parameter Estimate	Standard Error	p-value	HR	95% HR CL	
age	0.0268	0.0075	0.0004	1.027	1.012	1.042
Stage II vs I	0.5526	0.1814	0.0023	1.738	1.218	2.480
III vs I	0.9961	0.1866	<.0001	2.708	1.878	3.903
Neoadjuvant Yes vs No	0.2635	0.1874	0.1598	1.301	0.901	1.879
IGF1R_c	0.0042	0.0020	0.0416	1.004	1.000	1.008
Insulin_m_01 (Pos vs 0)	0.3879	0.1468	0.0082	1.474	1.105	1.965
pAMPK_c_01 (Pos vs 0)	-0.5052	0.1512	0.0008	0.603	0.449	0.811
pmTOR_c_01 (Pos vs 0)	-0.3171	0.1664	0.0567	0.728	0.526	1.009
CXCR2_c_01m (>=23.3 vs <23.3)	0.3329	0.1484	0.0249	1.395	1.043	1.866
EpCAM_c_01 (Pos vs 0)	-0.3513	0.1529	0.0216	0.704	0.522	0.950
CASK_m	-0.0041	0.0020	0.0377	0.996	0.992	1.000

Internal validation

$$\text{C-index} = 0.33/2 + 0.5 = 0.66$$

Calibration



Clinical Cancer Research



A five-gene and corresponding-protein signature for stage-I lung adenocarcinoma prognosis

Humam Kadara, Carmen Behrens, Ping Yuan, et al.

Clin Cancer Res Published OnlineFirst December 16, 2010.

Updated Version

Access the most recent version of this article at:
doi:[10.1158/1078-0432.CCR-10-2703](https://doi.org/10.1158/1078-0432.CCR-10-2703)

Author Manuscript

Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts

[Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions

To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.

A five-gene and corresponding-protein signature for stage-I lung adenocarcinoma prognosis

Humam Kadara¹, Carmen Behrens¹, Ping Yuan², Luisa Solis², Diane Liu³, Xuemin Gu³, John D. Minna⁴, J. Jack Lee³, Edward Kim¹, Waun-Ki Hong¹, Ignacio I. Wistuba^{1,2†*} and Reuben Lotan^{1†*}.

Departments of ¹Thoracic/Head and Neck Medical Oncology, ²Pathology and ³Biostatistics, The University of Texas M.D. Anderson Cancer Center, Houston TX and ⁴Hamon Center for Therapeutic Oncology the University of Texas Southwestern, Dallas, TX.

[†] These authors have contributed equally to this work and both should be considered as corresponding authors.

*To whom correspondence should be addressed: Reuben Lotan, Ph.D., Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M.D. Anderson Cancer Center, Tel: 713-792-8467, Fax: 713-745-5656, Email: rlotan@mdanderson.org; or Ignacio I. Wistuba, M.D., Departments of Pathology and Thoracic/Head and Neck Medical Oncology, the University of Texas M.D. Anderson Cancer Center, Tel: 713-563-9184, Fax: 713-563-1848, Email: iiwistuba@mdanderson.org.

Keywords: Lung adenocarcinoma, NSCLC, gene signature, prognosis

Running title: Lung adenocarcinoma five-gene signature

Grant support

This work was supported in part by the Jeffrey Lee Cousins Fellowship in Lung Cancer Research (to HK), Department of Defense (grant W81XWH-04-1-0142 to W.K.H., R.L. and I.I.W.), Jimmy Lane Hewlett Fund for Lung Cancer Research (to R.L. and I.I.W) and NCI lung cancer SPORE (P50 CA70907 to J.D.M. and I.I.W.).

Statement of Translational Relevance

Identification of molecular prognostic markers will improve the clinical management of NSCLC, the leading cause of cancer-related deaths in the United States and world wide. We derived a five-gene and corresponding immunohistochemical protein signature and tested their prognostic capacities in various publicly available microarray datasets and in an independent set of NSCLC histological tissue specimens, respectively. Both the five-gene transcript and corresponding protein signatures effectively predicted poor survival of all-stages or stage-I lung adenocarcinoma but not of squamous cell carcinoma patients. Moreover, the FILM protein signature specifically identified a subgroup of non-treated stage-IB lung adenocarcinoma patients with poor prognosis. These findings suggest that the derived five-gene signature may be assessed for expression by different methods (transcript versus protein) and for identifying early stage lung adenocarcinoma patients with poor prognosis that will benefit from adjuvant therapy following resective surgery.

Abstract

PURPOSE: Identification of effective markers for outcome is expected to improve the clinical management of non-small cell lung cancer (NSCLC). Here, we assessed in NSCLC the prognostic efficacy of genes, which we had previously found to be differentially expressed in an *in vitro* model of human lung carcinogenesis. **EXPERIMENTAL DESIGN:** Prediction algorithms and risk-score models were applied to the expression of the genes in publicly available NSCLC expression datasets. The prognostic capacity of the immunohistochemical expression of proteins encoded by these genes was also tested using formalin-fixed paraffin-embedded (FFPE) tissue specimens from 156 lung adenocarcinomas and 79 squamous cell carcinomas (SCCs). **RESULTS:** The survival of all-stages ($p<0.001$, HR=2.0) or stage-I ($p<0.001$, HR=2.84) adenocarcinoma patients that expressed the five-gene *in vitro* lung carcinogenesis model (FILM) signature was significantly poorer than that of patients who did not. No survival differences were observed between SCCs predicted to express or lack FILM signature. Moreover, all stages ($p<0.001$, HR=1.95) or stage-I ($p=0.001$, HR=2.6) adenocarcinoma patients predicted to be at high risk by FILM transcript exhibited significantly worse survival than patients at low risk. Furthermore, the corresponding protein signature was associated with poor survival (all stages, $p<0.001$, HR=3.6; stage-I, $p<0.001$, HR=3.5; stage-IB, $p<0.001$, HR=4.6) and mortality risk (all stages, $p=0.001$, HR=4.0; stage-I, $p=0.01$, HR=3.4; stage-IB, $p<0.001$, HR=7.2) in lung adenocarcinoma patients. **CONCLUSIONS:** Our findings highlight a gene and corresponding protein signature with effective capacity for identification of stage-I lung adenocarcinoma patients with poor prognosis that are likely to benefit from adjuvant therapy.

Introduction

Lung cancer is the leading cause of cancer-related deaths in the United States and worldwide (1). NSCLC accounts for the majority of the lung cancer cases and is comprised of two major subtypes, lung adenocarcinomas and SCCs (2). The average 5-year relative survival rate among NSCLC patients is only 15% (2-4). Mortality due to NSCLC is high because most cancers are diagnosed after regional or distant spread of the disease (3, 4). However, even the 5-year survival rate of stage I NSCLC patients (30-50%) is among the worst for early-stage disease of all other malignancies (4, 5). Therefore, identification of markers for early prediction of outcome is warranted for better clinical management of NSCLC patients, and in particular for those with early stage disease.

Although surgical resection is the first treatment, adjuvant therapy has been shown to improve the survival of lung cancer patients (6, 7). Several factors have been proposed to elevate risk and justify the use of adjuvant therapy for lung cancer patients of which the TNM staging is the most effective standard (5). However the potential benefits of adjuvant therapy are contentious, in particular in stage-I lung cancer patients (6, 7) as there are few if not any established clinical criteria to separate stage-I NSCLCs (8). Therefore, it is possible that additional molecular factors might help identify early or stage-I lung cancer patients with poor prognosis that may need to receive therapy versus patients with good prognosis that could be spared adjuvant therapy.

With the advent of microarray and high-throughput technology, several early studies have shown the significant association of gene expression profiles and signatures with the survival and outcome of NSCLC patients (9-12). Additionally, gene expression signatures have been derived for the prognosis of early (stages I and II) (13, 14) and all stages NSCLC patients (15-17), lung adenocarcinomas alone (18, 19) and to predict recurrence of NSCLC disease (20, 21). Moreover, a five-gene signature identified from a set of 672 genes differentially expressed among invasive lung adenocarcinoma cell lines identified NSCLC patients at high risk of poor survival (16). We have previously studied the gene expression profiles of normal, premalignant and tumorigenic lung epithelial cells constituting an *in vitro*

Lung adenocarcinoma five-gene signature model of lung carcinogenesis and highlighted prominent gene expression profiles and pathways relevant to the survival of lung adenocarcinoma patients (22). Specifically, we identified the progressive modulation of six key genes, ubiquitin conjugating enzyme 2C (*UBE2C*), minichromosome maintenance (*MCM*) 2 and 6, targeting protein for Xklp2 (TPX2), flap structure-specific endonuclease 1 (*FEN1*) and stratifin (*SFN*), among the cell lines and the significant association of their expression with the survival of lung adenocarcinoma patients (22).

In this study, we sought to assess the prognostic efficacy of the aforementioned six genes in NSCLC. We analyzed the association of the expression of the genes at the transcript level with patient survival in several publicly available microarray datasets of lung adenocarcinoma and SCC. We derived a five-gene signature that is predictive of survival of early stage lung adenocarcinoma but not SCC patients. Moreover, we analyzed the proteins encoded by the five-gene signature using immunohistochemistry in an independent FFPE tissue microarray (TMA) NSCLC set, and found that it was also effective in predicting the survival of lung adenocarcinoma patients including those with stage I disease. Moreover, risk of mortality assessed by both the transcript and protein versions of the derived five-gene expression signature was an independent predictor of poor survival in lung adenocarcinoma patients. Importantly, the FILM protein signature was capable of identifying stage-IB lung adenocarcinoma patients with dismal prognosis warranting the need of adjuvant therapy in these patients.

Methods

Analysis of publicly available NSCLC microarray datasets

To assess the expression and clinical relevance of the aforementioned six genes in human NSCLC we used publicly available lung adenocarcinoma microarray data sets from the studies by Shedden *et al* (National Cancer Institute (NCI) Director's Challenge, n=442) (19), Bild *et al* (Duke cohort, n=58) (15) and Bhattacharjee *et al* (Harvard cohort, n=125) (10). We also used publicly

Lung adenocarcinoma five-gene signature
 available microarray datasets generated from lung SCC patient samples from the studies by Raponi *et al* (n=130) (17) and Bild *et al* (n=53) (15). The raw microarray data for the Director's Challenge study were obtained from the National Cancer Institute Cancer Array database, experiment ID 1015945236141280:1 (<https://caarraydb.nci.nih.gov/caarray>). Raw data from all the published studies were obtained from the Gene Expression Omnibus (GEO). Raw microarray data from all data sets were analyzed using the BRB-ArrayTools v.3.8.0 developed by Dr. Richard Simon and BRB-ArrayTools Development Team (23). Robust multi-array analysis (RMA) was used for normalization of gene expression data in R language environment (24). Probe sets for *UBE2C*, *MCMs* 2 and 6, *FEN1*, *TPX2* and *SFN* in the microarray platforms used in the different studies (Affymetrix HG-U95Av2, HG-U133A and HG-U133 plus 2.0) were identified using NetAffx™ from Affymetrix (<http://www.affymetrix.com/analysis/index.affx>).

Prediction of class

To predict the class of independent patient cohorts, we adopted a previously developed model using six algorithms, compound covariate predictor (CCP), linear discriminator analysis (LDA), nearest neighbors 1 and 3 (NN-1 and NN-3), nearest centroid (NC) and support vector machines (SVM) (23, 25). Lung adenocarcinomas from the datasets in the study by the Shedden *et al* (19) study were designated together as a training set (n=442). Class separation into high versus low expression (Figure 1A) was maintained following analysis of the six genes in the 442 lung adenocarcinomas by hierarchical cluster analysis with average linkage (data not shown) using Cluster version 2.11. Samples from the Duke (15) and Harvard (10) cohorts were used as a combined lung adenocarcinoma test set and median-centered independently (n=183), whereas SCC samples from the studies by Raponi *et al* (17) and Bild *et al* (15) were designated as an SCC validation set (n=183) (Figure 1A). Assessment of the classification efficacy of the six genes as well as their capacity in estimating the probability of the identity of a particular sample was performed using a leave-one-out-cross-validation (LOOCV) approach with random

Lung adenocarcinoma five-gene signature permutation for accuracy estimation as previously described (23, 25). Five (all but *SFN*) of the six genes were capable of significant class prediction in training statistically assessed by a univariate t-test with a statistical cut-off of $p < 10^{-5}$ (Figure 1A). Furthermore and compared to the other five genes, *SFN* was down-regulated in tumors relative to normal tissue, exhibited relatively lower mean expression in tumors and increased misclassification. The six classification algorithms were then applied to the indicated test sets and the different patient groups or arms, predicted by the derived FILM signature, were analyzed for statistically significant differences in survival by Kaplan-Meier method for estimation of survival probability and log-rank tests in R language environment.

Computation of risk scores

We generated a model for estimation of mortality risk similar to what was described earlier (26). Using gene expression data from training cohorts during LOOCV, the Cox regression coefficients were computed for *UBE2C*, *MCMs* 2 and 6, *FEN1* and *TPX2*. A risk score was then derived for each patient in the training set by calculating the summation of the products of the Cox coefficients and normalized centered expression of each gene. Patients were dichotomized into high-risk and low-risk groups using the 50th percentile (median) cut-off of the risk score as the threshold value or were divided into tertiles or three groups(low, intermediate and high risk). In all cases, the coefficient and the threshold value derived from the training cohort were directly applied to the gene expression data from validation cohorts. Kaplan-Meier and log-rank test were then applied to assess the significance of prognostic difference between the risk groups without or with the inclusion of age and gender clinical covariates as previously described (27).

Tissue microarray

For this study, we obtained archived FFPE samples from surgically resected lung cancer specimens from the lung cancer tissue bank at The University of Texas M. D. Anderson Cancer Center

Lung adenocarcinoma five-gene signature

(Houston, TX). The tissue specimens originally had been collected between 2003 and 2005 and had been classified using the 2004 World Health Organization (WHO) classification system as described before (28). The tissue microarray analyzed in this study is comprised of 235 NSCLC tumor specimens (156 lung adenocarcinomas and 79 SCCs) obtained from patients, who underwent surgery at the same institution from 2003-2005 and who did not receive any adjuvant or neo adjuvant therapy, under a protocol that was approved by the MD Anderson Cancer Center institutional review board. Detailed clinical and pathologic information was available for most of these cases and included patients' demographic data, smoking history (never smokers or ever smokers, patients who had smoked at least 100 cigarettes in their lifetime) and pathologic tumor-node-metastasis (TNM) staging (29). After histological examination of NSCLC specimens, the NSCLC TMAs were constructed by obtaining three 1-mm in diameter cores from each tumor at three different sites (periphery, intermediate and central tumor sites). The TMAs were prepared with a manual tissue arrayer (Advanced Tissue Arrayer ATA100, Chemicon International, Temecula, CA).

Immunohistochemical analysis of the protein signature

Immunohistochemistry was done on histology sections of FFPE tissue samples, using the purified primary anti-human antibodies against UBE2C (1:500 dilution) (Boston Biochem, Cambridge, MA), FEN1 (1:50 dilution) (BD Biosciences, San Jose, CA), MCM2, MCM6, SFN (all 1:100 dilution) (Novus Biologicals, Littleton, CO) and TPX2 (1:400 dilution) (Novus Biologicals). The sections were deparaffinized, hydrated, subjected to antigen retrieval by heating in a steamer for 20 min with 10 mmol/L sodium citrate (pH 6.0), and then incubated in peroxidase blocking reagent (DAKO). Sections were then washed with Tris-containing buffer and incubated overnight at 4°C with the primary antibodies. Subsequently, the sections were washed and incubated with secondary antibodies using the Evision plus labeled polymer kit (DAKO) for 30 min followed by incubation with avidin-biotin-peroxidase complex (DAKO) and development with diaminobenzidine chromogen for 5 min. Finally,

Lung adenocarcinoma five-gene signature

the sections were rinsed in distilled water, counterstained with hematoxylin (DAKO), and mounted on glass slides before evaluation under the microscope. FFPE samples processed similarly, except for the omission of the primary antibody, were used as negative controls. Experienced lung cancer pathologists blinded to the clinical data examined the immunostainings jointly at the same time using light microscopy to generate one set of readings (P.Y. and I.I.W.). The antigens studied exhibited mainly nuclear immunoreactivity. The immunostainings were quantified using a four-value intensity score (0, 1+, 2+, and 3+) and the percentage (0-100%) of the extent of reactivity in each core. The final score was then obtained by multiplying the intensity and reactivity extension values (range, 0-300) as previously reported (22, 28, 30).

To understand the association of the expression of the studied proteins with the survival of NSCLC patients, a combined immunoreactivity score for each patient was computed by simple addition of the individual final immunoreactivity scores for each of the analyzed antigens. Lung adenocarcinoma and SCC patients were then dichotomized into a high and low expression group using the 50th percentile (median) cut-off of the combined final immunoreactivity score value. Alternatively for confirmation, patient samples were clustered by average linkage using Cluster v 2.11 program following median-centering of the antigens' immunoreactivity scores. Clusters were then identified and visualized using TreeView programs (Michael Eisen Laboratory, Lawrence Berkeley National Laboratory and University of California, Berkeley; <http://rana.lbl.gov/EisenSoftware.htm>). To further validate the prognostic relevance of the protein signature, lung adenocarcinomas were randomly divided into a training set (n=78) and a complete (n=78) or stage I only test set (n=62) and a mortality risk was estimated similar to what was described above using Cox coefficients in the training set and protein expression within each set. Statistically significant differences in the survival of the clusters, expression or risk groups were analyzed by the Kaplan-Meier and log-rank tests in R language environment without and with the inclusion of and adjustment for age and gender as previously described (27).

Results

Derivation of a five-gene signature predictive of survival in lung adenocarcinoma

We have previously identified genes that are expressed differentially among cells constituting an *in vitro* model of lung carcinogenesis (22). Functional pathways analysis of genes differentially expressed between the previously studied lung tumorigenic (1170-I) and normal lung epithelial cells aided us to identify a significantly modulated gene-interaction network comprised of six key genes based on level of modulation and number of interactions with neighboring molecules (22). Here, we sought to test the relevance of these six genes to NSCLC prognosis. The lung adenocarcinoma microarray datasets from the NCI Director's challenge study (Shedden *et al*) (19) were used as a training set (n=442). Lung adenocarcinomas from the Duke (15) and Harvard (10) cohorts served as a combined adenocarcinoma validation set, whereas SCCs from the studies by Raponi *et al* (17) and Bild *et al* (15) served as an SCC test set (Figure 1A). An LOOCV approach, described in more detail in the Methods section of this manuscript, was used to train the six genes based on class separation of high versus low expression identified by cluster analysis (not shown). The number of genes was reduced to five following application of a t-test $p < 10^{-5}$ statistical cut-off to minimize misclassification during LOOCV (Figure 1A) giving rise to a five-gene signature which we have designated as FILM. The FILM signature was efficacious as indicated by the sensitivity and specificity of the six prediction algorithms (Supplementary Table 1), which demonstrated that overall survival of all stages ($p < 0.001$) and stage-I only ($p < 0.001$) lung adenocarcinoma patients predicted to express higher levels of the FILM signature was significantly poorer compared to patients with lower expression of FILM using the linear-discriminator analysis method (LDA) for cross-validation (Figure 1B). Similar findings were obtained using the other methods outlined in Supplementary table 1 (data not shown).

We integrated genes (n=584) we had previously found to be differentially expressed between normal and lung tumorigenic cells (22) with the Shedden *et al* dataset which in turn subdivided patients into two clusters with significant differences in survival (data not shown). To compare to the

Lung adenocarcinoma five-gene signature performance of the FILM signature, we also analyzed the 584 genes using similar approaches and derived a five-gene signature by recursive feature selection and a fold difference in expression of at least two between two classes identified by cluster analysis (data not shown). The recursive feature-generated five-gene signature exhibited reduced sensitivity and specificity compared to the FILM signature (data not shown). These findings demonstrate the effectiveness of the FILM signature, despite being selected *a priori*, in predicting the survival of lung adenocarcinoma patients.

The FILM signature does not predict survival of lung SCC patients

We then examined the capacity of the FILM signature to predict survival in lung SCC. A similar strategy to that depicted in Figure 1A was employed except that a pooled SCC test set (n=183) was used. All six prediction algorithms depicted the inability of the FILM signature to predict survival in lung SCC (Figure 1C and data not shown). Similar results were obtained when only lung SCC patients were used in the training and test sets (data not shown). These findings exemplify the prognostic specificity of the FILM signature for lung adenocarcinoma.

Mortality risk assessed by the FILM expression signature predicts survival in all stages or stage-I lung adenocarcinoma

We then sought to further validate the robustness of the FILM expression signature in predicting survival in lung adenocarcinoma. We used a strategy similar to what was described before (26) to estimate mortality risk based on computation of risk scores. Lung adenocarcinomas from the Shedden *et al* (19) study were used as a training set whereas adenocarcinomas from the Harvard (10) and Duke (15) cohorts were pooled as a test set (Figure 2A). We developed risk scores for patients using the Cox regression coefficients of the genes in the FILM signature and their normalized expression data in the training cohort and patients were dichotomized by the median (50%) risk score (Figure 2B) or were divided into tertiles and three risk groups (low, intermediate and high; Supplementary Figure 1) as

Lung adenocarcinoma five-gene signature described before (19). Lung adenocarcinoma patients identified to be at high risk based on FILM signature exhibited significantly worse survival than patients at low risk ($p=5.9 \times 10^{-5}$) (Figure 2C). For validation of the risk score model, Cox regression coefficients and dichotomization cut-off threshold generated from the training cohort were directly applied to the validation set ($n=183$). All stages ($p<0.001$) or stage-I ($p=0.001$) lung adenocarcinoma patients predicted to be at high risk displayed significantly worse survival (Figure 2D). Similar results were obtained when patients in the training sets were divided into tertile risk groups (Supplementary Figure 1) along with improved capacity of the FILM transcript signature to separate stage-I but not all stages lung adenocarcinoma patients with poor survival from those with excellent survival (Supplementary Figures 1C and 1D). Multivariate Cox proportional hazard regression analyses of patients divided into three risk groups revealed that mortality risk assessed by the FILM transcript signature was, along with stage, an independent predictor of survival ($p=0.03$, HR=1.8, 95% CI=1.06-3.1) in all stages lung adenocarcinoma patients (Supplementary Table 2). Importantly, in stage-I only patients, high risk estimated by the FILM signature was an independent predictor of survival ($p<0.001$ HR=4.5, 95% CI 1.98-10.16) similar in significance to stage IA versus IB disease ($p=0.001$, HR=0.135, 95% CI 0.04-0.45) (Supplementary Table 2). Survival probability analysis was then adjusted by inclusion of clinical covariates, age and gender, in all stages and stage-I test set patients similar to what was described earlier by Shedden et al (19). Inclusion of age and gender as covariates, enhanced the capacity of the FILM transcript signature to separate patients in the test set with good survival from those with poorer survival more notably when analyzing events after seventy months and in particular in stage-I patients (Supplementary Figures 2B and 2C). In contrast, the FILM signature performed similarly with or without clinical covariates when analyzed in the Director's challenge study utilized as a training set (Supplementary Figure 2A).

We then asked whether there is an association between mortality risk assessed by the FILM signature and survival in stage IA and stage IB patients separately. Stage IB but not IA lung adenocarcinoma patients, from the datasets of the NCI Director's challenge, Duke and Harvard cohorts

Lung adenocarcinoma five-gene signature combined, at high risk of mortality as predicted by the FILM signature exhibited significantly poorer overall survival than patients at low risk ($p < 0.001$, HR=2.1, 95% CI=1.4-3.2) (Supplementary Figure 3A). All stage IA and IB patients were included in the analysis as the therapy status of patients in the Duke and Harvard cohorts is not available. We found similar results, albeit less effective, when we analyzed stage IA and IB lung adenocarcinoma patients from the NCI Director's challenge datasets dataset that are known to have not been treated with any form of therapy (Supplementary Figure 3B). These findings further validate the robustness of the FILM expression signature in predicting survival in lung adenocarcinoma and demonstrate the signature's capacity in identifying a subpopulation of stage-I patients with poor prognosis.

The corresponding FILM protein signature also predicts survival of non-treated all stages or stage-I lung adenocarcinoma patients

We then explored the prognostic capacity of the FILM signature in NSCLC at the protein level because the expression of transcripts and corresponding encoded proteins do not always correlate (31). We analyzed the protein expression of the FILM signature by immunohistochemistry analysis in FFPE histological tissue specimens obtained from 156 and 79 lung adenocarcinoma and SCC patients, respectively, and who did not receive any therapy before or after tumor resection (Figure 3A). A combined total immunoreactivity score for FILM protein expression was computed, and patients were then dichotomized based on the median FILM protein score. In accordance with the findings obtained with the transcript signature, all-stages ($p < 0.001$, HR=3.6, 95% CI=1.9-6.8) or stage-I only ($p < 0.001$, HR=3.5, 95% CI=1.7-7.5) lung adenocarcinoma patients with higher expression of FILM protein exhibited significantly poorer survival compared to patients with lower expression. In contrast, the protein signature was again not prognostic in lung SCC (Figure 3B). To confirm the prognostic capacity of the FILM protein signature, all stages (Figure 3C) or stage-I (Figure 3D) lung adenocarcinoma patients were analyzed by hierarchical cluster analysis by average linkage based on the centered

Lung adenocarcinoma five-gene signature expression of the immunoreactivity scores of FILM protein expression. Two clusters were identified with dissimilar expression of FILM protein. All stages or stage-I only lung adenocarcinoma patients in the cluster with higher FILM protein (high cluster) exhibited significantly poorer survival compared to patients in the low FILM protein cluster (both $p < 0.001$) (Figures 3C and 3D).

We then determined to further confirm the robustness of the FILM protein signature in predicting the survival of lung adenocarcinoma patients. Lung adenocarcinoma patients were randomized into a training ($n=78$) set and an all-stages ($n=78$) or stage-I only ($n=62$) test set (Figure 4A). As performed with the gene expression signature, risk scores were computed based on the FILM proteins centered immunoreactivity scores and Cox coefficients in the training set (Figure 4A). Patients were then dichotomized based on the median (50%) risk score into high versus low risk groups (Figure 4B) or into tertiles (low, intermediate and high risk groups; data not shown) and subsequently analyzed for survival differences. Lung adenocarcinoma patients in the training set and predicted to be at high mortality risk based on FILM protein exhibited significantly poorer survival than patients at low risk ($p=0.01$) (Figure 4C). Cox regression coefficients and dichotomization cut-off threshold generated from the training set were then directly applied to the test set ($n=78$) or only stage-I patients within the set ($n=62$) and analyzed similarly. In accordance, all stages ($p=0.001$, HR=4.0, 95% CI=1.6-9.8) or stage-I ($p=0.01$, HR=3.4, 95% CI=1.4-9.1) lung adenocarcinoma patients predicted to be at high mortality risk based on FILM protein exhibited significantly poor survival (Figure 4D and Table 1). Separation of patients by the FILM protein risk score into tertiles did not increase its capacity for survival prediction (data not shown). Moreover, multivariate Cox hazard regression analyses demonstrated that risk assessed by FILM protein was an independent predictor of survival in all stages ($p=0.005$, HR=5.78, 95% CI=1.7-20) almost similar in connotation to the most significant variable, age ($p=0.002$, HR=1.1, 95% CI=1.03-1.15) (Table 1). In addition, FILM protein risk was also an independent, and better than IB stage, predictor of survival in stage I patients ($p=0.01$, HR=6, 95% CI=1.4-26.2) (Table 1). Inclusion of clinical covariates (age and gender) enhanced the capacity of the FILM protein signature to separate

Lung adenocarcinoma five-gene signature patients with poor survival from those with excellent survival (Supplementary Figure 4). When comparing at 50 months follow up and onwards and following adjustment of survival probability plots for age and gender, no events or deaths were noted in low risk patients in contrast to survival analysis without the clinical covariates.

Since the FILM protein signature was effective in stage-I lung adenocarcinoma prognosis, we tested its prognostic capacity in stage IA and stage IB patients separately. When patients were dichotomized based on mortality risk computed by FILM FFPE protein signature, stage IB but not stage IA patients with higher risk exhibited significantly poorer survival than patients at low risk ($p < 0.001$, HR=4.2, 95% CI=1.8-9.9) (Figure 5A). Similarly, when patients were divided into tertile groups of low, intermediate and high risk, FILM protein signature was able to further separate stage-IB patients with poor survival from those with excellent survival ($p < 0.001$, HR=7.26, 95% CI=2.40-21.98) (Figure 5B). These results demonstrate that the FILM protein signature, like its transcript version, is valuable for predicting the survival of lung adenocarcinoma patients and that the protein signature may be valuable for identifying stage-I or -IB patients who may benefit from adjuvant therapy.

Discussion

We previously identified genes differentially expressed among cells constituting an *in vitro* model of lung carcinogenesis (22). Functional pathways analysis of the differentially expressed genes highlighted a significantly modulated gene-interaction network comprised of six key genes, *UBE2C*, *MCM2*, *MCM6*, *FEN1*, *TPX2* and *SFN*, and that were associated with poor survival in lung adenocarcinoma. In this study, we sought to validate and further examine the prognostic capacity of these genes using multiple methods, various publicly available expression datasets of lung adenocarcinoma and SCCs as well as an additional set of FFPE NSCLCs. We derived a five-gene signature, FILM (*UBE2C*, *MCM2*, *MCM6*, *FEN1*, *TPX2*), that was predictive of poor survival in lung adenocarcinoma including those with stage-I disease. In contrast, the FILM signature was not prognostic

Lung adenocarcinoma five-gene signature in lung SCCs. In addition, we developed a risk model based on the FILM signature with and without inclusion of clinical covariates (age and gender) and demonstrated that FILM transcript signature mortality risk score was predictive of poor survival in lung adenocarcinoma. Moreover, we analyzed and validated the expression of the protein version of the FILM classifier by immunohistochemistry in a series of FFPE NSCLCs and found that FILM protein signature, like the transcript signature, was associated with poor survival in lung adenocarcinomas and not SCCs. Furthermore, risk assessed by the FILM immunohistochemical protein signature was a significant predictor of poor survival in all stages or stage-I lung adenocarcinomas. Lastly, we demonstrated that the FILM protein signature effectively identified a subset of non-treated stage IB lung adenocarcinomas with poor prognosis that may benefit from adjuvant therapy.

The robustness of the FILM transcript signature was validated by deriving a risk score model using the genes' Cox coefficients and expression in the training set similar to what was described before for predicting recurrence in tamoxifen-treated node-negative breast cancer patients (26). Following direct application of the same Cox coefficients and dichotomization threshold from the training set onto an independent test set, mortality risk assessed by the FILM expression signature was an independent predictor of poor survival in all stages or stage-I lung adenocarcinomas. It is worthwhile to mention that mortality risk assessed by FILM signature was significantly predictive of poor survival regardless of the identity of the training and validation cohorts (data not shown). Moreover, these findings were not replicated in lung SCCs. It is unclear why the FILM signature is only prognostic in lung adenocarcinoma but not SCCs. One possible explanation is that the molecular constituents of the FILM signature are significantly up-regulated in 1170-I lung adenocarcinoma forming cells (32) compared to normal bronchial epithelial cells (22). However, one cannot neglect the significant expression of the genes in lung SCCs. For example, TPX2 protein was shown to be up-regulated in preneoplastic lesions representing lung SCC pathogenesis as well as in tumors (33). In addition, we have previously reported the significant up-regulation of UBE2C immunohistochemical protein in lung SCCs compared to normal

Lung adenocarcinoma five-gene signature bronchial epithelia (22). Interestingly, UBE2C protein immunohistochemical expression was significantly higher in lung SCCs compared to adenocarcinomas (22). It is plausible to suggest that increased expression of the FILM signature renders lung adenocarcinomas but not SCCs more clinically aggressive.

The prognostic effectiveness of the FILM expression signature was validated at an additional level by assessing the immunohistochemical expression of proteins encoding the FILM classifier in an independent set of FFPE NSCLCs. We deemed this approach to be important as the levels of proteins need not to match with the expression level of transcripts (31, 34). In clear accordance to the transcript signature, the FILM protein signature was prognostic in lung adenocarcinoma (all stages or stage-I) but not in SCC patients and an independent predictor of poor survival in stage-I lung adenocarcinomas. However, it is noteworthy that the semi-quantitative and subjective nature of immunohistochemistry analyses, unlike automated quantitative determination of histological protein expression (35), poses a limitation on using an immunohistochemical protein signature for prognostic purposes such as independent validation by other groups or studies. Moreover, a shortcoming of a protein prognostic classifier from paraffin embedded tissues is the possible variability in readings among different pathologists using the same or different tissue sets. Nevertheless, we attempted to overcome this potential limitation by averaging the score of three tissue cores representing three different sites per tumor specimen to generate final immunohistochemical scores. Moreover, we demonstrated the prognostic capacity of the FILM protein signature using various analytical methods; patient dichotomization based on median immunohistochemical score, hierarchical clustering of patients and derivation of a risk score model. In addition, the prognostic capacities of each protein within the FILM signature were analyzed independently and deemed in all cases weaker in comparison to the combined FILM protein signature. Furthermore and importantly, the results obtained using the transcript and protein versions of the FILM signature were nearly identical demonstrating the reliability and biological significance of this signature in lung adenocarcinoma prognosis.

Lung adenocarcinoma five-gene signature

We had previously derived global differential expression profiles between normal and tumorigenic lung epithelial cells that are also widely expressed in lung adenocarcinomas from the Shedden *et al* dataset (22). It is plausible that different five-gene signatures developed from all genes differentially expressed among the *in vitro* model of lung carcinogenesis may as well be equally or more efficacious in prognosis compared to the FILM signature. To test this hypothesis, we derived a five-gene signature by recursive feature selection and a fold difference in expression of at least two between the classes identified by cluster analysis of the Shedden *et al* dataset following integration of the genes (data not shown). The FILM signature, despite being originally based on genes selected *a priori*, was superior to the recursive feature-generated five-gene signature in sensitivity and specificity as well as in predicting poor survival in lung adenocarcinoma. The effectiveness of the FILM signature in lung adenocarcinoma prognosis may be due to the nature of its genes that are proliferation-related despite acting at different points and through distinct mechanisms in cell cycle control (33, 36-42). It is noteworthy that in a large meta analysis using breast tumors from several publicly available breast cancer expression datasets, Wirapati *et al* demonstrated that all nine prognostic signatures compared exhibited similar prognostic capacities which were mostly driven by proliferation-related genes (43, 44). Moreover, in the NCI Director's Challenge study by Shedden *et al*, the most effective gene classifiers in survival prediction (methods A and H as designated by the authors) were mainly comprised of proliferation related genes and performed well with or without clinical covariates. Interestingly, all genes within the FILM signature were found in methods A and H developed by Shedden *et al* which may explain why the FILM transcript signature performed similarly with or without clinical covariates when analyzed in the Director's Challenge study.

In recent reviews of studies geared towards the development of prognostic expression signatures in NSCLC, Subramanian and Simon have suggested guidelines to follow in the design and testing of an expression signature for lung cancer prognosis including the stratification of stage IA and stage IB NSCLC patients independently in an effort to identify patients who may benefit from adjuvant therapy

Lung adenocarcinoma five-gene signature

(6, 43, 44). Towards this, we had analyzed stage IA and stage IB patients alone using both the transcript and protein FILM signatures in publicly available microarray datasets and a tissue microarray of FFPE NSCLC specimens, respectively. We found that the FILM signature identified a subgroup of stage IB but not stage IA lung adenocarcinoma patients with poor prognosis, which was consistent as it was found by analyzing the transcript signature in publicly available microarray datasets and the parallel protein signature in a tissue microarray of FFPE specimens. However, it is noteworthy that we analyzed, as depicted in Supplementary Figure 3, the entire set of stage IA and IB lung adenocarcinomas from the Director's challenge (19), Duke (15) and Harvard (10) datasets as the therapy status for all the patients from the two latter cohorts and some from the former is not known. Nevertheless, when analyzed only in patients from the Director's challenge dataset that were not treated with any form of therapy, the FILM transcript signature was also prognostic, albeit less, in stage IB lung adenocarcinoma. Furthermore, we analyzed the FILM protein signature in a FFPE tissue microarray with more complete and annotated clinicopathological information and found that the FILM protein signature effectively identifies non-treated stage IB lung adenocarcinoma patients with poor prognosis pinpointing to a potential benefit of adjuvant treatment in this patient population. It is worthwhile to mention that the FILM protein signature, in comparison to its transcript counterpart, was more effective in identifying stage-I and in particular -IB lung adenocarcinoma patients with very poor survival and separating them from patients with excellent survival. Therefore, it is plausible to assume that the FILM protein signature may be clinically more useful for stage-I or specifically -IB lung adenocarcinoma prognosis. However and as mentioned before, it cannot be neglected that the relatively more qualitative nature of immunohistochemistry analysis of FFPE tissues and more difficult external validation compared to real-time PCR assessment of RNA poses a limitation on the translation of FFPE protein signatures to the clinic. Unless the capacity of the FILM protein signature is cross-validated in independent tissue microarrays and by different investigators, a combination of both mRNA and protein signatures may be powerful for stage-I lung adenocarcinoma prognosis.

Lung adenocarcinoma five-gene signature

In conclusion, our study describes the development and testing of an effective five-gene signature in lung adenocarcinoma prognosis. This expression signature has several unique and valuable attributes; 1) it identifies stage-I lung adenocarcinomas with poor prognosis; 2) exhibits prognostic specificity towards lung adenocarcinomas only; 3) the protein variant of the signature can effectively predict survival in all stages or stage I lung adenocarcinoma following analysis of FFPE tissue specimens by immunohistochemistry; and 4) the protein FFPE variant is effective in identifying non-treated stage-IB lung adenocarcinoma patients with dismal prognosis that are most likely to benefit from adjuvant therapy. Therefore, further independent studies are warranted to externally validate the potential clinical use of this five-gene signature, and in particular, the prognostic capacity of the FILM protein FFPE signature towards translation to the clinic for identifying non-treated stage-I and specifically IB lung adenocarcinoma patients with poor survival that will benefit from adjuvant therapy.

References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer Statistics, 2010. *CA Cancer J Clin*.
2. Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med* 2008;359: 1367-80.
3. Gabrielson E. Worldwide trends in lung cancer pathology. *Respirology* 2006;11: 533-8.
4. Sun S, Schiller JH, Spinola M, Minna JD. New molecularly targeted therapies for lung cancer. *J Clin Invest* 2007;117: 2740-50.
5. Tanoue LT. Staging of non-small cell lung cancer. *Semin Respir Crit Care Med* 2008;29: 248-60.
6. Subramanian J, Simon R. Gene expression-based prognostic signatures in lung cancer: ready for clinical use? *J Natl Cancer Inst*;102: 464-74.
7. Tsuboi M, Ohira T, Saji H, *et al*. The present status of postoperative adjuvant chemotherapy for completely resected non-small cell lung cancer. *Ann Thorac Cardiovasc Surg* 2007;13: 73-7.
8. Kratz JR, Jablons DM. Genomic prognostic models in early-stage lung cancer. *Clin Lung Cancer* 2009;10: 151-7.
9. Beer DG, Kardia SL, Huang CC, *et al*. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002;8: 816-24.
10. Bhattacharjee A, Richards WG, Staunton J, *et al*. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 2001;98: 13790-5.
11. Garber ME, Troyanskaya OG, Schluens K, *et al*. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A* 2001;98: 13784-9.
12. Wigle DA, Jurisica I, Radulovich N, *et al*. Molecular profiling of non-small cell lung cancer and correlation with disease-free survival. *Cancer Res* 2002;62: 3005-8.
13. Lu Y, Lemon W, Liu PY, *et al*. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med* 2006;3: e467.

Lung adenocarcinoma five-gene signature

14. Potti A, Mukherjee S, Petersen R, *et al.* A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N Engl J Med* 2006;355: 570-80.
15. 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.
16. Chen HY, Yu SL, Chen CH, *et al.* A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med* 2007;356: 11-20.
17. 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.
18. Guo L, Ma Y, Ward R, Castranova V, Shi X, Qian Y. Constructing molecular classifiers for the accurate prognosis of lung adenocarcinoma. *Clin Cancer Res* 2006;12: 3344-54.
19. Shedden K, Taylor JM, Enkemann SA, *et al.* Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008;14: 822-7.
20. Larsen JE, Pavey SJ, Passmore LH, *et al.* Expression profiling defines a recurrence signature in lung squamous cell carcinoma. *Carcinogenesis* 2007;28: 760-6.
21. Larsen JE, Pavey SJ, Passmore LH, Bowman RV, Hayward NK, Fong KM. Gene expression signature predicts recurrence in lung adenocarcinoma. *Clin Cancer Res* 2007;13: 2946-54.
22. Kadara H, Lacroix L, Behrens C, *et al.* Identification of gene signatures and molecular markers for human lung cancer prognosis using an in vitro lung carcinogenesis system. *Cancer Prev Res (Phila Pa)* 2009;2: 702-11.
23. Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y. Analysis of Gene Expression Data Using BRB-Array Tools. *Cancer Inform* 2007;3: 11-7.
24. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003;31: e15.
25. Lee JS, Heo J, Libbrecht L, *et al.* A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med* 2006;12: 410-6.

Lung adenocarcinoma five-gene signature

26. Paik S, Shak S, Tang G, *et al.* A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351: 2817-26.
27. Lee J, Yoshizawa C, Wilkens L, Lee HP. Covariance adjustment of survival curves based on Cox's proportional hazards regression model. *Comput Appl Biosci* 1992;8: 23-7.
28. Behrens C, Feng L, Kadara H, *et al.* Expression of interleukin-1 receptor-associated kinase-1 in non-small cell lung carcinoma and preneoplastic lesions. *Clin Cancer Res*;16: 34-44.
29. Mountain CF. The international system for staging lung cancer. *Semin Surg Oncol* 2000;18: 106-15.
30. Prudkin L, Behrens C, Liu DD, *et al.* Loss and reduction of FUS1 protein expression is a frequent phenomenon in the pathogenesis of lung cancer. *Clin Cancer Res* 2008;14: 41-7.
31. Maier T, Guell M, Serrano L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett* 2009;583: 3966-73.
32. Klein-Szanto AJ, Iizasa T, Momiki S, *et al.* A tobacco-specific N-nitrosamine or cigarette smoke condensate causes neoplastic transformation of xenotransplanted human bronchial epithelial cells. *Proc Natl Acad Sci U S A* 1992;89: 6693-7.
33. Ma Y, Lin D, Sun W, *et al.* Expression of targeting protein for xklp2 associated with both malignant transformation of respiratory epithelium and progression of squamous cell lung cancer. *Clin Cancer Res* 2006;12: 1121-7.
34. de Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C. Global signatures of protein and mRNA expression levels. *Mol Biosyst* 2009;5: 1512-26.
35. Zheng Z, Chen T, Li X, Haura E, Sharma A, Bepler G. DNA synthesis and repair genes RRM1 and ERCC1 in lung cancer. *N Engl J Med* 2007;356: 800-8.
36. Reddy SK, Rape M, Margansky WA, Kirschner MW. Ubiquitination by the anaphase-promoting complex drives spindle checkpoint inactivation. *Nature* 2007;446: 921-5.

Lung adenocarcinoma five-gene signature

37. Sato M, Girard L, Sekine I, *et al.* Increased expression and no mutation of the Flap endonuclease (FEN1) gene in human lung cancer. *Oncogene* 2003;22: 7243-6.
38. Stegmeier F, Rape M, Draviam VM, *et al.* Anaphase initiation is regulated by antagonistic ubiquitination and deubiquitination activities. *Nature* 2007;446: 876-81.
39. Tan DF, Huberman JA, Hyland A, *et al.* MCM2--a promising marker for premalignant lesions of the lung: a cohort study. *BMC Cancer* 2001;1: 6.
40. Warbrick E, Coates PJ, Hall PA. Fen1 expression: a novel marker for cell proliferation. *J Pathol* 1998;186: 319-24.
41. Whitfield ML, Sherlock G, Saldanha AJ, *et al.* Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell* 2002;13: 1977-2000.
42. Zheng L, Dai H, Qiu J, Huang Q, Shen B. Disruption of the FEN-1/PCNA interaction results in DNA replication defects, pulmonary hypoplasia, pancytopenia, and newborn lethality in mice. *Mol Cell Biol* 2007;27: 3176-86.
43. Subramanian J, Simon R. What should physicians look for in evaluating prognostic gene-expression signatures? *Nat Rev Clin Oncol*;7: 327-34.
44. Wirapati P, Sotiriou C, Kunkel S, *et al.* Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res* 2008;10: R65.

Figure legends

Figure 1. Development of a five-gene signature, FILM, effective in predicting the survival of all stages or stage-I lung adenocarcinoma patients. **A.** Derivation and validation of the FILM expression signature. Lung adenocarcinomas from the NCI Director's challenge study were used as a training set (n=442) and adenocarcinoma samples from the Duke and Harvard cohorts were pooled as an adenocarcinoma test set (n=183). Lung SCCs from the studies by Bild *et al* and Raponi *et al* were pooled as an SCC test set (n=183). Gene selection by the prediction algorithms is detailed in the Materials and Methods section. **B.** All stages (n=183, left panel) or stage-I lung adenocarcinomas (n=110, right panel) and lung SCCs (n=183) (**C**) were dichotomized into groups of high (red) versus low (blue) expression of FILM signature expression as predicted by the linear discriminator analysis (LDA) method. Statistically significant differences in survival between the groups were assessed by the Kaplan-Meier method of estimation of survival probability and the long-rank test.

Figure 2. Risk assessed by FILM signature predicts poor survival in all stages or stage-I lung adenocarcinoma. **A.** Development of FILM risk score model. Lung adenocarcinomas from the NCI Director's challenge study were used as training set (n=442) whereas samples from the Duke and Harvard cohorts were pooled as a test set. Patient risk scores were computed as detailed in the Materials and Methods section. **B.** Patients in the training set were dichotomized into high versus low risk groups based on the median (50%) value (dashed line). **C.** Kaplan-Meier plots depicting survival differences between high (black) and low (grey) risk groups in the training set. **D.** Gene Cox coefficients and dichotomization threshold from the training set were directly applied onto all stages and stage-I lung adenocarcinomas in the test set to separate patients into high (black) and low (grey) risk groups.

Figure 3. FILM immunohistochemical protein signature is significantly associated with poor survival. **A.** Representative photomicrographs depicting immunohistochemical expression of the

Lung adenocarcinoma five-gene signature indicated antigens in FFPE histological tissue specimens of lung adenocarcinoma. **B.** Following analysis of FILM protein expression by immunohistochemistry, lung adenocarcinomas (n=156) and SCCs (n=79) were dichotomized into groups with high (red) versus low (blue) expression of FILM protein based on the combined immunoreactivity score of the antigens. For further validation of the FILM protein prognostic signature, all stages (n=156) (**C**) or stage I (n=123) (**D**) lung adenocarcinomas were clustered following median-centering of the antigens' immunoreactivity scores, after which survival differences between the high (red) and low (blue) FILM protein clusters were analyzed.

Figure 4. Mortality risk by FILM protein signature predicts poor survival in non-treated all stages or stage-I lung adenocarcinomas. **A.** Tissue microarray lung adenocarcinomas were randomly divided into a training (n=78) and all stages (n=78) or stage I test set (n=62) and risk scores were computed as detailed in the Materials and Methods section. **B.** Training set lung adenocarcinomas were then dichotomized into high and low risk groups based on the median (50%) value and analyzed for survival differences (**C**). **D.** The univariate Cox coefficients and dichotomization threshold from the training set were directly applied onto the all stage and stage-I lung adenocarcinoma test sets to separate patients similarly into high (black) and low (grey) risk groups which were then analyzed for differences in survival.

Figure 5. FILM protein signature predicts poor survival in non-treated stage-IB lung adenocarcinoma. Non-treated stage IA or stage IB tissue microarray FFPE lung adenocarcinomas were dichotomized into groups of (**A**) high (black) versus low (grey) risk or (**B**) divided into three groups of high (black), intermediate (dotted), low (grey) risk as predicted by the FILM protein signature. Survival differences between the different groups were statistically assessed by the Kaplan-Meier method and log-rank test.

Fig. 1

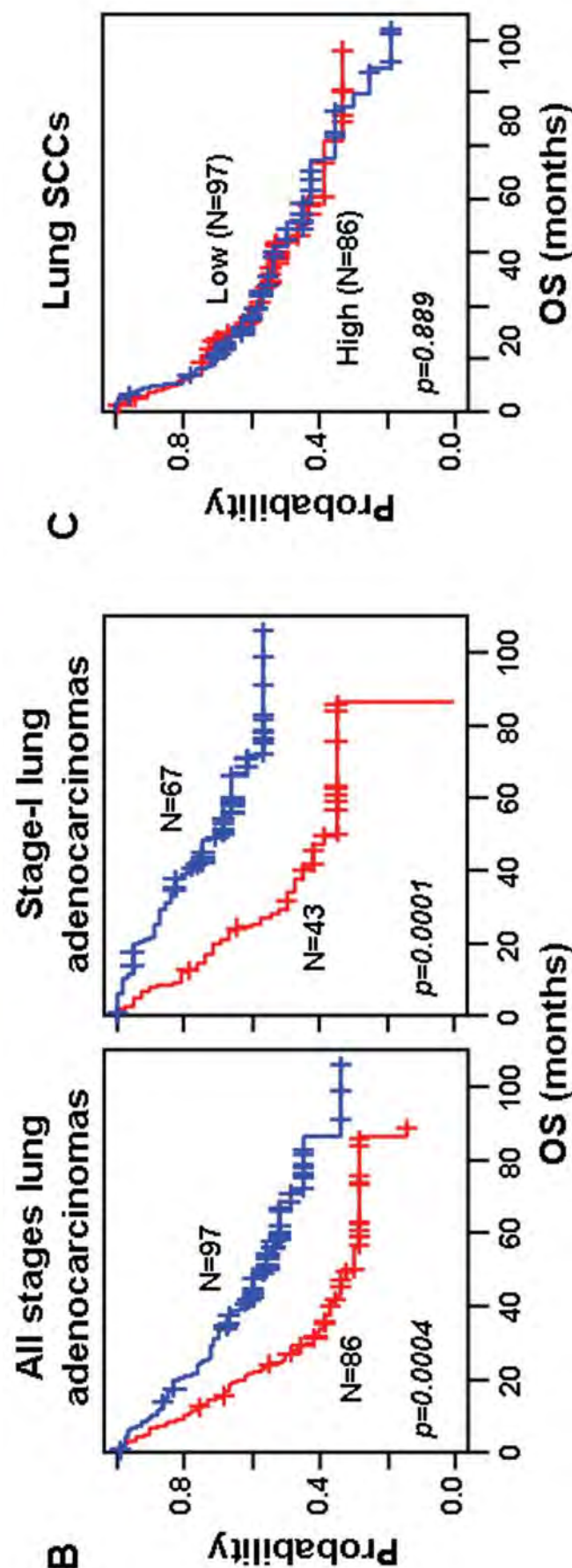
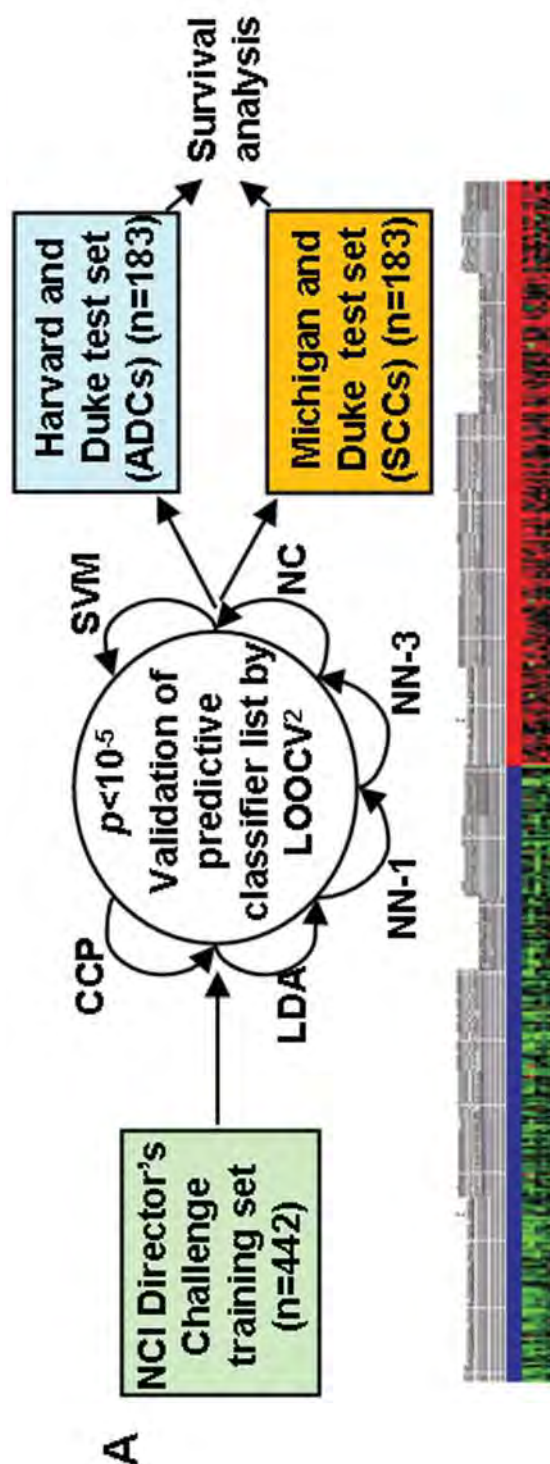
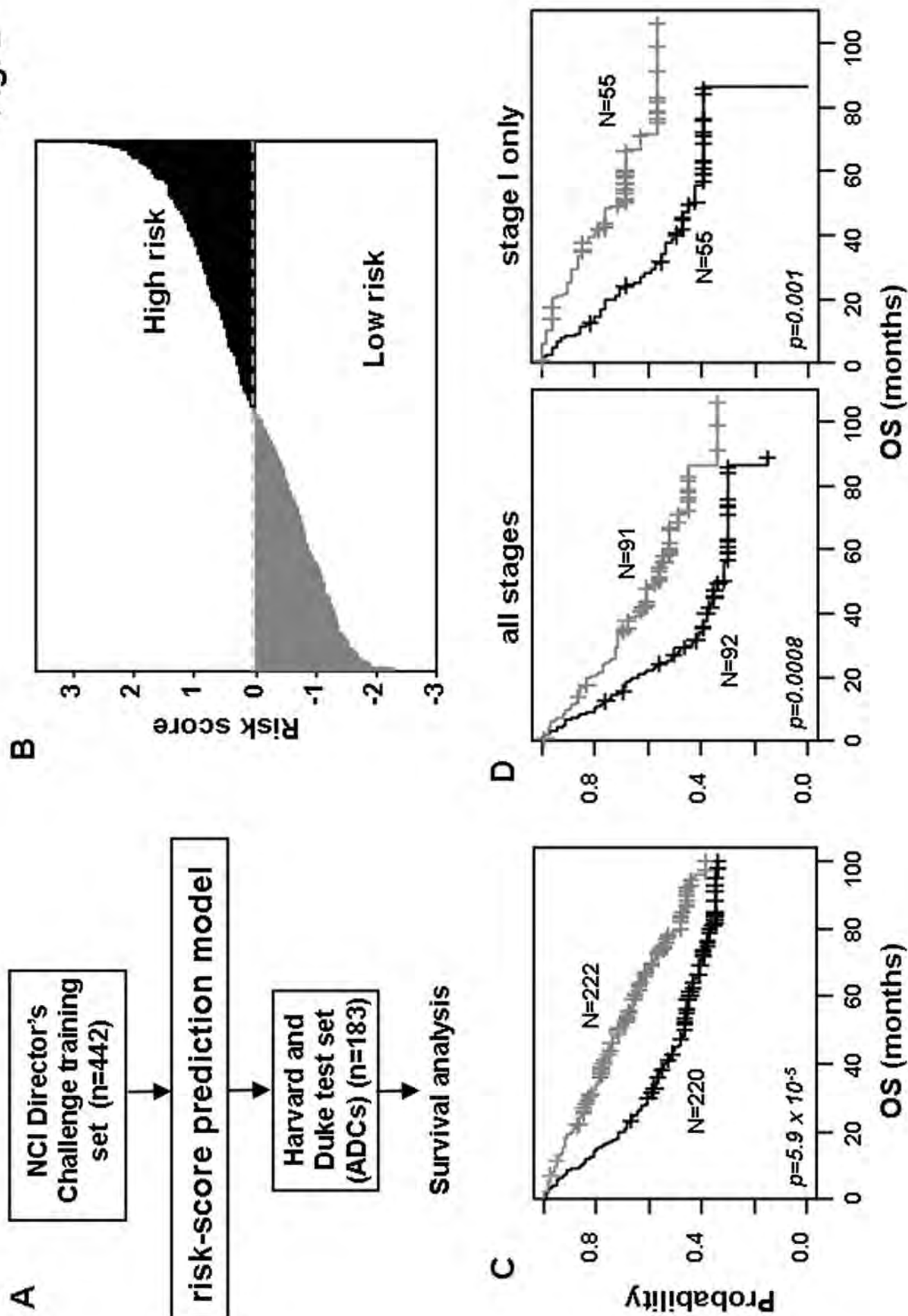


Fig. 2



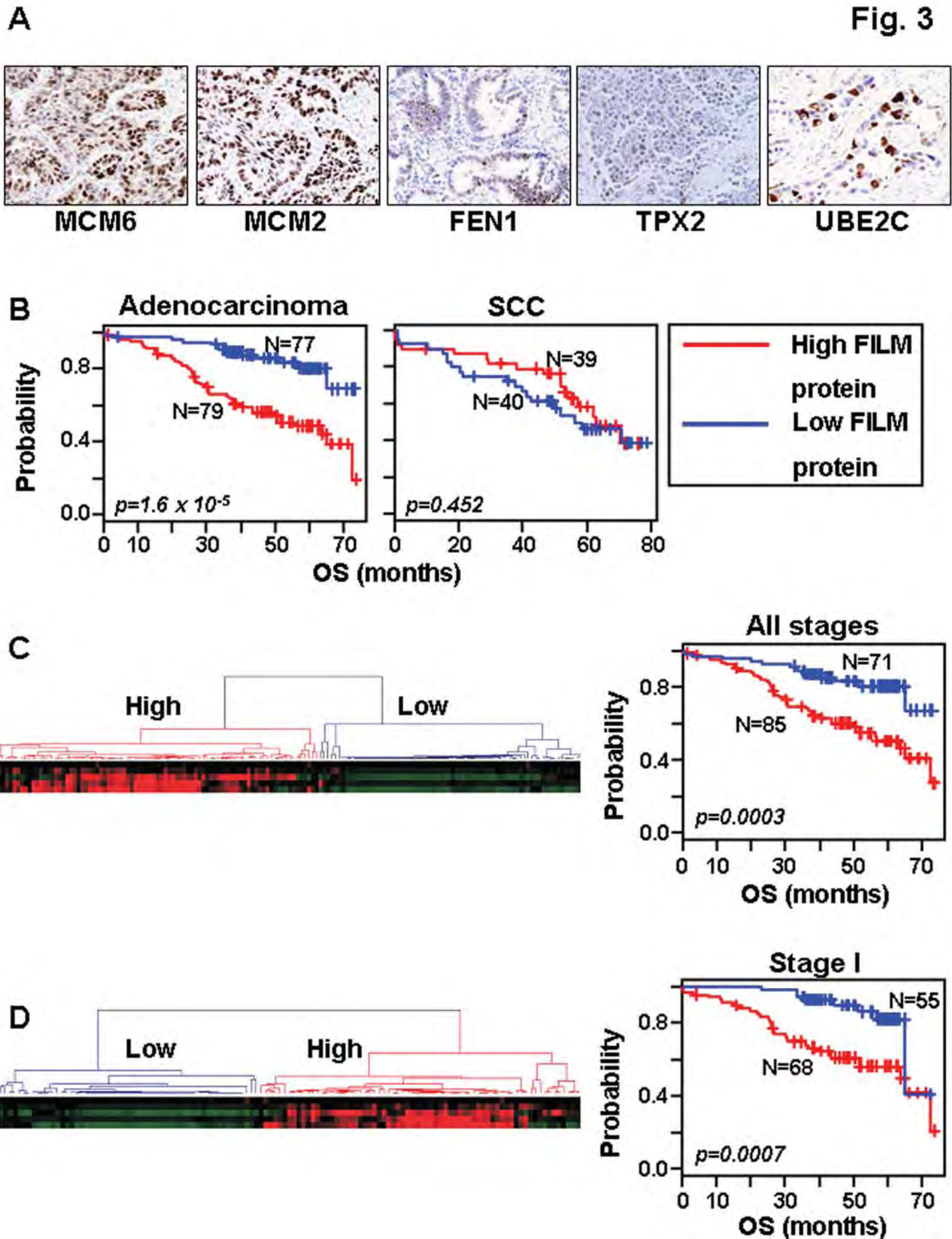


Fig. 4

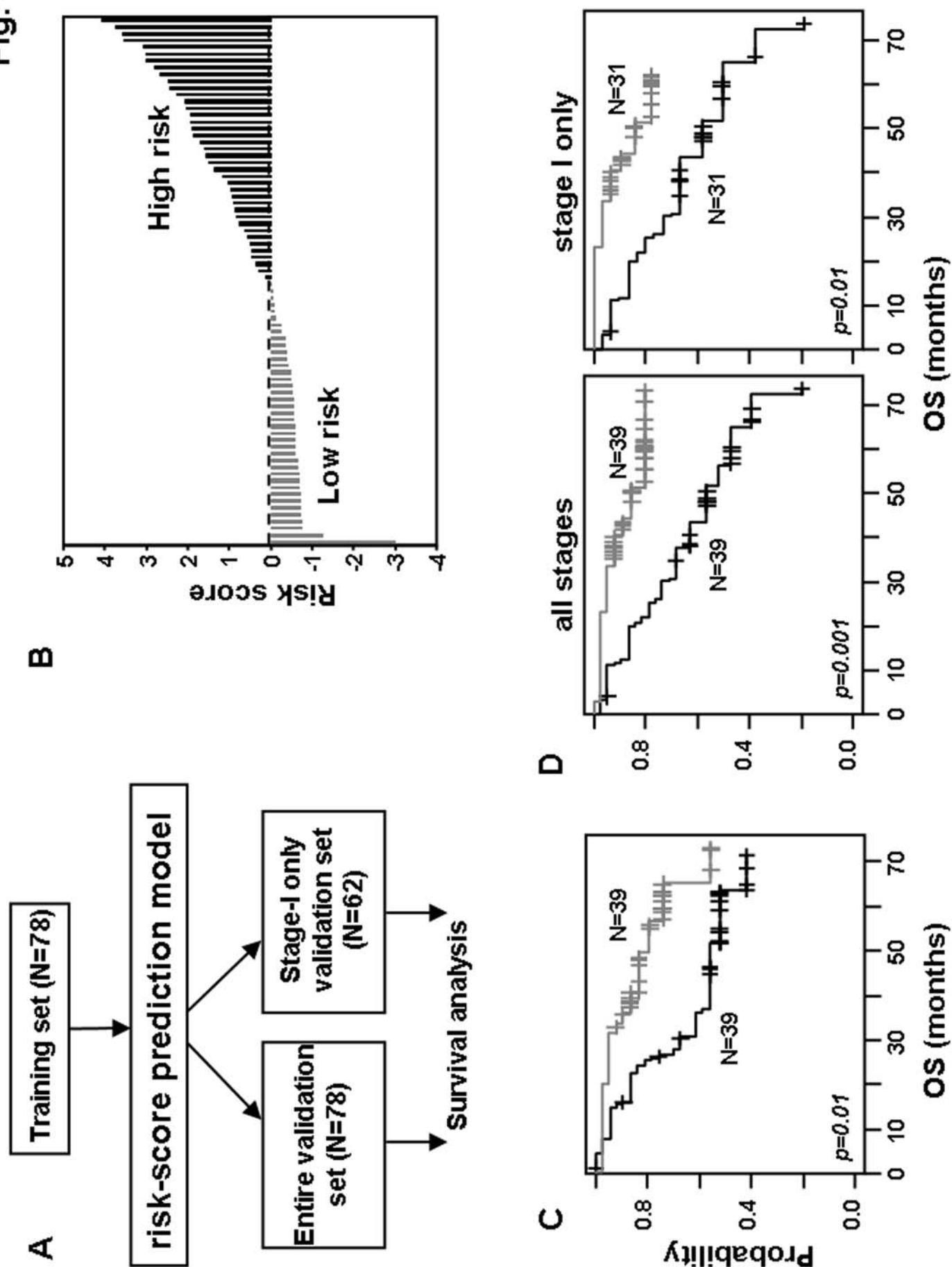


Fig. 5

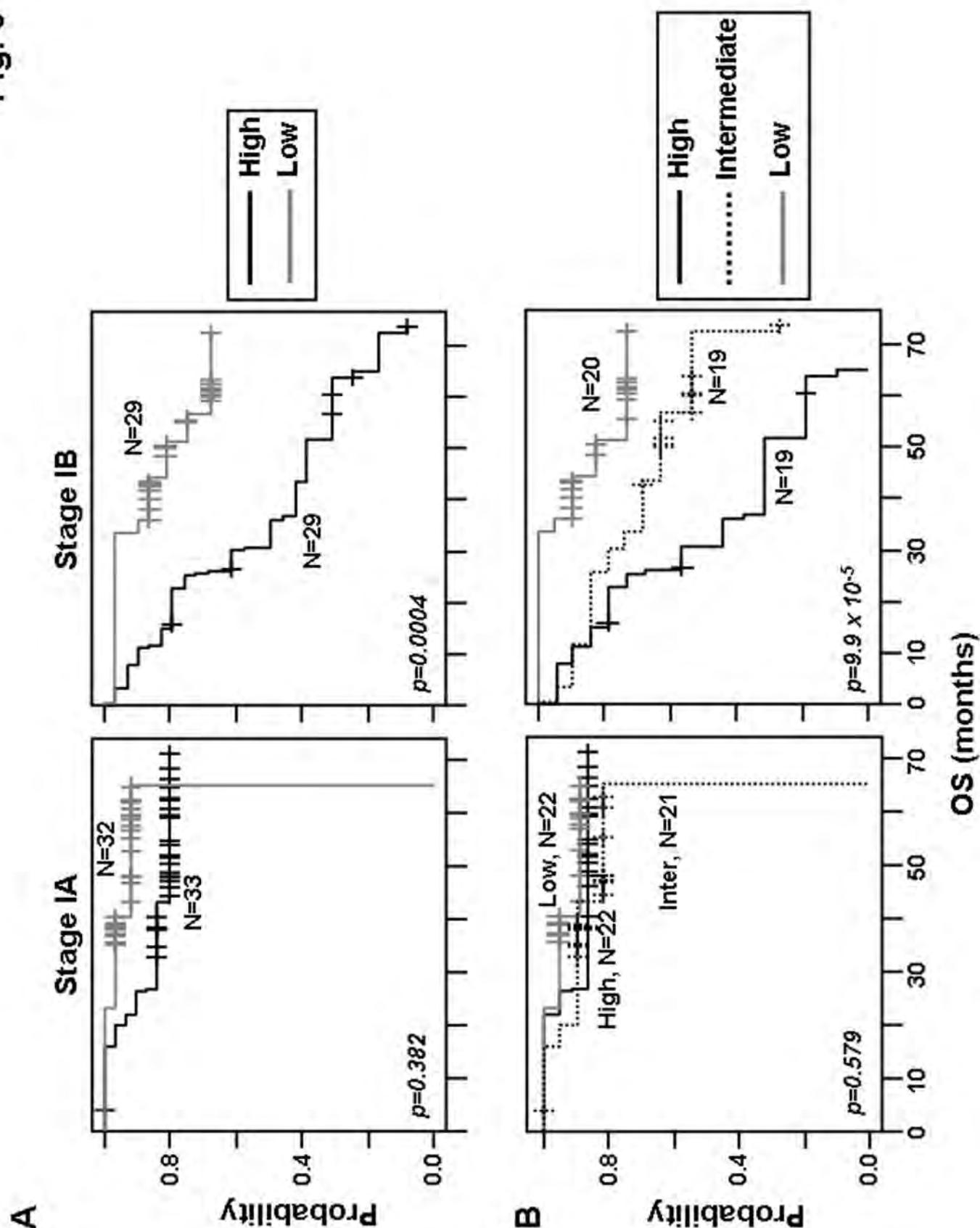


Table 1

FILM protein signature risk is an independent predictor of poor survival in non-treated all stage or stage-I lung adenocarcinoma

All stages				
	Univariate		Multivariate	
	H.R. (95% C.I.)	p-value	H.R. (95% C.I.)	p-value
High risk	3.998 (1.575-9.846)	0.001	5.776 (1.668-20.002)	0.005
Stage III vs I, II or IV	2.505 (0.731-8.583)	0.144	7.011 (1.467-33.506)	0.018
Poor differentiation	1.401 (0.580-3.383)	0.453	0.814 (0.176-3.762)	0.793
Solid subtype	1.007 (0.994-1.020)	0.275	1.003 (0.983-1.023)	0.777
BAC subtype	0.997 (0.983-1.012)	0.71	1.014 (0.995-1.034)	0.14
Never vs ever smoker	0.969 (0.386-2.432)	0.946	1.366 (0.419-4.453)	0.605
Male gender	1.816 (0.823-4.003)	0.139	1.226 (0.491-3.062)	0.662
Age	1.032 (0.987-1.079)	0.161	1.088 (1.031-1.149)	0.002

Stage I				
	Univariate		Multivariate	
	H.R. (95% C.I.)	p-value	H.R. (95% C.I.)	p-value
High risk	3.359 (1.392-9.110)	0.011	6.049 (1.399-26.154)	0.01
Stage IB vs IA	2.556 (0.917-7.124)	0.072	3.154 (0.977-10.191)	0.055
Poor differentiation	2.135 (0.773-5.900)	0.143	1.034 (0.187-5.722)	0.97
Solid subtype	1.010 (0.996-1.025)	0.161	0.997 (0.972-1.026)	0.921
BAC subtype	0.996 (0.998-1.012)	0.646	1.007 (0.988-1.026)	0.459
Never vs ever smoker	0.728 (0.210-2.517)	0.616	0.770 (0.148-4.024)	0.757
Male gender	2.210 (0.871-5.604)	0.095	1.940 (0.597-6.289)	0.269
Age	1.037 (0.984-1.094)	0.173	1.060 (0.990-1.135)	0.09

Univariate and multivariate Cox proportional Hazard regression analysis
 H.R., hazard risk; C.I., confidence interval

Biological Activity of Celecoxib in the Bronchial Epithelium of Current and Former Smokers

Edward S. Kim¹, Waun K. Hong¹, J. Jack Lee², Li Mao¹, Rodolfo C. Morice³, Diane D. Liu², Carlos A. Jimenez³, Georgie A. Eapen³, Reuben Lotan¹, Ximing Tang¹, Robert A. Newman⁴, Ignacio I. Wistuba¹, and Jonathan M. Kurie¹

Abstract

Non-small cell lung cancer is the primary cause of cancer-related death in Western countries. One important approach taken to address this problem is the development of effective chemoprevention strategies. In this study, we examined whether the cyclooxygenase-2 inhibitor celecoxib, as evidenced by decreased cell proliferation, is biologically active in the bronchial epithelium of current and former smokers. Current or former smokers with at least a 20 pack-year (pack-year = number of packs of cigarettes per day times number of years smoked) smoking history were randomized into one of four treatment arms (3-month intervals of celecoxib then placebo, celecoxib then celecoxib, placebo then celecoxib, or placebo then placebo) and underwent bronchoscopies with biopsies at baseline, 3 months, and 6 months. The 204 patients were primarily (79.4%) current smokers: 81 received either low-dose celecoxib or placebo and 123 received either high-dose celecoxib or placebo. Celecoxib was originally administered orally at 200 mg twice daily and the protocol subsequently increased the dose to 400 mg twice daily. The primary end point was change in Ki-67 labeling (from baseline to 3 months) in bronchial epithelium. No cardiac toxicities were observed in the participants. Although the effect of low-dose treatment was not significant, high-dose celecoxib decreased Ki-67 labeling by 3.85% in former smokers and by 1.10% in current smokers—a significantly greater reduction ($P = 0.02$) than that seen with placebo after adjusting for metaplasia and smoking status. A 3- to 6-month celecoxib regimen proved safe to administer. Celecoxib (400 mg twice daily) was biologically active in the bronchial epithelium of current and former smokers; additional studies on the efficacy of celecoxib in non-small cell lung cancer chemoprevention may be warranted. *Cancer Prev Res*; 3(2); 148–59. ©2010 AACR.

Introduction

Non-small cell lung cancer (NSCLC) is the leading cause of death from cancer among both men and women in the United States, accounting for ~28% of such deaths. Indeed, an estimated 160,000 Americans died of NSCLC in 2007. In recent years, the incidence of NSCLC has begun to decline among men (1). However, smoking-related NSCLC has continued to increase among women, surpassing even breast cancer as the leading cause of cancer death in this group (2). Despite aggressive treatment strategies, the 5-year survival rate for NSCLC remains only ~15% (1). These grim facts underscore the urgent need for a change in our approach to NSCLC.

Smoking prevention and cessation have been emphasized as ways to reduce deaths from cancer. Despite the reduction in NSCLC risk observed with smoking cessation, however, several studies have shown that former smokers still have a higher NSCLC risk than nonsmokers have (3–7) and consequently account for a large proportion of NSCLC patients in this country. Chemoprevention strategies, especially for high-risk populations such as former smokers, are appropriate in NSCLC. However, large-scale NSCLC chemoprevention trials, including the Alpha-Tocopherol Beta-Carotene trial, Beta-Carotene and Retinol Efficacy Trial, and Lung Intergroup Trial, have yet to show that any agent can reduce lung cancer risk (8–11).

One of the changes identified in premalignant bronchial tissues that has potential therapeutic significance is an increase in expression of cyclooxygenase-2 (COX-2). COX-2 converts arachidonic acid to prostaglandin H_2 , a precursor of prostaglandin E_2 that has been implicated in a variety of biochemical processes, required for cell proliferation and survival, and whose expression increases in response to growth factors, oncogenes, and carcinogens (12–18). COX-2 has been extensively studied in epithelial tumors, including colorectal cancer and NSCLC (19–22). COX-2 overexpression has prognostic value, predicting a worse

Authors' Affiliations: Departments of ¹Thoracic/Head and Neck Medical Oncology, ²Biostatistics, ³Pulmonary Medicine, and ⁴Pharmacology, The University of Texas M.D. Anderson Cancer Center, Houston, Texas

Corresponding Author: Jonathan M. Kurie, Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M.D. Anderson Cancer Center, Box 432, 1515 Holcombe Boulevard, Houston, TX 77030. Phone: 713-792-6363; Fax: 713-792-1220; E-mail: jkurie@mdanderson.org.

doi: 10.1158/1940-6207.CAPR-09-0233

©2010 American Association for Cancer Research.

outcome in NSCLC patients with stage I disease whose tumors have been surgically resected (23, 24) and thus suggesting that COX-2 is an important biological determinant in NSCLC. Tellingly, COX-2 expression is higher in bronchial premalignant lesions than in adjacent normal lung tissue (22, 25), raising the possibility that COX-2 promotes malignant progression in the lung. Moreover, COX-2 inhibitors have shown efficacy as NSCLC chemopreventive agents in preclinical studies, reducing the size and number of carcinogen-induced NSCLC tumors in mice (26). These findings provide a compelling rationale to investigate the activity of COX-2 inhibitors as chemopreventive agents for lung cancer.

In this study, our goal was to determine whether a 6-month treatment with celecoxib, a COX-2 inhibitor, would be safe and have biological activity in the lungs of current and former smokers. We did a randomized, placebo-controlled study to examine the toxicity and efficacy of celecoxib; bronchial epithelial cell proliferation, as measured by the Ki-67 labeling index after 3 months, was the primary end point. We chose this primary end point on the basis of evidence that bronchial premalignant lesions increase epithelial cellular proliferation and that COX-2 promotes cellular proliferation and survival (27–29).

Materials and Methods

For the present study, we recruited current smokers (those actively smoking or those who had quit within the previous 12 mo) and former smokers (those who had quit at least 12 mo before study entry) who had at least a 20 pack-year history of smoking (pack-year = number of packs of cigarettes per day times number of years smoked). Patients could have had prior stage I NSCLC or stage I or II laryngeal cancer but had to have been free of disease for at least 6 mo before study entry. Other exclusion criteria included the chronic use of steroids, the use of H₂ blockers for active ulcers, the use of nonsteroidal anti-inflammatory drugs other than low-dose aspirin of ≤ 81 mg/d, and a history of stroke, uncontrolled hypertension, and/or angina pectoris. Patients were recruited through local community groups, health fairs, and advertisements distributed to referring practitioners and patients at The University of Texas M.D. Anderson Cancer Center. The study was approved by the Institutional Review Board and by the U.S. Department of Health and Human Services. All patients provided written informed consent before entering the study.

Trial design and treatment

The clinical study design was a four-arm, double-blind, placebo-controlled, randomized study to evaluate the biological effects of celecoxib as a chemopreventive agent in current and former smokers. The primary end point was modulation of Ki-67 in the bronchial epithelium after a 3-mo period of treatment. Patients were treated for up to 6 mo and were randomized onto one of four treatment arms: celecoxib daily for 3 mo, then placebo daily for 3

mo (CCX + PCB); celecoxib daily for 3 mo, then celecoxib daily for 3 mo (CCX + CCX); placebo daily for 3 mo, then celecoxib daily for 3 mo (PCB + CCX); and placebo daily for 3 mo, then placebo daily for 3 mo (PCB + PCB). The research pharmacy randomly assigned each patient to one of the four treatment arms and recorded this assignment by using a computer-generated treatment code that was available only to the pharmacist. Pfizer Corp. provided the 200-mg celecoxib capsules and the matching placebo capsules.

After a bronchoscopy at 3 mo, patients received treatment for an additional 3 mo. A bronchoscopy was then done at the 6-mo time point.

On completing informed consent and enrollment, patients were screened with a chest X-ray and bronchoscopy, which included bronchial washings, brushings, and biopsies from six predetermined sites (carina, right lower, middle, and upper lobes and left lower and upper lobe regions) as well as from any abnormal sites suspicious for cancer. Metaplasia indices (MI) were calculated from the biopsies done at the predetermined sites. The presence of dysplasia was confirmed by histologic evaluation of all biopsy samples. Patients with severe dysplasia at initial or subsequent bronchoscopy were strongly encouraged to undergo additional bronchoscopies at 6 mo.

Patients were stratified for randomization according to smoking history (current versus former), prior cancer (prior versus no prior), and MI ($<15\%$ versus $\geq 15\%$ and/or dysplasia). Toxicity was monitored using the National Cancer Institute Common Toxicity Criteria 2.0, and patients who experienced grade 2 or higher toxicity had their dose reduced. Clinic visits occurred before treatment and during treatment at 1-mo intervals. A complete physical examination, which included asking about the patient's relevant medical history and history of tobacco and alcohol exposure, was done at each clinic visit. Patients were referred to smoking cessation programs on request.

Celecoxib dose

In the original study design, celecoxib was to be administered at 200 mg twice daily. At a scheduled External Advisory Board meeting, the committee raised the concern that in a recently completed colon polyp prevention study (20), a 100-mg dose of celecoxib did not differ from placebo in terms of polyp reduction. On the basis of this updated data, the External Advisory Board recommended a higher dose of celecoxib (400 mg). Therefore, the starting dose level, effective December 2003 (after 81 patients had been enrolled at the low-dose celecoxib level), was set at 400 mg for new patients randomized to receive celecoxib. The first subject was randomized using the new schedule on January 23, 2004.

On December 17, 2004, reports of cardiovascular toxicity in other COX-2 inhibitor trials were released (30–35). At that point, a total of 150 patients had been registered on the current study and 143 had been randomized to treatment. New subject entry and celecoxib treatment were put on hold by the principal investigator of the study and

The M.D. Anderson Cancer Center Institutional Review Board. During the subsequent months, efforts were made to follow-up with participants, audit clinical data and the laboratory specimen inventories, modify the eligibility criteria to exclude patients with preexisting cardiovascular conditions, and include additional procedures to screen and monitor for cardiovascular toxicities. The protocol was amended to address cardiovascular safety issues to ensure the safety of study patients. The M.D. Anderson Institutional Review Board approved the amended protocol, the study was reactivated in April 2005, and we began registering new patients; seven of the patients whose treatment was put on hold reentered the trial. We stopped patient accrual for this trial as of January 2007.

Biopsy specimens

Per the protocol, patients underwent bronchoscopies at the time of enrollment before randomization. All evaluable study patients then had repeat bronchoscopies with biopsies, brushings, and washes at the completion of the first stage of treatment (3 mo) and again at 6 mo. Biopsy, brushing, and wash samples were obtained from the same predetermined sites sampled in the initial bronchoscopy. As noted, these biopsy specimens were taken at six predetermined sites in the bronchial tree: the main carina, the bifurcation of the right upper lobe and the main stem bronchus, the bifurcation of the right middle lobe and right lower lobe, the bifurcation of the left upper lobe and lingula, the medial bronchus of the right lower lobe, and the anterior bronchus of the left lower lobe. We fixed the biopsy specimens in 10% buffered formalin, embedded them in paraffin, and sectioned them. The first two 4- μ m tissue sections from each biopsy site were stained with H&E and evaluated for the presence of squamous metaplasia and dysplasia. We did histologic assessments to determine whether the MI had changed during the 3-mo period. The MI was calculated as the percentage of biopsy sections exhibiting squamous metaplasia out of the total number of sections examined. A single pathologist (X.T.) who was blinded to the study treatment served as the reference pathologist.

We cytologically analyzed sputum samples acquired by sputum induction from all patients before treatment and after 3 and 6 mo of treatment. Additionally, we did buccal brushings on all patients before treatment and after 3 and 6 mo of treatment to look for evidence of tobacco-induced histologic and genetic alterations.

Immunohistochemical analysis of Ki-67

We calculated the fraction of Ki-67-positive cells in the bronchial epithelium, including the basal, parabasal, and superficial layers of the biopsy specimens. Ki-67 labeling indices were expressed as the percentage of cells with positive nuclear staining, as detailed in our prior reports (27, 28). Slides that lacked evaluable epithelium were excluded from the analyses. Ki-67 labeling indices were analyzed on a per-biopsy-site basis and on a per-subject basis (the

average of all biopsy sites that could be evaluated from a participant at a particular time point).

The immunohistochemical analysis was done as follows: one 4- μ m tissue section was deparaffinized in xylene and rehydrated through a series of alcohols. Peroxide blocking was done by immersing the section in 3% hydrogen peroxide in methanol for 15 min. Antigen retrieval was accomplished by placing slides in a steamer for 10 min with 10 mmol/L sodium citrate (pH 6.0) and washing them in Tris buffer. The slides were then blocked in 10% fetal bovine serum for 35 min. The Ki-67 antibody used was MIB-1 (DAKO), and incubation occurred at room temperature at 1:200 dilution for 65 min. Secondary antibody was provided and detection was done using the Envision Link+ kit (DAKO) for 30 min. Diaminobenzidine chromogen was applied for 5 min. The slides were then counterstained with hematoxylin and topped with a coverslip. We used NSCLC cell line pallet sections that had been formalin fixed and paraffin embedded and that evidenced confirmed antigen expressions as the control cell lines.

Statistical design and analysis

This study was designed as a randomized, double-blinded, placebo-controlled trial to evaluate the efficacy and toxicity of celecoxib as a chemopreventive agent in current and former smokers. The planned duration of treatment was a total of 6 mo. The primary end point of the study was modulation of Ki-67, measured after a 3-mo treatment intervention. The secondary end point of the study was the change in Ki-67 labeling at 6 mo. The stratified Z test was applied for comparing the reduction of Ki-67 from baseline to 3 mo between the treatment and placebo groups. The target number of randomized and evaluable patients was 182, which would require a total of 216 randomized patients, allowing for a 15% dropout rate. The study design had at least 80% power to detect a 1.2% difference in the reduction of Ki-67 between the COX-2 inhibitor and placebo, with a two-sided 5% level of significance.

Summary statistics, including frequency, tabulation, mean (and SD), and median (and range), were used to characterize the distribution of Ki-67 labeling indices in the basal layer, parabasal layer, and all layers. The mean Ki-67 index across all six potential biopsy sites was computed with the patient used as the analysis unit. The Wilcoxon rank sum test was used to compare continuous variables between two groups. The Kruskal-Wallis test was used to compare continuous variables among three groups. The χ^2 test or the Fisher's exact test was used to test the statistical significance of the association between two categorical variables. The Wilcoxon signed-rank test was used to test changes in Ki-67 labeling indices by patient before and after treatment within each treatment group. To increase the efficiency of the statistical analysis, we also used the biopsy site as a unit of analysis under the assumption that the site was nested within the patient.

For these analyses, we used a mixed-effect model to test the effect of treatment with celecoxib against placebo on Ki-67 labeling indices, adjusting for covariates that affect Ki-67 levels, such as number of years since smoking cessation (in categories), squamous metaplasia (presence or absence), treatment arm, and time point (0 or 3 mo). When the mixed-effect model was used, a logarithmic transformation was applied to Ki-67 labeling indices to satisfy the Gaussian distribution assumption. All statistical tests were two-sided, with a 5% type I error rate. Statistical analysis was done using standard statistical software, including SAS release 9.1 (SAS Institute) and S-Plus version 7 (Insightful, Inc.).

Results

Patient characteristics

From November 2001 to September 2006, a total of 212 patients registered onto the study, with 204 patients randomized to treatment with either agent or placebo. Eight patients were not randomized to treatment arms: two patients because they declined bronchoscopies, two patients because they had concurrent medical conditions,

and four patients because of the temporary protocol suspension on December 17, 2004.

Of the 204 patients randomized to study sections, 127 patients (61 receiving low-dose celecoxib and 66 receiving high-dose celecoxib) received baseline and 3-month bronchoscopies and thus had data evaluable for the primary end point analysis (Fig. 1). There were 104 patients who received all three (the baseline, 3-month, and 6-month) bronchoscopies. The characteristics of the patients who were randomized to study sections are listed in Table 1. Although the treatment groups generally had similar characteristics, there were fewer women in the arm treated with PCB + CCX ($P = 0.52$), no Hispanic patients in the CCX + PCB arm ($P = 0.06$), and less dysplasia at baseline in the CCX + PCB and PCB + CCX arms ($P = 0.10$).

More patients than expected dropped out of the study because of the temporary protocol suspension, which may explain why the accrual goal of 182 patients with evaluable data was not reached. Common reasons for study dropout in the low-dose celecoxib (200 mg) group included personal reasons (10 patients), being lost to follow-up (8 patients), and concurrent medical conditions (5 patients).

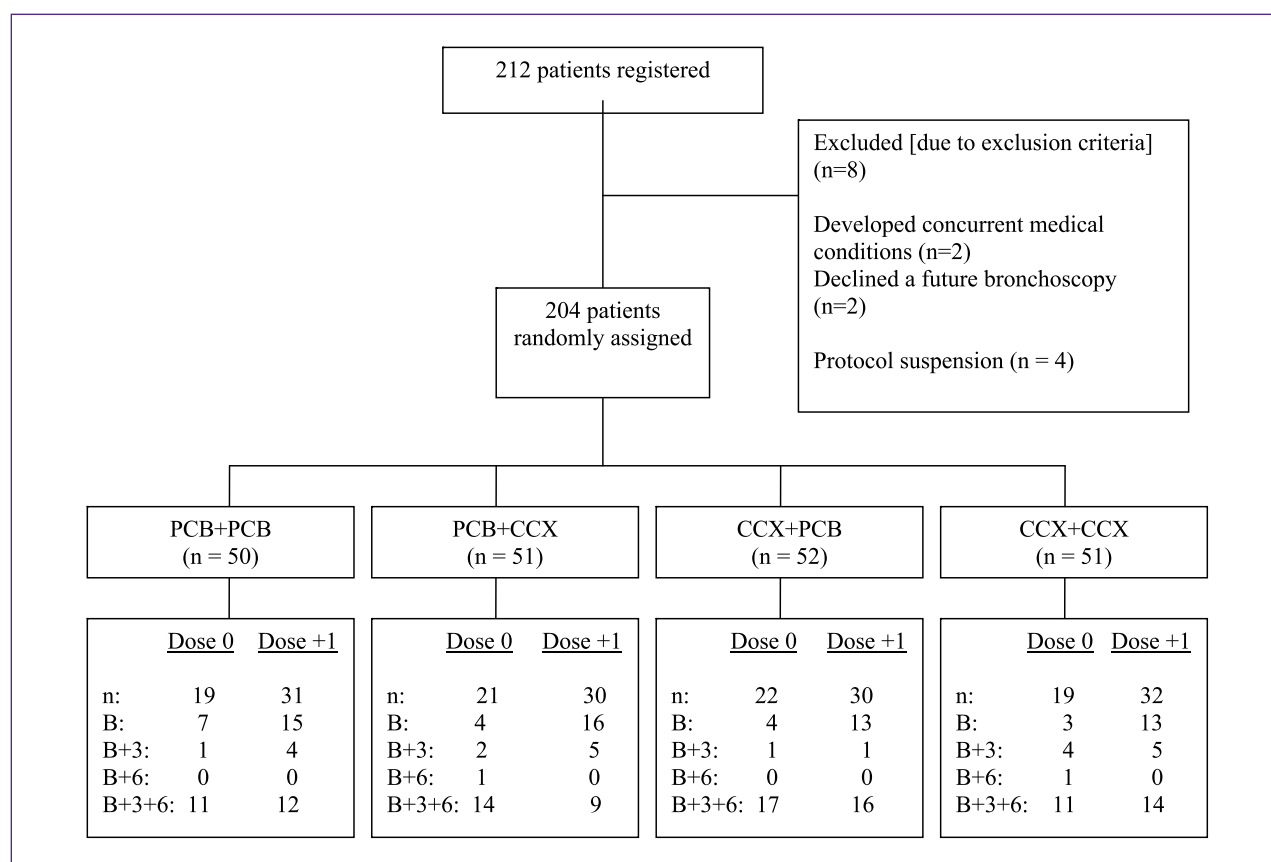


Fig. 1. CONSORT flow diagram of subject accrual into the trial. Patients were randomized to receive the following daily for 3-mo intervals: placebo then placebo (PCB + PCB), placebo then celecoxib (PCB + CCX), celecoxib then placebo (CCX + PCB), or celecoxib then celecoxib (CCX + CCX). Number of patients (n) who completed baseline (B), 3-mo (B+3), and 6-mo (B+6) evaluations are listed. Reasons for leaving the study are also listed.

Table 1. Characteristics of randomized patients by treatment arm

Characteristic	PCB + PCB	PCB + CCX	CCX + PCB	CCX + CCX	Total
Total patients treated	50	51	52	51	204
Age (y)					
Mean $\pm$ SD	53.6 $\pm$ 7.9	52.5 $\pm$ 9.0	54.3 $\pm$ 8.4	53.0 $\pm$ 9.5	53.4 $\pm$ 8.7
Median (range)	52.8 (39.6-70.4)	52.4 (32.9-73.2)	54.9 (39.9-73.6)	52.4 (32.0-71.8)	53.3 (32.0-73.6)
Gender					
Female	25 (50.0%)	20 (39.2%)	26 (50.0%)	27 (52.9%)	98 (48.0%)
Male	25 (50.0%)	31 (60.8%)	26 (50.0%)	24 (47.1%)	106 (52.0%)
Race					
Black	4 (8.9%)	3 (5.9%)	4 (7.7%)	6 (11.8%)	17 (8.3%)
Hispanic	1 (2.0%)	3 (5.9%)	0	5 (9.8%)	9 (4.4%)
White	45 (90.0%)	44 (86.3%)	46 (88.5%)	40 (78.4%)	175 (85.8%)
Other	0	1 (2.0%)	2 (3.8%)	0	3 (1.5%)
Cancer history					
No	46 (92.0%)	45 (88.2%)	43 (82.7%)	48 (94.1%)	182 (89.2%)
Yes	4 (8.0%)	6 (11.8%)	9 (17.3%)	3 (5.9%)	22 (10.8%)
Smoking-related cancer					
No	49 (98.0%)	49 (96.1%)	51 (98.1%)	51 (100%)	200 (98.0%)
Yes	1 (2.0%)	2 (3.9%)	1 (1.9%)	0	4 (2.0%)
Smoking status					
Former smokers	8 (16.0%)	13 (25.5%)	11 (21.2%)	10 (19.6%)	42 (20.6%)
Current smokers	42 (84.0%)	38 (74.5%)	41 (78.8%)	41 (80.4%)	162 (79.4%)
Pack-years, mean $\pm$ SD (range)					
Former smokers	46 $\pm$ 24.4 (21.3-85.2)	39.8 $\pm$ 13.7 (20.6-61.5)	39.8 $\pm$ 18.7 (20.4-73.9)	43.6 $\pm$ 24.2 (20.0-92.6)	41.9 $\pm$ 19.4 (19.9-92.6)
Current smokers	44.0 $\pm$ 14.7 (19.7-81.5)	38.4 $\pm$ 18.2 (20.8-85.6)	47.2 $\pm$ 17.7 (20.0-87.9)	45.0 $\pm$ 22.3 (21.3-132.8)	43.8 $\pm$ 18.5 (19.7-132.8)
Smoking quit-years, mean $\pm$ SD (range)	9.5 $\pm$ 7.2 (1.5-18.1)	7.2 $\pm$ 10.0 (1.0-33.4)	9.9 $\pm$ 9.6 (1.2-35.3)	9.0 $\pm$ 7.9 (1.3-23.8)	8.8 $\pm$ 8.7 (1.0-35.3)
MI, mean $\pm$ SD (range)					
Former smokers	6.3 $\pm$ 12.4 (0-33.3)	2.6 $\pm$ 6.3 (0-16.7)	4.5 $\pm$ 7.8 (0-16.7)	10.0 $\pm$ 14.1 (0-33.3)	5.6 $\pm$ 10.2 (0-33.3)
Current smokers	17.0 $\pm$ 24.3 (0-83.3)	14.9 $\pm$ 18.9 (0-66.7)	14.8 $\pm$ 19.3 (0-66.7)	14.0 $\pm$ 18.5 (0-60.0)	15.2 $\pm$ 20.3 (0-83.3)
Dysplasia					
No	46 (92.0%)	50 (98.0%)	51 (98.1%)	45 (88.2%)	192 (94.1%)
Yes	4 (8.0%)	1 (2.0%)	1 (1.9%)	6 (11.8%)	12 (5.9%)

Abbreviations: PCB, placebo; CCX, celecoxib.

During the protocol suspension, treatment was suspended for 29 patients, all of whom subsequently left the study. Four patients were randomized but never started the study drug because of safety concerns.

Common reasons for dropout in the high-dose celecoxib group (400 mg) included nonadherence as judged by pill counts (12 participants), having or developing concurrent medical conditions (10 participants), and personal reasons (5 participants).

Patients were stratified into statistical groups according to smoking status (current or former smokers). To monitor smoking status, patients were asked at each visit whether they were actively smoking, and serum cotinine levels

were measured. In general, serum cotinine values agreed with the patient reports, but Fig. 2 shows two patients who reported having stopped smoking but had cotinine levels >20 ng/mL at baseline. An additional two former smokers admitted to resuming smoking during the study.

Adherence to treatment

Pill counts were done on a monthly basis to measure treatment adherence, which was excellent. Based on pill counts over the first 3 months of treatment, participants enrolled in the PCB, CCX0, and CCX1 arms took 93.7% (± 15.2), 92.1% (± 11.78), and 94.5% (± 15.3) of the prescribed doses, respectively. Comparable results were

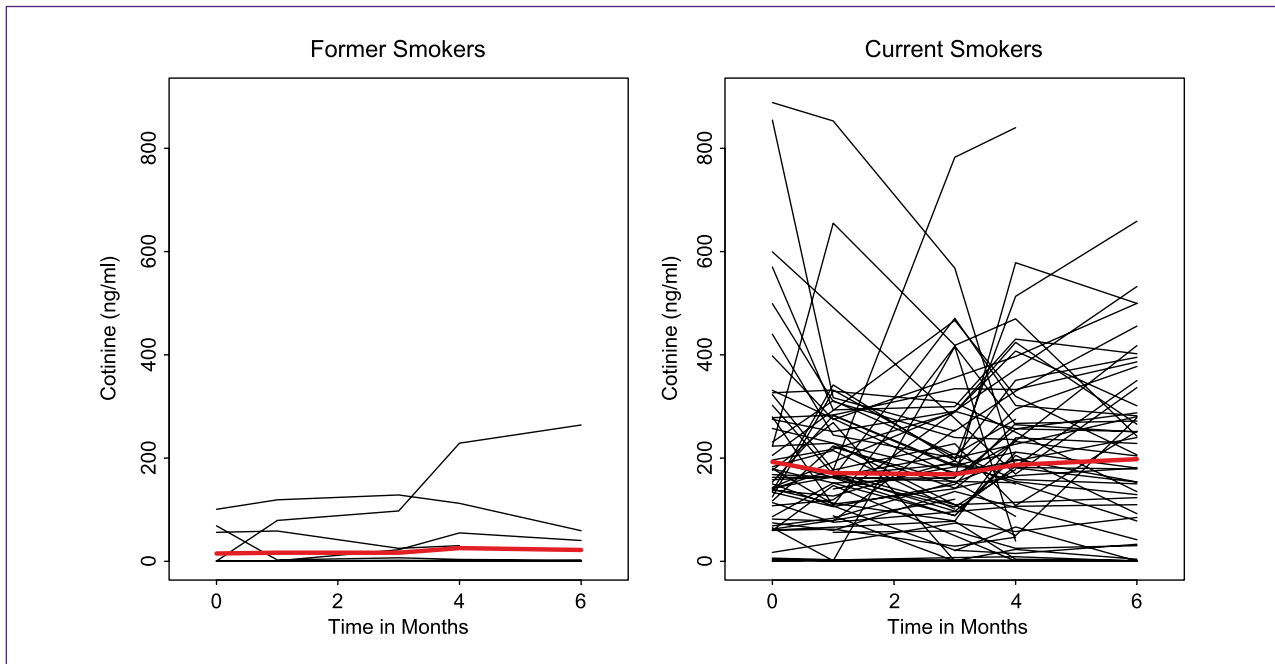


Fig. 2. Cotinine levels by smoking status over time in both former smokers and current smokers. Each black line represents one participant's data. The red line represents the average. The Y axis is the measured cotinine level.

observed at 6 months of follow-up (data not shown). There were no differences in adherence levels between the treatment groups.

Treatment-related toxicity

Of the 204 patients who were randomized to study sections, 92 experienced at least one toxic effect, and a total of

196 toxicity episodes were reported. Fifty-eight patients experienced grade 1 toxicities, 28 patients experienced grade 2 toxicities, and 6 patients reported grade 3 to 4 toxicities (confusion, thrombosis, hyperglycemia, allergic reaction, hypertension, and nausea/abdominal pain), but only hyperglycemia, hypertension, and nausea/abdominal pain were considered to be possibly treatment related (Table 2).

Table 2. Toxic effects of protocol treatment over time period

Treatment group	Time period	No. patients experiencing each grade of toxicity			
		1	2	3	4
PCB + PCB (n = 50)	All	17	7	1	0
PCB + CCX					
Dose 0 (n = 21)	1st 3 mo	6	0	0	0
	2nd 3 mo	4	1	0	0
Dose +1 (n = 30)	1st 3 mo	10	2	1	0
	2nd 3 mo	1	3	2	0
CCX + PCB					
Dose 0 (n = 22)	1st 3 mo	2	0	0	0
	2nd 3 mo	7	1	0	0
Dose +1 (n = 30)	1st 3 mo	5	6	0	0
	2nd 3 mo	2	4	0	0
CCX + CCX					
Dose 0 (n = 19)	All	4	0	0	1
Dose +1 (n = 32)	All	5	4	1	0

Abbreviations: PCB, placebo; CCX, celecoxib; Dose 0, low-dose; Dose +1, high-dose.

Table 3. Celecoxib levels at 3 months by smoking status

Smoking status	Celecoxib levels						
	Treatment	n observed	Mean ($\mu\text{mol/L}$)	SD	Minimum	Median	Maximum
Current	CCX: 0	23	2.11	3.40	0.00	1.26	16.53
	CCX: +1	17	2.71	1.34	0.00	2.43	5.16
Former	CCX: 0	7	2.86	0.79	1.92	2.62	3.98
	CCX: +1	4	3.23	3.08	0.64	2.38	7.50

NOTE: $P = 0.78$ and $P = 0.004$, Wilcoxon rank sum test, comparing celecoxib levels between dose levels in former (2.86 ± 0.79 versus 3.23 ± 3.08) and current (2.11 ± 3.40 versus 2.71 ± 1.34) smoker groups, respectively. $P = 0.008$ and $P = 0.89$, Wilcoxon rank sum test, comparing celecoxib levels between smoking status for dose 0 (2.11 ± 3.40 versus 2.86 ± 0.79) and +1 (2.71 ± 1.34 versus 3.23 ± 3.08) levels, respectively.

Abbreviations: CCX: 0, low-dose celecoxib; CCX: +1, high-dose celecoxib.

No cardiac toxicities were observed in the study. According to protocol guidelines, patients who experienced toxicities of grade 2 or greater had their dose levels reduced to the -1 dose level (100 mg).

Serum celecoxib levels

Serum celecoxib levels were collected via blood draw and measured at baseline and at defined treatment time points (1, 3, 4, and 6 months) before the patient taking the morning dose. At 3 months, mean celecoxib levels were generally in the low micromolar range (Table 3). Serum celecoxib levels were dose dependent in current smokers (2.71 ± 1.34 with high dose versus 2.11 ± 3.40 with low dose; $P = 0.004$, Wilcoxon rank sum test) but not in former smokers (Table 3). On the basis of findings from an *in vitro* study (29), low-micromolar celecoxib levels would be sufficient to have biological effects on NSCLC cells.

Squamous metaplasia and dysplasia in the bronchial epithelium

A total of 212 patients underwent at least one bronchoscopic procedure each, adding up to a sum of 443 bronchoscopic procedures generating 2,658 biopsy samples. Among them, 1,272 biopsy samples were done at baseline, 762 at 3 months of time, and 624 at 6 months of time. Eighteen baseline biopsy samples were inadequate for histologic interpretation. Of the remaining 1,254 baseline samples, 1,086 (86.6%) had normal histology, 152 (12.1%) had squamous metaplasia, and 16 (1.3%) had dysplasia. Squamous metaplasia or dysplasia was detected in 15.5% (148 of 958) of the samples obtained from current smokers and in 5.7% (14 of 248) of the samples obtained from former smokers. The corresponding MI was higher in current smokers [$15.2 (\pm 20.3)$, $n = 162$] than in former smokers [$5.6 (\pm 10.2)$, $n = 42$; $P = 0.004$, Wilcoxon

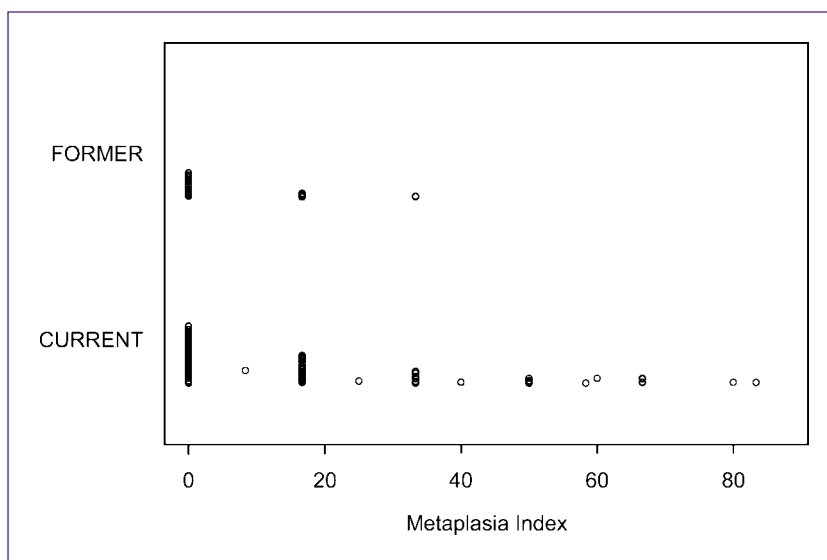


Fig. 3. Baseline squamous metaplasia. Current smokers had a higher percentage of squamous metaplasia than former smokers. Each dot represents one participant's information in relation to MI.

Table 4. Modulation of MI from baseline to 3 mo by treatment arm at each dose level, by smoking status

Smoking status	Treatment	Variable	n	Mean	SD	Minimum	Median	Maximum
Former	CCX, dose 0	MI ₀	8	6.25	8.63	0.00	0.00	16.67
		MI ₃	8	6.25	12.40	0.00	0.00	33.33
		MI ₃₀	8	0.00	12.60	-16.67	0.00	16.67
	CCX, dose +1	MI ₀	7	4.76	8.13	0.00	0.00	16.67
		MI ₃	7	0.00	0.00	0.00	0.00	0.00
		MI ₃₀	7	-4.76	8.13	-16.67	0.00	0.00
	Placebo	MI ₀	13	5.13	10.51	0.00	0.00	33.33
		MI ₃	13	1.28	4.62	0.00	0.00	16.67
		MI ₃₀	13	-3.85	12.08	-33.33	0.00	16.67
Current	CCX, dose 0	MI ₀	25	15.00	21.97	0.00	0.00	66.67
		MI ₃	25	16.13	18.95	0.00	16.67	66.67
		MI ₃₀	25	1.13	14.87	-33.33	0.00	33.33
	CCX, dose +1	MI ₀	29	12.87	14.93	0.00	16.67	50.00
		MI ₃	29	15.92	17.88	0.00	16.67	66.67
		MI ₃₀	29	3.05	18.44	-33.33	0.00	33.33
	Placebo	MI ₀	44	15.45	23.18	0.00	0.00	83.33
		MI ₃	44	16.06	22.45	0.00	8.33	100.00
		MI ₃₀	44	0.61	22.43	-50.00	0.00	66.67

NOTE: One patient did not have a 3-mo MI reading from the biopsy.

Abbreviations: CCX, dose 0, low-dose celecoxib; CCX, dose +1, high-dose celecoxib; MI₀, MI baseline; MI₃, MI at 3 mo; MI₃₀, MI difference (MI₃ - MI₀).

rank sum test; Fig. 3]. There were no differences in MI values among the treatment groups (Table 4).

Effect of treatment on Ki-67 labeling index

The primary end point of the study was modulation of the Ki-67 index from the baseline level after 3 months of treatment. Ki-67 values were measurable in 2,202 biopsy samples (1,069 at baseline, 627 at 3 months of time, and 506 at 6 months of time) obtained from 200 patients randomized to one of the four study groups (Table 5). Wilcoxon rank sum test shows that baseline Ki-67 expression was significantly higher in current than in former smokers among all epithelial layers ($6.15 \pm 6.01\%$ versus $3.86 \pm 5.56\%$; $P = 0.002$), the basal layer ($6.07 \pm 6.77\%$ versus $3.49 \pm 4.69\%$; $P = 0.003$), and the parabasal layer ($9.23 \pm 10.42\%$ versus $6.31 \pm 12.43\%$; $P = 0.009$). Other variables that affected Ki-67 labeling were the presence of squamous metaplasia ($P < 0.0001$) and the number of quit-years (1 to <5, $P = 0.0004$; >5, $P < 0.0001$). We first examined the effect of celecoxib treatment on Ki-67 labeling in all epithelial layers, which was the primary study end point, by combining the low- and high-dose treatment cohorts. Mixed-model analysis revealed that Ki-67 labeling was not significantly different between the celecoxib and placebo groups ($P = 0.12$). However, although the effect of low-dose treatment was not significant ($P = 0.79$), 3 months of high-dose treatment decreased Ki-67 labeling in all epithelial layers in both former smokers (3.85% decrease) and current smokers (1.10% decrease), which was a significantly greater reduction in both groups ($P = 0.02$,

mixed-model analysis) than that in the placebo group after adjusting for metaplasia and smoking status (Table 6; Fig. 4A). This treatment effect persisted at the 6-month time point (Fig. 4B), which further supports the idea that there was a biological effect resulting from high-dose treatment. Additional analysis was then done to examine the effect of high-dose treatment on specific epithelial layers. Although changes in the parabasal layer did not reach significance, Ki-67 labeling decreased in the basal layer by 4.14% in former smokers and 1.41% in current smokers, which was a significantly greater reduction than that observed in the placebo arm ($P = 0.008$, mixed-model analysis).

Discussion

In this first-ever randomized clinical trial of a 6-month celecoxib regimen in current and former smokers, we found that celecoxib is safe to administer and biologically active in the bronchial epithelium. The effects of treatment on the primary end point, bronchial epithelial proliferation after 3 months of time, are noteworthy given that the participant accrual goal was not reached. Moreover, the biological activity and safety of celecoxib in this cohort warrant additional studies on the efficacy of celecoxib in NSCLC chemoprevention.

Problems encountered during the conduct of this trial highlight several important feasibility issues in planning lung chemoprevention studies. The unanticipated cardiac toxicities reported in large trials examining the efficacy of

Table 5. Distribution of Ki-67 index (all layers) in patients by smoking status and treatment at baseline, 3 mo, and difference from baseline to 3 mo ($n = 200$)

Smoking status	Treatment	Variable	<i>n</i>	Mean	SD	Minimum	Median	Maximum	<i>P</i> *
Former	CCX, dose 0	Baseline	11	2.90	2.36	0.50	2.69	8.63	0.55
		3 mo	8	3.85	4.57	0.17	1.31	11.06	
		Difference	8	1.04	5.43	-8.19	0.55	9.88	
	CCX, dose +1	Baseline	10	4.89	7.68	0.93	1.87	26.24	0.38
		3 mo	7	2.47	2.75	0.10	1.62	8.18	
		Difference	7	-3.85	10.13	-25.6	-1.79	6.32	
	Placebo	Baseline	18	3.88	5.77	0.08	1.50	23.81	0.95
		3 mo	13	2.80	2.38	0.14	2.23	7.45	
		Difference	13	-1.23	6.16	-20.1	0.03	6.89	
Current	CCX, dose 0	Baseline	30	6.62	7.37	0.38	4.16	30.89	0.97
		3 mo	25	6.48	4.89	0.25	5.46	16.23	
		Difference	25	0.06	5.22	-11.5	-0.68	12.97	
	CCX, dose +1	Baseline	50	6.03	4.84	0.35	4.59	23.66	0.27
		3 mo	29	6.76	5.88	0.00	5.43	19.82	
		Difference	28	-1.10	6.91	-13.5	-1.71	14.12	
	Placebo	Baseline	80	6.05	6.18	0.00	4.65	32.13	0.89
		3 mo	44	7.66	7.10	0.00	6.15	38.10	
		Difference	44	0.15	8.58	-26.4	-0.15	34.42	

NOTE: One patient did not have a baseline Ki-67 reading due to inadequate tissue.

*Wilcoxon signed-rank test comparing modulation of Ki-67 index within each subgroup.

celecoxib and other nonsteroidal anti-inflammatory drugs in colon cancer chemoprevention (30–35) negatively affected the conduct of this study in several respects. First, participant accrual was interrupted for 6 months. Second, data from patients who had been actively receiving treatment at the time of protocol suspension were deemed inevaluable due to early treatment cessation. Third, patient accrual after the trial reopened proceeded at a slower rate than it had before trial suspension, which suggests that the negative publicity associated with the cardiac toxicity re-

ports adversely affected patient accrual. In fact, we did not observe any cardiovascular toxicity in this cohort. This may have been related to the short duration of celecoxib treatment in this study relative to that of the trials reporting these toxicities, which required treatments of more than 12 months of duration (20, 30, 36–38).

The findings reported here on Ki-67 labeling in the bronchial epithelium are noteworthy for several reasons. First, Ki-67 labeling decreased in participants treated with high-dose but not low-dose celecoxib. Similarly, celecoxib

Table 6. Mixed-model analysis on the effects of Ki-67 ($n = 200$)

Covariates	Estimate	SE	<i>P</i>
SQM (+ vs -)	0.62	0.05	<0.0001
Quit-years			
1 to <5 y vs current smokers	-0.21	0.06	0.0004
≥5 y vs current smokers	-0.37	0.06	<0.0001
Treatment			
CCX (dose 0 vs placebo)	0.09	0.05	0.12
CCX (dose +1 vs placebo)	0.10	0.05	0.03
Time (3 mo vs baseline)	0.09	0.05	0.049
Treatment time			
CCX (dose 0 vs placebo at 3 mo)	-0.02	0.08	0.79
CCX (dose +1 vs placebo at 3 mo)	-0.17	0.07	0.02

Abbreviation: SQM, any squamous metaplasia.

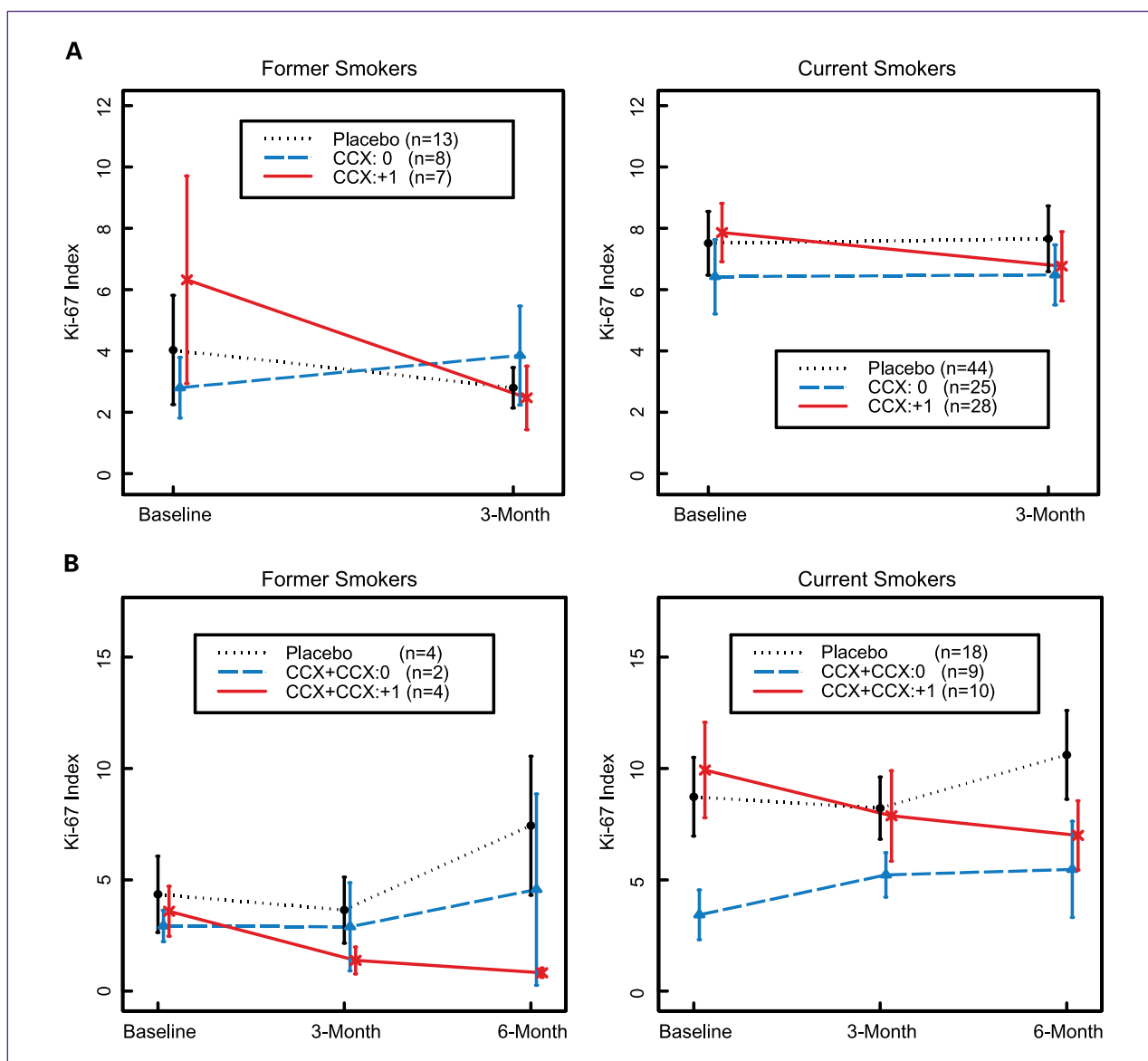


Fig. 4. Mean Ki-67 over time in all layers. A, baseline and 3-mo time points show decreasing expression of Ki-67 with high-dose celecoxib over time in both current and former smokers who had both baseline and 3-mo Ki-67 measurements. Total evaluable patients are 28 in former smoker group and 97 in current smoker group. Y axis, Ki-67 index. B, baseline, 3-mo, and 6-mo time periods show a similar trend for Ki-67 expression with high-dose celecoxib in both current and former smokers who had baseline, 3-mo, and 6-mo Ki-67 measurements. Total evaluable patients are 10 in former smoker group and 37 in current smoker group. Placebo and low-dose celecoxib follow similar patterns, especially in current smokers. Y axis, Ki-67 index.

is more efficacious in colon cancer chemoprevention when administered at 400 mg twice daily than at 200 mg twice daily (20, 30, 38). These findings suggest that, although it may have less treatment-related toxicity, low-dose (200 mg) celecoxib has no efficacy in NSCLC chemoprevention and argue against such a trial design.

Second, Ki-67 levels decreased more prominently in former smokers than in current smokers, especially in those patients who completed baseline and 3-month bronchoscopies (Fig. 4A). This effect was also observed in patients who completed baseline, 3-month, and 6-month bronch-

oscopies (Fig. 4B) and received high-dose celecoxib. It is important to note that these are subset analyses and the number of patients is low, especially in the former smokers. Serum celecoxib levels did not differ in current versus former smokers treated with high-dose celecoxib (data not shown), but detailed pharmacokinetic studies were not done, so we cannot exclude the possibility that the pharmacokinetics of celecoxib contributed to this outcome. Current and former smokers may differ with respect to the role that COX-2 plays in maintaining bronchial epithelial proliferation. In fact, other studies have reported

differences between these two groups with respect to bronchial epithelial biology (28, 39, 40).

Third, the reduction in Ki-67 labeling was not accompanied by a decrease in MI, and the effect of celecoxib on Ki-67 did not vary on the basis of histology, indicating that the decrease in Ki-67 was not due to a reduction in bronchial metaplasia, which has been reported to increase Ki-67 labeling (28, 41–43). Dysplasia was uncommon in this cohort, so no conclusions can be made about the effect of celecoxib on this histologic abnormality.

Fourth, Ki-67 decreased more prominently in the basal layer than it did in the parabasal layer, a strikingly different finding from those reported in chemoprevention studies using retinoids, which reduce bronchial metaplasia and are active primarily in the parabasal layer of the bronchial epithelium (27, 28, 44). Collectively, these findings suggest that the basal and parabasal compartments of the bronchial epithelium are biologically distinct, which is consistent with evidence that cells in the basal layer have a low proliferation rate, express progenitor cell markers, and have multipotent differentiation potential (45), whereas cells in the parabasal layer have a higher proliferation rate and have undergone differentiation into mucous-secreting and other epithelial cell types.

Progress in the field of NSCLC chemoprevention research will require the ability to identify individuals at high risk for the development of NSCLC, a way to isolate premalignant bronchial epithelial cells in danger of malignant progression, and a method to elucidate the mechanisms by which these premalignant cells maintain their proliferation and survival. With respect to the latter, findings presented here and elsewhere (46) raise the possibility that COX-2 is one mediator of bronchial epithelial proliferation in current and former smokers. Additional studies are warranted to examine the importance of COX-2 in NSCLC development and explore NSCLC prevention with COX-2 inhibitors.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

National Cancer Institute grants CA 091844 and CA 16672 (J.M. Kurie), Department of Defense grant W81XWH-04-0142 (W.K. Hong), and Pfizer Pharmaceuticals.

Received 10/28/09; accepted 11/3/09; published OnlineFirst 1/26/10.

References

- Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. *CA Cancer J Clin* 2008;58:71–96.
- U.S. Department of Health and Human Services. The health benefits of smoking cessation: a report of the Surgeon General. Atlanta (GA): U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 1990. DHHS Publication No. (CDC) 90-8416.
- Halpern MT, Gillespie BW, Warner KE. Patterns of absolute risk of NSCLC mortality in former smokers. *J Natl Cancer Inst* 1993;85:457–64.
- Burns DM. Primary prevention, smoking, and smoking cessation: implications for future trends in NSCLC prevention. *Cancer* 2000;89:2506–9.
- Gottfredsen NS, Prescott E, Osler M. Effect of smoking reduction on NSCLC risk. *JAMA* 2005;294:1505–10.
- Lubin JH, Blot WJ. NSCLC and smoking cessation: patterns of risk. *J Natl Cancer Inst* 1993;85:422–3.
- Tong L, Spitz MR, Fueger JJ, Amos CA. Lung carcinoma in former smokers. *Cancer* 1996;78:1004–10.
- Kim ES, Hong WK, Khuri FR. Chemoprevention of aerodigestive tract cancers. *Annu Rev Med* 2002;53:223–43.
- The α -Tocopherol, β -Carotene Cancer Prevention Study Group. The effect of vitamin E and β carotene on the incidence of NSCLC and other cancers in male smokers. *N Engl J Med* 1994;330:1029–35.
- Omenn GS, Goodman GE, Thornquist MD, et al. Effects of a combination of β carotene and vitamin A on NSCLC and cardiovascular disease. *N Engl J Med* 1996;334:1150–5.
- Lippman SM, Lee JJ, Karp DD, et al. Randomized phase III intergroup trial of isotretinoin to prevent second primary tumors in stage I NSCLC. *J Natl Cancer Inst* 2001;93:605–18.
- Herschman HR. Prostaglandin synthase 2. *Biochim Biophys Acta* 1996;1299:125–40.
- Subbaramiah K, Telang N, Ramonetti JT, et al. Transcription of cyclooxygenase 2 is enhanced in transformed mammary epithelial cells. *Cancer Res* 1996;56:4424–9.
- Kelly DJ, Mestre JR, Subbaramiah K, et al. Benzo[a]pyrene upregulates cyclooxygenase-2 gene expression in oral epithelial cells. *Carcinogenesis* 1997;18:795–9.
- Kutcher W, Jones DA, Matsunami N, et al. Prostaglandin H synthase-2 is expressed abnormally in human colon cancer: evidence for a transcriptional effect. *Proc Natl Acad Sci U S A* 1996;93:4816–20.
- Ebert CE, Coffey RJ, Radhika A, Giardiello FM, Ferrenbach S, DuBois RN. Up-regulation of cyclooxygenase 2 expression in human colorectal adenomas and adenocarcinomas. *Gastroenterology* 1994;107:1183–8.
- Ristimäki A, Honkanen N, Jankala H, Sipponen P, Harkonen M. Expression of cyclooxygenase-2 in human gastric carcinoma. *Cancer Res* 1997;57:1276–80.
- Hida T, Yatabe Y, Achiwa H, et al. Increased expression of cyclooxygenase 2 occurs frequently in human NSCLCs, specifically in adenocarcinomas. *Cancer Res* 1998;58:3761–4.
- Oshima M, Dinchuk SL, Kargman H, et al. Suppression of intestinal polyposis in *Apc*^{Δ176} knockout mice by inhibition of cyclooxygenase 2 (COX-2). *Cell* 1996;87:803–9.
- Steinbach G, Lynch PM, Phillips RK, et al. The effect of celecoxib, a cyclooxygenase-2 inhibitor, in familial adenomatous polyposis. *N Engl J Med* 2000;342:1946–52.
- Wolff H, Saukkonen K, Anttila S, Karjalainen A, Vainio H, Ristimäki A. Expression of cyclooxygenase-2 in human lung carcinoma. *Cancer Res* 1998;58:4997–5001.
- Koki AT, Khan NK, Woerner BM, et al. Characterization of cyclooxygenase-2 (COX-2) during tumorigenesis in human epithelial cancers: evidence for potential clinical utility of COX-2 inhibitors in epithelial cancers. *Prostaglandins Leukot Essent Fatty Acids* 2002;66:13–8.
- Lu C, Soria JC, Tang X, et al. Prognostic factors in resected stage I non-small-cell NSCLC: a multivariate analysis of six molecular markers. *J Clin Oncol* 2004;22:4575–83.

24. Khuri FR, Wu H, Lee JJ, et al. Cyclooxygenase-2 overexpression is a marker of poor prognosis in stage I non-small cell NSCLC. *Clin Cancer Res* 2001;7:861–7.
25. Hosomi Y, Yokose T, Hirose Y, et al. Increased cyclooxygenase 2 (COX-2) expression occurs frequently in precursor lesions of human adenocarcinoma of the lung. *Lung Cancer* 2000;30:73–81.
26. Rioux N, Castonguay A. Prevention of NNK-induced lung tumorigenesis in A/J mice by acetylsalicylic acid and NS-398. *Cancer Res* 1998;58:5354–60.
27. Hittelman WN, Liu DD, Kurie JM, et al. Proliferative changes in the bronchial epithelium of former smokers treated with retinoids. *J Natl Cancer Inst* 2007;99:1603–12.
28. Lee JJ, Liu D, Lee JS, et al. Long-term impact of smoking on lung epithelial proliferation in current and former smokers. *J Natl Cancer Inst* 2001;93:1081–8.
29. Hida T, Kozaki K, Muramatsu H, et al. Cyclooxygenase-2 inhibitor induces apoptosis and enhances cytotoxicity of various anticancer agents in non-small cell NSCLC cell lines. *Clin Cancer Res* 2000;6:2006–11.
30. Bertagnolli MM, Eagle CJ, Zauber AG, et al. APC Study Investigators. Celecoxib for the prevention of sporadic colorectal adenomas. *N Engl J Med* 2006;355:873–84.
31. Solomon SD, McMurray JV, Pfeffer MA, et al. Cardiovascular risk associated with celecoxib in a clinical trial for colorectal adenoma prevention. *N Engl J Med* 2005;352:1071–80.
32. Bresalier RS, Sandler RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. *N Engl J Med* 2005;352:1092–102.
33. Nussmeier NA, Whelton AA, Brown MT, et al. Complications of the COX-2 inhibitors parecoxib and valdecoxib after cardiac surgery. *N Engl J Med* 2005;352:1081–91.
34. Fitzgerald GA. Coxibs and cardiovascular disease. *N Engl J Med* 2004;351:1709–11.
35. Solomon SD, Wittes J, Finn PV, et al. Cross Trial Safety Assessment Group. Cardiovascular risk of celecoxib in 6 randomized placebo-controlled trials: the cross trial safety analysis. *Circulation* 2008;117:2104–13.
36. Arber N, Eagle CJ, Spicak J, et al. Celecoxib for the prevention of colorectal adenomatous polyps. *N Engl J Med* 2006;355:885–95.
37. Baron JA, Sandler RS, Bresalier RS, et al. A randomized trial of rofecoxib for the chemoprevention of colorectal adenomas. *Gastroenterology* 2006;131:1674–82.
38. Bertagnolli MM, Zauber AG, Hawk ET. The Adenoma Prevention with Celecoxib (APC) trial: five-year efficacy and safety results [abstract LB-141]. *Proceedings of the 99th Annual Meeting of the American Association for Cancer Research*. 2008.
39. Spira A, Beane J, Shah V, et al. Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci U S A* 2004;101:10143–8.
40. Zhang L, Lee JJ, Tang H, et al. Impact of smoking cessation on global gene expression in the bronchial epithelium of chronic smokers. *Cancer Prev Res* 2008;1:112–8.
41. Martin B, Paesmans M, Mascaux C, et al. Ki-67 expression and patients' survival in NSCLC: systematic review of the literature with meta-analysis. *Br J Cancer* 2004;91:2018–25.
42. Miller YE, Blatchford P, Hyun DS, et al. Bronchial epithelial Ki-67 index is related to histology, smoking, and gender, but not NSCLC or chronic obstructive pulmonary disease. *Cancer Epidemiol Biomarkers Prev* 2007;16:2425–31.
43. Szabo E. Lung epithelial proliferation: a biomarker for chemoprevention trials? *J Natl Cancer Inst* 2001;93:1042–43.
44. Kurie JM, Lotan R, Lee JJ, et al. Treatment of former smokers with 9-*cis*-retinoic acid reverses loss of retinoic acid receptor- β expression in the bronchial epithelium: results from a randomized placebo-controlled trial. *J Natl Cancer Inst* 2003;95:206–14.
45. Vaughan MB, Ramirez RD, Wright WE, et al. A three-dimensional model of differentiation of immortalized human bronchial epithelial cells. *Differentiation* 2006;74:141–8.
46. Mao JT, Fishbein MC, Adams B, et al. Celecoxib decreases Ki-67 proliferative index in active smokers. *Clin Cancer Res* 2006;12:314–20.

Prevention of Bronchial Hyperplasia by EGFR Pathway Inhibitors in an Organotypic Culture Model

Jangsoon Lee^{1,4}, Seung-Hee Ryu^{1,5}, Shin Myung Kang^{1,6}, Wen-Cheng Chung¹, Kathryn Ann Gold², Edward S. Kim¹, Walter N. Hittelman³, Waun Ki Hong¹, Ja Seok Koo¹

¹Department of Thoracic/Head and Neck Medical Oncology, ²Hematology and Oncology, ³Experimental Therapeutics, The University of Texas M. D. Anderson Cancer Center, Houston, Texas.

Grant support: Department of Defense VITAL grant W81XW-04-1-0142 (J.S. Koo, E.S. Kim, W.K. Hong), National Heart, Lung and Blood Institute grant R01-HL-077556 (J.S. Koo), and NCI Cancer Center Support grant CA-16672 (The University of Texas M. D. Anderson Cancer Center).

Corresponding Author: Ja Seok Koo, Department of Thoracic/Head and Neck Medical Oncology, Unit 432, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030. Phone: 713-792-8454; Fax: 713-794-5997; E-mail: jskoo@mdanderson.org

Note: Jangsoon Lee and Seung-Hee Ryu contributed equally to this work.

Present address: ⁴Breast Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas, ⁵Department National Investment Project, Asan Institute for Life Sciences, Asan Medical Center, Seoul, Korea, ⁶Center for Lung Cancer, National Cancer Center, Goyang-si, Korea.

Running title: Prevention of Bronchial Hyperplasia.

Keywords: NHBE, Bronchial hyperplasia, Dysplasia, Erlotinib, MEK inhibitor

Disclosure of Potential Conflicts of Interest.

No potential conflicts of interest were disclosed.

Abstract

Lung cancer is the leading cause of cancer-related mortality worldwide. Early detection or prevention strategies are urgently needed to increase survival. Hyperplasia is the first morphological change that occurs in the bronchial epithelium during lung cancer development, followed by squamous metaplasia, dysplasia, carcinoma in situ and invasive tumor. Current study was designed to understand molecular mechanisms that control the hyperplasia of the bronchial epithelium. Using primary normal human tracheobronchial epithelial (NHTBE) cells cultured by the 3-dimensional organotypic method, we found that the epidermal growth factor receptor (EGFR) ligands epidermal growth factor (EGF), transforming growth factor- α (TGF- α), and amphiregulin (AR) induce hyperplasia as determined by cell proliferation and formation of multilayered epithelium. We also found that EGF induced the increased expression of cyclin D1, which plays a critical role in bronchial hyperplasia, and the overexpression was mediated by activating the mitogen-activated protein kinase (MAPK) pathway but not by the phosphoinositide 3-kinase/Akt (PI3-K/Akt) signaling pathway. Erlotinib, an EGFR tyrosine kinase inhibitor, and U0126, a MEK inhibitor, completely inhibited EGF-induced hyperplasia. Furthermore, promoter analysis revealed that the activator protein-1 (AP-1) transcription factor regulates EGF-induced cyclin D1 overexpression. Depletion of AP-1 using siRNA targeting its c-Jun component completely abrogated EGF-induced cyclin D1 expression. In conclusion, we demonstrated that bronchial hyperplasia can be modeled *in vitro* using primary NHTBE cells maintained in a 3-dimensional (3-D) organotypic culture. Inhibitors of EGFR and MEK completely blocked EGF-induced bronchial hyperplasia, suggesting a potential chemopreventive role of these inhibitors.

Introduction

Hyperplasia, as evidenced by increased cell proliferation, in the bronchial epithelium is associated with various clinical settings, such as trauma, smoking, chronic cough, chronic inflammatory airway disease, and cancer. It is the first of several progressive, cumulative genetic and morphological changes associated with development of squamous cell carcinoma in the lung, followed by squamous metaplasia, dysplasia, and carcinoma in situ (1-3). These extensive and multifocal changes occur throughout the respiratory tree when the lungs are exposed chronically to common carcinogens, a phenomenon referred to as field cancerization (4).

In the completely developed lung, growth factors and their signaling receptors are balanced to support cellular activities at equilibrium and thereby preserve normal lung structure and function (5). However, this homeostatic control can be compromised during the accumulation of genetic and molecular alterations that lead to lung cancer. Several decades of cancer research have revealed that the ErbB system is one of the most critical growth factor systems involved in the normal and abnormal proliferation of epithelial cells (6, 7). The ErbB family, ErbB1-4, plays an important role in lung cancer development. In addition, several ligands of the ErbB system are known to be aberrantly regulated in cancer cells. Therefore, we hypothesized that bronchial hyperplasia is a consequence of a hyperactivation of the ErbB system in bronchial epithelial cells. To test this hypothesis, we examined a panel of ErbB ligands for their ability to induce bronchial hyperplasia using a 3-dimensional (3-D) organotypic air-liquid interface primary bronchial epithelial cell culture system (8-10). We then determined which downstream signaling pathways and genes were involved in the development of bronchial hyperplasia. Our data showed that EGFR ligands induce bronchial hyperplasia via the MEK/ERK signaling pathway. EGF-induced cyclin D1 overexpression plays a critical role in the

development of bronchial hyperplasia. Inhibitors of EGFR and MEK completely blocked EGF-induced bronchial hyperplasia.

As monotherapy, erlotinib, a small molecular inhibitor targeting the intracellular tyrosine kinase domain of EGFR, significantly prolonged the survival of previously treated patients with advanced non-small cell lung cancer (NSCLC) compared to placebo (13) and was recently approved by the FDA. Erlotinib has anti-proliferative effects arising from G1 arrest and pro-apoptotic effects on cancer cells (14). However, the effect of erlotinib on normal or hyperplastic bronchial epithelial cells is not known. Our study clearly showed that erlotinib blocks bronchial hyperplasia induced by EGF and also can reverse hyperplasia, thereby restoring normal bronchial epithelial morphology. Our data help to identify some of the mechanisms involved in the onset of changes leading to lung cancer, such as abnormal cell proliferation, and provide potential targets for preventing for preventing the progression to the invasive malignant state.

Materials and Methods

Chemicals

The EGFR ligands epidermal growth factor (EGF), transforming growth factor- α (TGF- α), amphiregulin (AR), and heregulin (HR), cAMP response element-binding (CREB) small interfering RNA (siRNA), and c-Jun siRNA were purchased from Sigma-Aldrich. Erlotinib (LKT Laboratories, Inc.), U0126, LY294002, and Akt inhibitor VIII (Calbiochem) were each dissolved in dimethylsulfoxide (DMSO).

3-D organotypic air-liquid interface cell culture and treatment

Normal human tracheobronchial epithelial (NHTBE) cells (Lonza, MA) were cultured by the 3-D organotypic air-liquid interface method described previously (9, 11). The medium in the

Prevention of Bronchial Hyperplasia

bottom chamber was changed every 24 h during experiment period. Seven-day old confluent NHTBE cells grown on a porous membrane of Transwell plate were treated with various ligands for ErbB receptors EGF (10 ng/ml), TGF- α (10 ng/ml), AR (50 ng/ml) or HR (100 ng/ml) for 4 days. The ligands were included only in basal media and apical side of the cultures was exposed to air by removing media overlaying the cells in the upper side of the well. For dose dependent experiments, we used 0.5, 2.0, 5.0, 10 and 25 μ g/ml of EGF. For time dependent experiment, we cultured the cells with 5 ng/ml of EGF for 1 to 4 days.

Western blot analysis

Total protein extracts were prepared using cold radioimmunoprecipitation assay lysis buffer (50 mM HEPES, pH 7.4, 1% NP-40, 150 mM NaCl, 1 mM EDTA, phosphates inhibitors and protease inhibitors). Total 15 μ g of protein were resolved by 10% SDS-PAGE gel. Membranes were incubated with rabbit polyclonal antibodies against CREB, phospho-CREB-133 (Upstate Biotechnology), ERK, phospho-ERK-202/204, cyclin A1, cyclin B1, cyclin D1, cyclin E2, Akt, phospho-Akt-473, c-Jun, phospho-c-Jun-73, p-EGFR-1068, and EGFR (Cell Signaling Technology) overnight. β -actin (clone AC-15; Sigma-Aldrich) was used as a loading control. The proteins reactive with primary antibodies were visualized with horseradish peroxidase (HRP)-conjugated secondary antibody and enhanced chemiluminescence reagents (GE Health Care).

Small interfering RNA

Human c-Jun (Accession no. NM_002228) siRNAs were used to knock-down expression of c-Jun, according to the manufacturer's instructions. A mixture of several siRNAs ensured that the targeted gene product was effectively deleted. Cells at 60% to 70% confluency were

transfected for 48 h with a final concentration of 100 nM c-Jun siRNA or nonspecific control pooled siRNAs using the Dharmafect 1 transfection reagent (Dharmacon) according to the manufacturer's instructions. The cells were treated with EGF (5 ng/ml) for 24 h, when target protein levels had been reduced more than 70%, as assessed by Western blot analysis.

Immunohistochemistry (IHC)

The NHTBE cells were fixed in neutral-buffered formalin and embedded in paraffin. Sections (5 μ m each) were prepared using a microtome, mounted on slides, deparaffinized in xylene, rehydrated in graded alcohols, and washed in distilled water. Endogenous peroxidases were quenched by incubation in 3% H₂O₂. Antigens were retrieved by microwaving the sections in 10 mM citric acid (pH 6.0) for 5 min. The slides were washed three times with phosphate-buffered saline (PBS) and blocked for 30 min with 10% normal goat serum in 1% bovine serum albumin (BSA)/PBS. IHC staining was visualized using the Histostain-Bulk-SP kit and the AEC red substrate kit (Zymed Laboratories). IHC staining without a primary antibody was performed as a negative control. Stained slides were visualized with an Axioskop 40 fluorescence microscope (Carl Zeiss, NY), and the images were captured at a magnification of 200 $\times$ and stored using Axiovision LE software v4.5 (Carl Zeiss, NY) according to the manufacturer's instructions.

Luciferase reporter assay

The NHTBE cells (5×10^4) were cultured in 12-well tissue culture plates (Corning, MA) overnight, and co-transfected with cyclin D1 promoter-luciferase constructs (wild type, AP-1 site mutant, or CRE sites mutant; kindly provided by Dr. Richard Pestell at the Thomas Jefferson University) and *Renilla* luciferase control vector using Lipofectamin 2000 (Invitrogen). After 24 h, the culture medium was changed to 0.1% BSA in BEBM, and cells were treated with or

without EGF (5 ng/ml) for 24 h. Luciferase activity was detected using the Dual-Luciferase Reporter assay (Promega) and measured using a Lumat LB 9507 tube luminometer (Berthold, TN). All assays were performed in triplicate and repeated at least three times, and figures show representative results.

Evaluation of hyperplasia via cell layer thickness and cell number counting

To determine the effect of ErbB receptors ligands on the histomorphology of NHTBE cell cultures, NHTBE cells were incubated with EGF (10 ng/ml), TGF- α (10 ng/ml), AR (50 ng/ml) or HR (100 ng/ml) for 4 days. After making the paraffin-embedded block, we captured three images from each block within 10 mm area from the center of transwell membrane with an Axioskop 40 microscope (Carl Zeiss, NY) under light microscopy (200 $\times$). For evaluation of hyperplasia, we measured thickness by Axiovision LE software v4.5 (Carl Zeiss, NY) according to the manufacturer's instructions. Total cell numbers in the captured area were measured by manual counting under light microscopy (200X). bars, standard error (SE); ** $P < 0.01$ and *** $P < 0.001$.

Statistical analysis

The results were summarized by descriptive statistics (mean, SE, and median) and box-plots were generated for each experiment with the different cell groups compared side-by-side. An analysis of variance (ANOVA) model was used to detect any differences among treatment and control groups.

Results

EGFR ligands induce hyperplasia of bronchial epithelial cells grown in 3-D organotypic culture

Prevention of Bronchial Hyperplasia

To determine whether the ErbB receptor system is involved in the morphological changes associated with bronchial hyperplasia, NHTBE cells were cultured by organotypic air-liquid interface method with various ligands of ErbB receptors, including EGF (10 ng/ml), TGF- α (10 ng/ml), AR (50 ng/ml) or HR (100 ng/ml), for 4 days. We examined the effect of ErbB ligands on the morphological change of NHTBE cells by histochemical evaluation of the cultures. The thickness of the NHTBE cell layer was statistically significantly increased by following treatment with EGFR ligands, EGF, TGF- α , and AR, but not by treatment with HR (Fig. 1A). Quantitative changes are shown in the figure as average thicknesses: control=13 ($\pm$ 2) μ m, EGF-treated cells=42 ($\pm$ 6) μ m, TGF- α -treated cells=31 ($\pm$ 4) μ m, AR-treated cells=28 ($\pm$ 3) μ m and HR-treated cells=15 ($\pm$ 2) μ m. These data clearly demonstrate that high concentrations of EGFR ligands induce hyperplasia of bronchial epithelial cells. The most prominent hyperplastic morphological changes in the histologic pattern of NHTBE cell culture were induced by EGF; therefore, we selected EGF for subsequent study.

To assess the dose-dependent effect of EGF on NHTBE cell hyperplasia, NHTBE cells were incubated with EGF at indicated concentrations for 4 days. Immunohistochemical analysis and cell quantitation clearly indicated that EGF induced hyperplasia in a dose-dependent manner. We detected a significant increase in the NHTBE cell layer thickness after treatment with 5 ng/ml of EGF, and cell quantitation showed a similar pattern (Fig. 1B). To determine the time-dependent effect of EGF, NHTBE cells were incubated with 5 ng/ml of EGF for indicated time periods (1 to 4 days). When treated with EGF, the NHTBE cell layer gradually expanded during for 4 days treatment in a time-dependent manner. Four-day treatment with EGF resulted in approximately 2.5-fold increase in cell number and thickness of cell layers (Fig. 1C).

EGF induces cell proliferation at only the basal layer of NHTBE cell cultures

Multilayered hyperplasia is believed to be a consequence of uncontrolled proliferation of bronchial epithelial cells. To identify the specific population of NHTBE cells following exposure to EGFR ligands, we performed Ki-67 immunostaining in NHTBE cells grown in 3-D organotypic culture system. We used A549 cells as a positive control. As shown in Fig. 2, immunofluorescence analysis of cell proliferation showed that cells in the basal layer of the NHTBE cultures were stained with anti-Ki67, but A459 cells were randomly stained in both basal and parabasal layers (Fig. 2). These results clearly indicate that only NHTBE cells in the basal layer divide and grow in response to EGF, while dividing cancer cells are not limited to the basal layer in response to EGF.

The MEK/ERK pathway is a critical signaling pathway for EGF-induced hyperplasia of NHTBE cells

The MEK/ERK and PI3-K/Akt pathways are well established downstream signal pathways of the EGF-EGFR pathway (12). To determine the relative importance of the two pathways for transmitting signals for EGF-induced hyperplasia in NHTBE cells, NHTBE cells were treated with 5 ng/ml of EGF for 2 hr. EGF induced phosphorylation of Akt, ERK, and CREB in NHTBE cells (Fig. 3A). This result demonstrates that potentially both Akt and ERK pathways can participate in induction of cell proliferation and layer thickening by EGF. To further examine the relative role of the two pathways, we treated NHTBE cells with EGF along with pharmacological inhibitors targeting select molecules in the PI3-K/Akt and ERK signaling pathways: erlotinib (EGFR-TKI), LY294002 (PI3-K inhibitor), Akt inhibitor VIII, and U0126 (MEK inhibitor). Erlotinib completely inhibited the phosphorylation of EGFR, Akt, and ERK (Fig. 3B). MEK inhibitor U0126 inhibited the phosphorylation of ERK and CREB. However,

LY294002 and Akt inhibitor VIII had no effect on the EGF-induced phosphorylation of ERK and CREB (Fig. 3B).

To further determine the morphologic consequence of inhibiting the critical molecules in the signaling pathways, NHTBE cells were cultured in EGF (5 ng/ml) with/without inhibitors for 4 days. NHTBE cells treated with EGF alone were subjected to hyperplastic changes and increased thickness, while those treated with EGF in the presence of erlotinib or U0126 underwent no such changes. However, LY294002- and Akt inhibitor VIII-treated NHTBE cells still showed hyperplasia and increased thickness (Fig. 3C). These results strongly suggest that the MEK/ERK pathway is the main pathway by which EGF induces hyperplasia in NHTBE cells. We next asked whether inhibitors can reverse EGF-induced hyperplasia in NHTBE cells. After they became hyperplastic, cells were treated with erlotinib, LY294002, U0126 or Akt inhibitor VIII for 72 hr. Only erlotinib successfully counteracted EGF-induced hyperplasia of NHTBE cell cultures such that normal morphology was restored (Fig. 3D).

Cyclin D1 is increased during EGF-induced NHTBE cell hyperplasia

Increased cell proliferation is partly responsible for the induction of hyperplasia. Since elevated cyclin levels are known to play an important role in enhancing cell proliferation, we determined the levels of various cyclins in NHTBE cells after EGF treatment. EGF treatment robustly increased cyclin D1 expression and slightly increased cyclin E2 (Fig. 4A). However, cyclins A1 and B1 were not significantly increased by EGF. The EGFR and MEK inhibitors erlotinib and U0126, respectively, markedly blocked EGF-induced expression of cyclin D1, but LY294002 and Akt inhibitor VIII did not (Fig. 4B).

To determine which transcription factors are involved in EGF-induced *cyclin D1* gene expression, we performed a *cyclin D1* promoter-luciferase activity assay (Fig. 4C). We

transfected NHTBE cells with various *cyclin D1* promoter-luciferase reporters (wild type, AP-1 site mutant, and CRE sites mutant in the cyclin D1 promoter region) and treated the transfected cells with or without EGF. EGF increased luciferase activity by over 20 times when wild-type or mutated CRE *cyclin D1* promoter-reporters were introduced in the cells (Fig. 4C). However, when the recognition sequence of AP-1 in the promoter was mutated or removed, the response to EGF in the promoter was dramatically lower than the response in the wild-type promoter. This result demonstrates the importance of the AP-1 transcription factor in EGF-induced cyclin D1 overexpression.

Next, we investigated activation status of AP-1 component, c-Jun and c-Fos. We found that EGF induced the expression and phosphorylation of c-Jun (Fig. 4D) but not c-Fos (data not shown). In addition, erlotinib and U0126 markedly blocked EGF-induced c-Jun phosphorylation (Fig. 4E). Knockdown of c-Jun with c-Jun siRNA prevented EGF-induced cyclin D1 expression, suggesting that EGF-induced cyclin D1 expression is mediated by c-Jun (Fig. 4F). Taken these data together, we concluded that EGF induces cyclin D1 overexpression and this overexpression is mediated by AP-1 (c-Jun) transcription factor. In addition, EGF-induced cyclin D1 overexpression is blocked by inhibitors of EGFR and MEK.

Discussion

We demonstrated that bronchial hyperplasia can be modeled and manipulated *in vitro* using primary NHTBE cells maintained in a 3-D organotypic air-liquid interface culture. Our results clearly demonstrate that EGFR ligands EGF, TGF- α , and AR induce hyperplasia in NHTBE cells. This histomorphological change is regulated by the MEK/ERK signaling pathway, but not the PI3-K/Akt signaling pathway. The MEK/ERK signaling pathway induces cyclin D1

expression level through activation of AP-1 transcription factor. Inhibitors of EGFR and MEK, erlotinib and U0126, completely blocked EGF-induced hyperplasia.

In view of multistep lung carcinogenesis and field cancerization, our results suggest the possible role of erlotinib as a chemopreventive agent, as such agents inhibit, delay, or reverse carcinogenesis. First, it is conceivable that the use of erlotinib will be beneficial for high-risk patients, such as those with a strong smoking history. Erlotinib is currently being studied in the adjuvant setting after surgery and chemotherapy in NSCLC. Lung cancer develops in a field with extensive and multifocal hot spots throughout the respiratory trees which are consistently exposed to common carcinogens. Even after the resection of primary tumors, hot spots in the remaining bronchial trees have the potential to develop into lung cancer. After resection of an NSCLC tumor, the risk of developing a second primary lung cancer is approximately 1% to 2% per patient per year, with a cumulative risk up to 20% at 6 to 8 years following resection (15). If erlotinib could be used in the treatment of early-stage NSCLC after resection, it would target a dual cellular population at once: micrometastatic NSCLC cells and evolving bronchial epithelial cells with the potential to develop into lung cancer. Erlotinib's inhibition and reversal of the first step in lung carcinogenesis in NHTBE cells warrant further investigation. We are currently conducting a clinical study and have enrolled 50 patients with early-stage lung cancer who have undergone neoadjuvant chemotherapy with cisplatin and docetaxel followed by surgical resection and 1 year of adjuvant erlotinib. Patients will undergo bronchoscopy at 6 months and 1 year to assess possible changes in the bronchial epithelium after erlotinib treatment (16).

Our study showed that EGF robustly increases cyclin D1 in primary NHTBE cells grown in an organotypic culture. The malignant transformation of bronchial epithelial cells is driven by the dysregulation of oncogenes, growth factors, or tumor suppressor genes. Cyclin D1 is strongly

implicated as an oncogene in lung cancer and several other human cancers, including B-cell lymphomas, squamous cell carcinomas of the head and neck, esophageal cancer, and breast cancer (17). Cyclin D1 is part of the cyclin-dependent kinase (CDK)--cyclin complex that increases retinoblastoma (Rb) protein phosphorylation at the G1-S transition and may also play an important role in transcriptional regulation. The deregulation of cyclin D either by amplification or by transcriptional upregulation has been demonstrated in many tumor types (18-22). The p53-Rb pathway that mediates G1 arrest is the most commonly affected pathway in lung cancer. Defects in G1 regulatory proteins, especially the deregulation of p53-p21WAF1, p16-Rb-cyclin D1 and cyclin E-p16 pathways, seem to be essential events in lung cancer development (23, 24). Immunohistochemical analysis has demonstrated overexpression of cyclins D1 and E in bronchial preneoplasia that precedes the development of squamous cell carcinoma (25). These data imply that increased cyclin levels could play a critical role in the progression of preneoplastic bronchial lesions. This conclusion has been confirmed in bronchial epithelial cellular models (14) and carcinogen-induced lung tumors in animal models (26, 27).

Cyclin D1 overexpression may portend a worse prognosis in patients with resected lung cancer (28), though results have not been consistent across studies (29). Cyclin D1 appears to be regulated by EGFR in EGFR mutant cell lines that are resistant to gefitinib, and these cell lines are sensitive to flavopiridol, a CDK inhibitor (30). Repression of cyclin D1 is an indirect marker of response to erlotinib treatment in aerodigestive tract cancers (31). Chemoprevention trials have found that rexinoid, selective retinoid X receptor agonist, suppresses cyclin D1 expression in NSCLC (32), and low cyclin D1 expression predicts longer cancer-free survival in laryngeal premalignancy patients (33). Cyclin D1 levels thus have been studied as markers for abnormal cell growth in chemoprevention trials (34). Regulation of Cyclin D1 gene expression have been

Prevention of Bronchial Hyperplasia

reported to include the ras/raf/MAPK cascade in fibroblast cells (35, 36), p60Src pathways through CREB/activating transcription factor 2 activation in breast cancer cells (37) and PI3-K/Akt/NF- κ B pathway involved pro-oncogenic effects in human bronchial epithelial cells (38). Our data showed that EGF activates both the PI3-K/Akt and MEK/ERK pathways in a time-dependent manner. However, only MEK/ERK pathway is involved in EGF-induced cyclin D1 expression, suggesting MEK/ERK pathway is involved in the early stage of lung carcinogenesis.

In summary, we demonstrate that bronchial hyperplasia can be modeled *in vitro* using a 3-D organotypic culture method and can be prevented by blocking the EGFR/MEK signaling pathway. We further show that bronchial hyperplasia is dependent on cyclin D1, which is in turn regulated by AP-1 activation through the MEK/ERK pathway rather than the PI3-K/Akt pathway. Our model system and results will help to elucidate the molecular mechanisms of lung carcinogenesis at its early stages, and may support the prophylactic usage of EGFR targeting agents for patients with high risk of tumor development.

References

1. Saccomanno G, Archer VE, Auerbach O, Saunders RP, Brennan LM. Development of carcinoma of the lung as reflected in exfoliated cells. *Cancer* 1974; 33:256-70.
2. Auerbach O, Hammond EC, Garfinkel L. Changes in bronchial epithelium in relation to cigarette smoking, 1955-1960 vs. 1970-1977. *N Engl J Med* 1979; 300:381-5.
3. Saccomanno G, Archer VE, Saunders RP, Auerbach O, Klein MG. Early indices of cancer risk among uranium miners with reference to modifying factors. *Ann N Y Acad Sci* 1976; 271:377-83.
4. Strong MS, Incze J, Vaughan CW. Field cancerization in the aerodigestive tract--its etiology, manifestation, and significance. *J Otolaryngol* 1984; 13:1-6.
5. Warburton D, Schwarz M, Tefft D, Flores-Delgado G, Anderson KD, Cardoso WV. The molecular basis of lung morphogenesis. *Mech Dev* 2000; 92:55-81.
6. Scagliotti GV, Selvaggi G, Novello S, Hirsch FR. The biology of epidermal growth factor receptor in lung cancer. *Clin Cancer Res* 2004; 10:4227s-32s.
7. Vlahovic G, Crawford J. Activation of tyrosine kinases in cancer. *The oncologist* 2003; 8:531-8.
8. Koo JS, Jetten AM, Belloni P, Yoon JH, Kim YD, Nettesheim P. Role of retinoid receptors in the regulation of mucin gene expression by retinoic acid in human tracheobronchial epithelial cells. *Biochem J* 1999; 338:351-7.
9. Aggarwal S, Kim SW, Cheon K, Tabassam FH, Yoon JH, Koo JS. Nonclassical action of retinoic acid on the activation of the cAMP response element-binding protein in normal human bronchial epithelial cells. *Mol Biol Cell* 2006; 17:566-75.

10. Kim SW, Cheon K, Kim CH, et al. Proteomics-based identification of proteins secreted in apical surface fluid of squamous metaplastic human tracheobronchial epithelial cells cultured by three-dimensional organotypic air-liquid interface method. *Cancer Res* 2007; 67:6565-73.
11. Koo JS, Yoon JH, Gray T, Norford D, Jetten AM, Nettesheim P. Restoration of the mucous phenotype by retinoic acid in retinoid-deficient human bronchial cell cultures: changes in mucin gene expression. *Am J Respir Cell Mol Biol* 1999; 20:43-52.
12. Loupakis F, Vasile E, Santini D, Masi G, Falcone A, Graziano F. EGF-receptor targeting with monoclonal antibodies in colorectal carcinomas: rationale for a pharmacogenomic approach. *Pharmacogenomics* 2008; 9:55-69.
13. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, et al. Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* 2005; 353:123-32.
14. Moyer JD, Barbacci EG, Iwata KK, et al. Induction of apoptosis and cell cycle arrest by CP-358,774, an inhibitor of epidermal growth factor receptor tyrosine kinase. *Cancer Res* 1997; 57:4838-48.
15. Johnson BE. Second lung cancers in patients after treatment for an initial lung cancer. *J Natl Cancer Inst* 1998; 90:1335-45.
16. Gold KA, Lee JJ, Rice D, et al. Phase II pilot study of neoadjuvant docetaxel and cisplatin followed by adjuvant erlotinib in patients with stage I-III non-small cell lung cancer (NSCLC). *Journal of Clinical Oncology*, 2009 ASCO Annual Meeting Proceedings (Post-Meeting Edition) 2009; 27.
17. Motokura T, Arnold A. Cyclin D and oncogenesis. *Curr Opin Genet Dev* 1993; 3:5-10.

18. Sheyn I, Noffsinger AE, Heffelfinger S, Davis B, Miller MA, Fenoglio-Preiser CM. Amplification and expression of the cyclin D1 gene in anal and esophageal squamous cell carcinomas. *Hum Pathol* 1997; 28:270-6.
19. Bellacosa A, Almadori G, Cavallo S, et al. Cyclin D1 gene amplification in human laryngeal squamous cell carcinomas: prognostic significance and clinical implications. *Clin Cancer Res* 1996; 2:175-80.
20. Yatabe Y, Nakamura S, Seto M, et al. Clinicopathologic study of PRAD1/cyclin D1 overexpressing lymphoma with special reference to mantle cell lymphoma. A distinct molecular pathologic entity. *Am J Surg Pathol* 1996; 20:1110-22.
21. Gillett C, Fantl V, Smith R, et al. Amplification and overexpression of cyclin D1 in breast cancer detected by immunohistochemical staining. *Cancer Res* 1994; 54:1812-7.
22. Bartkova J, Lukas J, Strauss M, Bartek J. Cyclin D1 oncoprotein aberrantly accumulates in malignancies of diverse histogenesis. *Oncogene* 1995; 10:775-8.
23. Reissmann PT, Koga H, Takahashi R, et al. Inactivation of the retinoblastoma susceptibility gene in non-small-cell lung cancer. The Lung Cancer Study Group. *Oncogene* 1993; 8:1913-9.
24. Kratzke RA, Greatens TM, Rubins JB, et al. Rb and p16INK4a expression in resected non-small cell lung tumors. *Cancer Res* 1996; 56:3415-20.
25. Lonardo F, Rusch V, Langenfeld J, Dmitrovsky E, Klimstra DS. Overexpression of cyclins D1 and E is frequent in bronchial preneoplasia and precedes squamous cell carcinoma development. *Cancer research* 1999; 59:2470-6.
26. Witschi H, Espiritu I, Suffia M, Pinkerton KE. Expression of cyclin D1/2 in the lungs of strain A/J mice fed chemopreventive agents. *Carcinogenesis* 2002; 23:289-94.

27. Boyle JO, Langenfeld J, Lonardo F, et al. Cyclin D1 proteolysis: a retinoid chemoprevention signal in normal, immortalized, and transformed human bronchial epithelial cells. *Journal of the National Cancer Institute* 1999; 91:373-9.
28. Ratschiller D, Heighway J, Gugger M, et al. Cyclin D1 overexpression in bronchial epithelia of patients with lung cancer is associated with smoking and predicts survival. *J Clin Oncol* 2003; 21:2085-93.
29. Filipits M, Pirker R, Dunant A, et al. Cell cycle regulators and outcome of adjuvant cisplatin-based chemotherapy in completely resected non-small-cell lung cancer: the International Adjuvant Lung Cancer Trial Biologic Program. *J Clin Oncol* 2007; 25:2735-40.
30. Kobayashi S, Shimamura T, Monti S, et al. Transcriptional profiling identifies cyclin D1 as a critical downstream effector of mutant epidermal growth factor receptor signaling. *Cancer Res* 2006; 66:11389-98.
31. Petty WJ, Dragnev KH, Memoli VA, et al. Epidermal growth factor receptor tyrosine kinase inhibition represses cyclin D1 in aerodigestive tract cancers. *Clin Cancer Res* 2004; 10:7547-54.
32. Dragnev KH, Petty WJ, Shah SJ, et al. A proof-of-principle clinical trial of bexarotene in patients with non-small cell lung cancer. *Clin Cancer Res* 2007; 13:1794-800.
33. Papadimitrakopoulou V, Izzo JG, Liu DD, et al. Cyclin D1 and cancer development in laryngeal premalignancy patients. *Cancer Prev Res (Phila Pa)* 2009; 2:14-21.
34. Freemantle SJ, Guo Y, Dmitrovsky E. Retinoid chemoprevention trials: cyclin D1 in the crosshairs. *Cancer Prev Res (Phila Pa)* 2009; 2:3-6.

35. Takuwa N, Takuwa Y. Ras activity late in G1 phase required for p27kip1 downregulation, passage through the restriction point, and entry into S phase in growth factor-stimulated NIH 3T3 fibroblasts. *Molecular and cellular biology* 1997; 17:5348-58.
36. Lavoie JN, L'Allemain G, Brunet A, Muller R, Pouyssegur J. Cyclin D1 expression is regulated positively by the p42/p44MAPK and negatively by the p38/HOGMAPK pathway. *J Biol Chem* 1996; 271:20608-16.
37. Lee RJ, Albanese C, Stenger RJ, et al. pp60(v-src) induction of cyclin D1 requires collaborative interactions between the extracellular signal-regulated kinase, p38, and Jun kinase pathways. A role for cAMP response element-binding protein and activating transcription factor-2 in pp60(v-src) signaling in breast cancer cells. *J Biol Chem* 1999; 274:7341-50.
38. Han SW, Roman J. Fibronectin induces cell proliferation and inhibits apoptosis in human bronchial epithelial cells: pro-oncogenic effects mediated by PI3-kinase and NF-kappa B. *Oncogene* 2006; 25:4341-9.

Acknowledgements

We thank Dr. Richard Pestell (Thomas Jefferson University, Philadelphia, PA) for the gift of cyclin D1 luciferase vectors.

Figure Legends

Fig. 1. EGFR ligands induces hyperplasia in NHTBE cells. NHTBE cells were cultured on the Transwell® plate until make monolayer, and then changed culture condition to air-liquid interface as described in the Materials and Methods section. Starting at day 7, cells were treated with various ErbB-ligands (EGF 10 ng/ml, TGF- α 10 ng/ml, AR 50 ng/ml and HR 100 ng/ml) for 4 days (A). EGF induces hyperplasia in a dose- (B) and time-dependent (C) manner. Seven-day-old NHTBE cell cultures were treated with indicated concentrations (B) of EGF for the indicated time periods (C). Cells were then fixed with phosphate-buffered formalin and embedded in paraffin. The result was visualized using hematoxylin/eosin staining method. Cell number was quantified by counting as described in the Materials and Methods section. Data shown are representative of three experiments with similar results.

Fig. 2. Proliferating cells are differentially localized in NHTBE or A549 cancer cells. Both cells were grown by ALI method as described in the Materials and Methods section. The cultures were starved without supplements in their respective culture media for 24 h, then treated with 5 ng/ml of EGF (0.5 ng/ml as control). The cultures were fixed and paraffin blocks were prepared. Slide sections were stained with anti-Ki-67. The result was visualized using immunofluorescence and immunohistochemistry. Data shown are representative of three experiments with similar results.

Fig. 3. MEK/ERK pathway is involved in EGF-induced hyperplasia in NHTBE cells. (A) EGF induced phosphorylation of Akt, ERK, and CREB. Seven-day-old NHTBE cell cultures

were maintained in BEBM without supplements for overnight, and then treated with EGF (5 ng/ml) for indicated period of time. (B) Effect of pharmacological inhibitors on EGF induced phosphorylation of Akt, ERK, and CREB. (C) Histological evaluation of the preventive effect with pharmacological inhibitors. (D) Treatment effects of erlotinib on EGF-induced hyperplasia. Seven-day-old NHTBE cell cultures were treated with EGF (5 ng/ml) in combination with indicated pharmacological inhibitors for 4 days. The result was evaluated by histochemistry using hematoxylin/eosin staining method. Cell number was counted as described in the Materials and Methods section.

Fig. 4. AP-1 is a critical transcription factor in EGF-induced cyclin D1 overexpression in NHTBE cells. (A, B, D and E) NHTBE cells were cultured on the Transwell plate until make monolayer, and then changed culture condition to air-liquid interface. Before stimulation with EGF or treated with pharmacological inhibitors, cells were incubated with 0.1% BSA/BEBM medium for 24 h before treatment. (A) EGF upregulates cyclin D1 expression in time-dependent manner. (B) Erlotinib and U0126 inhibit EGF-induced cyclin D1 expression. (C) AP-1 binding site is important for cyclin D1 promoter activity. Cells were transfected with indicated plasmids using 12-well tissue culture plate. After 48hr incubation, cells were maintained in 0.1% BSA/BEBM without supplements for overnight, and then treated with EGF (5 ng/ml) for 24hr. (D) EGF induces phosphorylation of c-Jun in a time-dependent manner. (E) Erlotinib and U0126 inhibit EGF-induced c-Jun phosphorylation. (F) c-Jun siRNA abrogated EGF-induced cyclin D1 expression. Cells were transfected with siRNA using 12-well tissue culture plate. After 48hr incubation, cells were maintained in 0.1% BSA/BEBM without supplements for overnight, and

then treated with EGF (5 ng/ml) for 24hr. Data shown are representative of three experiments with similar results.

Fig. 5. EGF induced the increased expression of cyclin D1, which plays a critical role in bronchial hyperplasia, and the overexpression was mediated by activating the mitogen-activated protein kinase (MAPK) pathway.

Fig. 1.

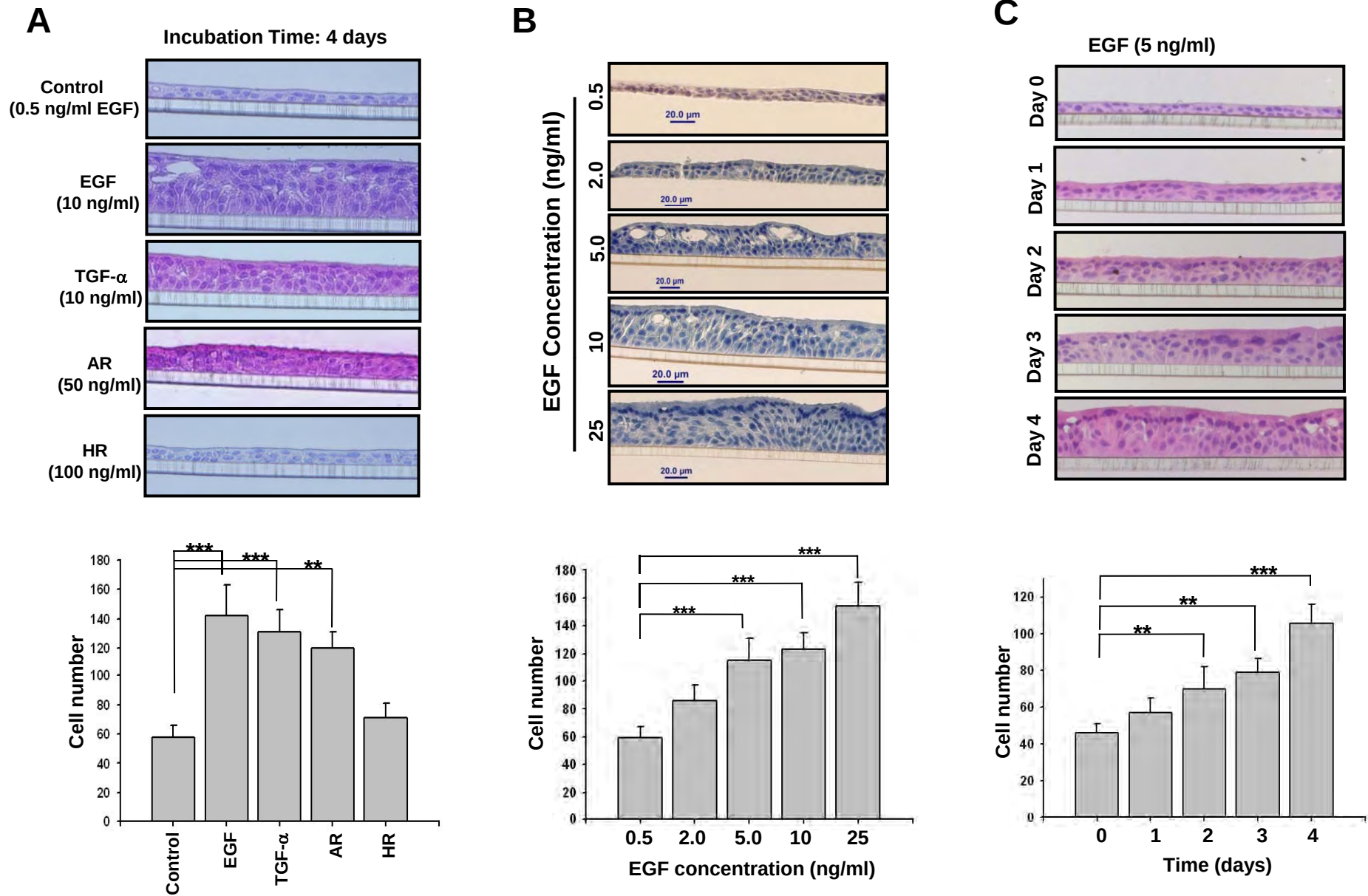
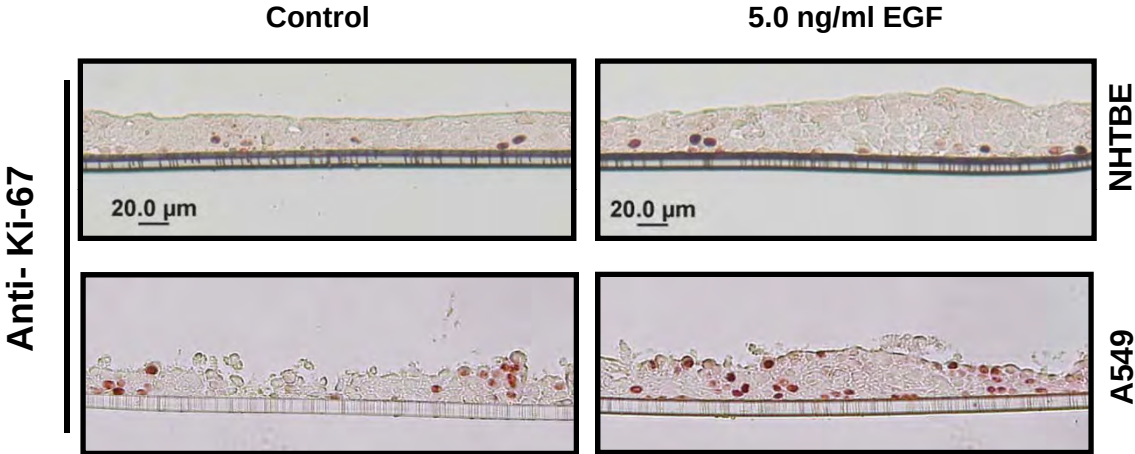
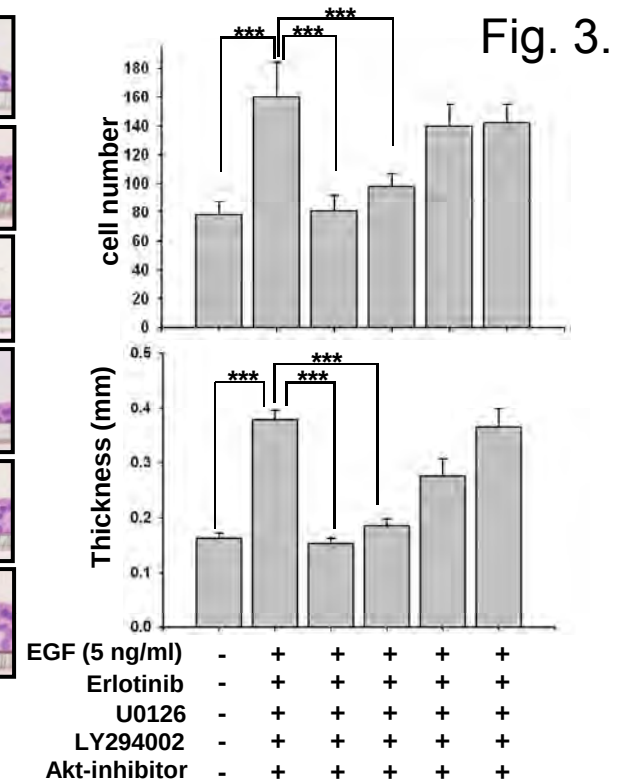
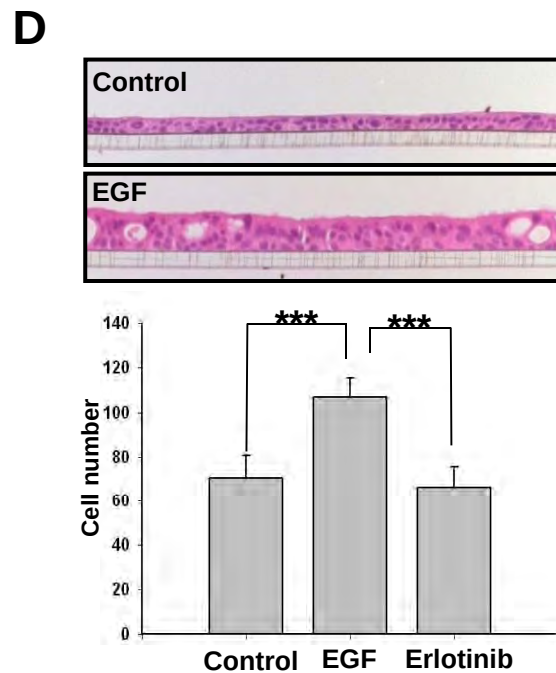
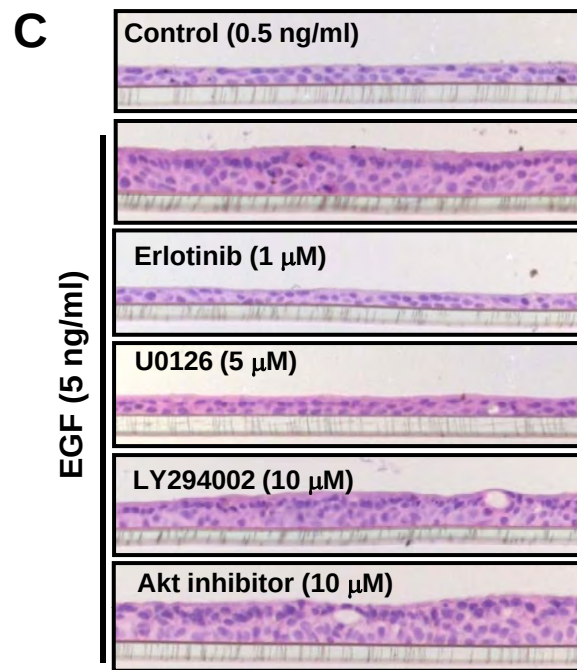
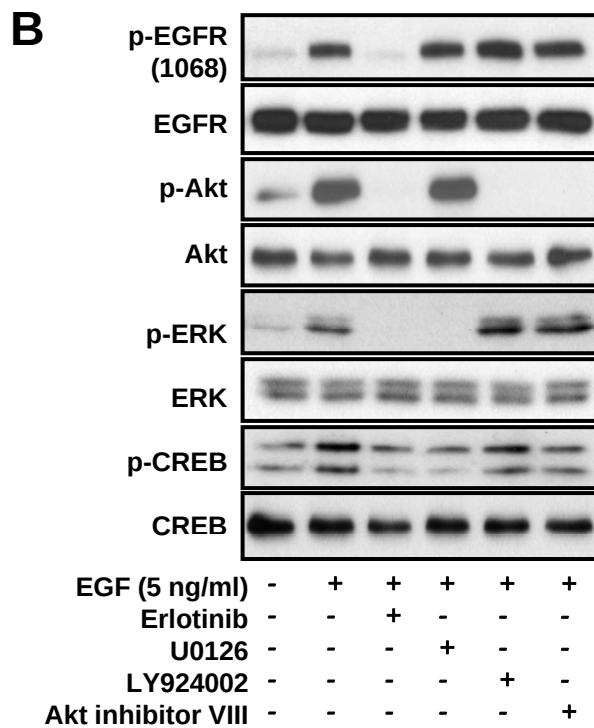
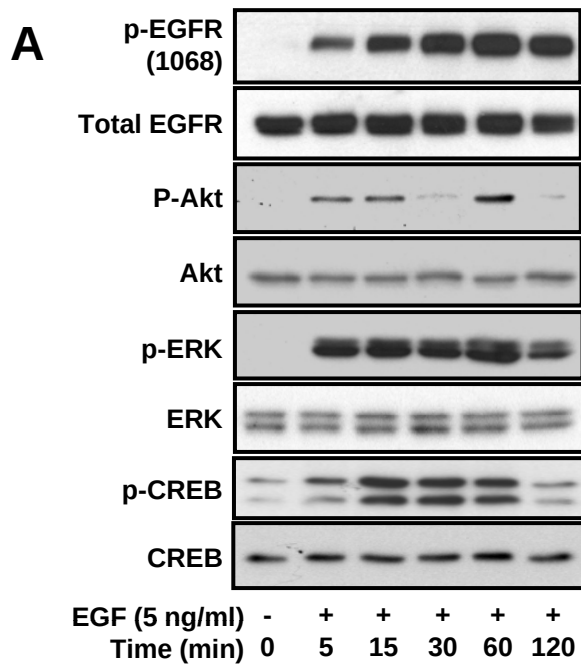


Fig. 2.





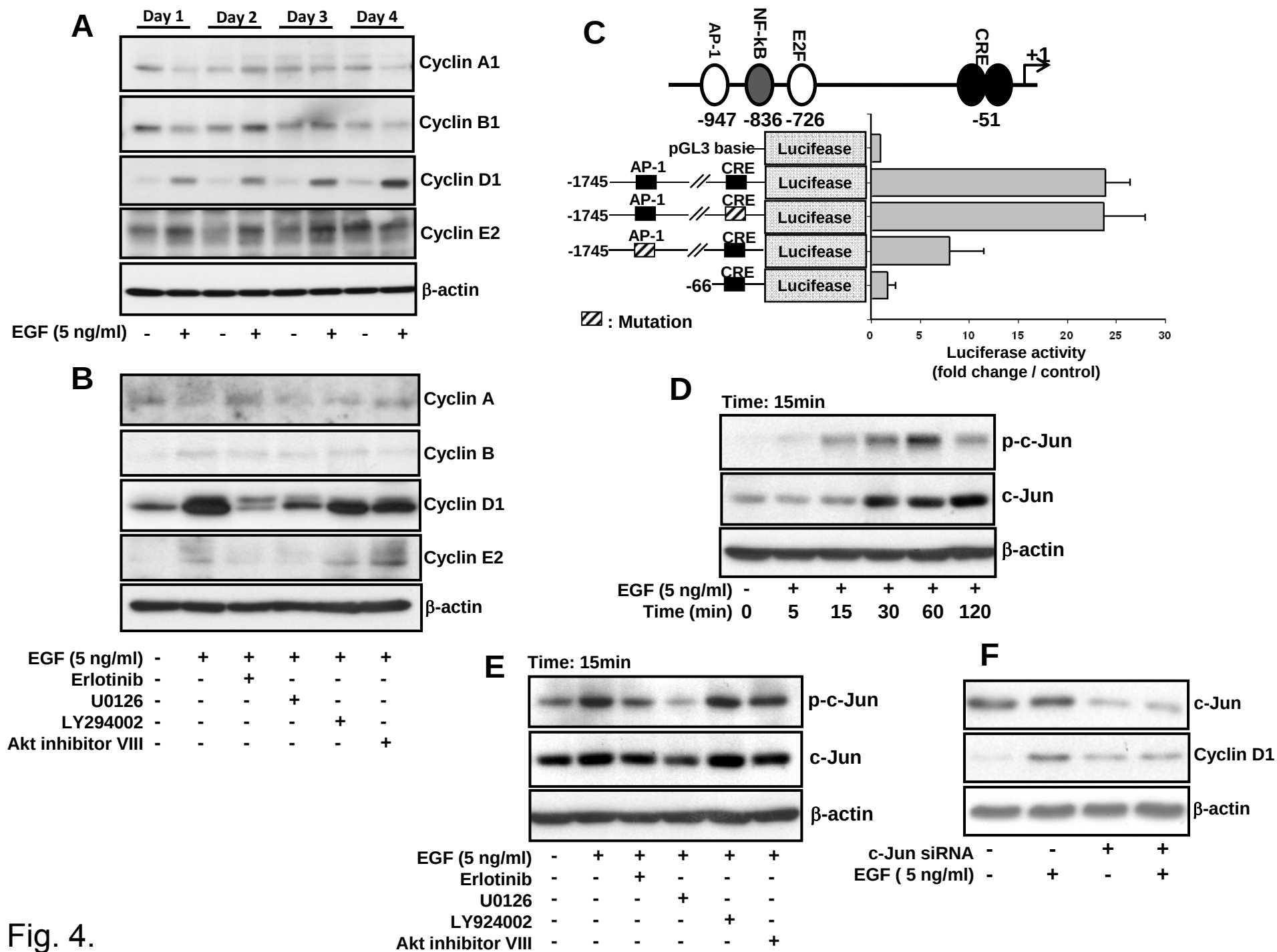
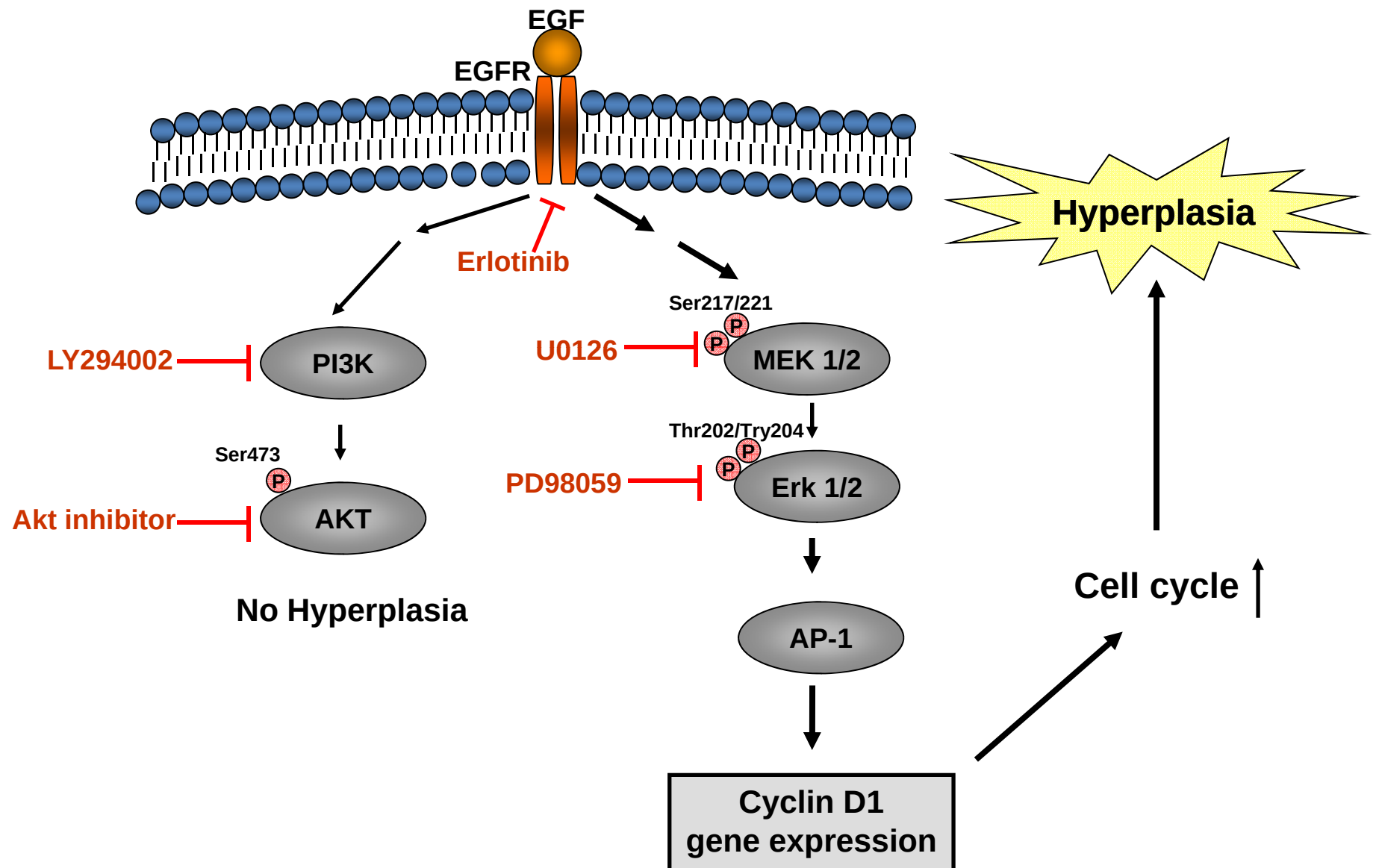


Fig. 4.

Fig. 5.



[Print this Page](#)

Presentation Abstract

Abstract
Number: 787

Presentation
Title: Differences in protein expression patterns in lung adenocarcinomas arising in never versus ever smokers

Presentation
Time: Sunday, Apr 18, 2010, 2:00 PM - 5:00 PM

Location: Exhibit Hall A-C, Poster Section 30

Poster
Section: 30

Poster Board
Number: 6

Author
Block: Carmen Behrens¹, Heather Lin¹, Maria Nunez¹, Ping Yuan¹, Luisa Solis¹, Maria G. Raso¹, Ludmila Prudkin¹, Menghong Sun¹, Xiaoling Li¹, Ximing Tang¹, Jack A. Roth¹, John D. Minna², David J. Stewart¹, Waun K. Hong¹, J. Jack Lee¹, Ignacio I. Wistuba¹. ¹UT M.D. Anderson Cancer Ctr., Houston, TX; ²UT Southwestern Medical Center, Dallas, TX

Abstract
Body: **Background.** Approximately 25% of lung cancer cases worldwide, mostly adenocarcinomas, are not attributable to tobacco use. Despite that some striking differences have been identified in the epidemiological, clinical and molecular characteristics of lung cancer arising in never smokers versus smokers, our current knowledge of lung adenocarcinoma in never smokers is still limited.

Methods. We examined the immunohistochemical (IHC) expression of 101 proteins in surgically resected lung adenocarcinoma tissue microarray specimens obtained from 52 never smokers and compared the findings with 152 tumors obtained from ever smokers. The markers examined included a wide variety of tumor-related proteins, representing all hallmarks of cancer (Hanahan and Weinberg, Cell 2000). The IHC expression was assessed at cytoplasm [c], membrane [m], and nucleus [n] of malignant cells, and in stromal cells. Univariate and multivariate (adjusting by patients' sex, and tumor stage and EGFR mutation status) statistical analyses were performed to assess the statistical differences in the expression of markers according to smoking status. The expression of the markers was correlated with patients' clinical characteristics and tumors' pathological features and *EGFR* mutational status.

Results. In the multivariate analysis, tumors from smokers showed a relatively high number of markers (n=32) with significant higher expression compared with tumors from never smokers. Interestingly, in the univariate analysis, nine markers showed significantly higher expression in tumors from never smokers compared with ever smokers, including FGFR-1 [n], FGFR-2 [n], ER-alpha [n], CD44 [c], FOLR1 [m], IGFBP3 [n], IL-1alpha [c], NF-kB [n], survivin [n] and RGS17 [n]. In the multivariate analysis, six markers showed significantly higher expression in tumors from never smokers, including FGFR-2 [n] ($P=0.018$), CD44 [c] ($P=0.001$), c-Met [c] ($P=0.045$) and [m] ($P=0.017$), E-Cadherin [m] ($P=0.003$), IGFBP3 [n] ($P=0.0009$) and p-HER3 [m] ($P=0.035$). Twenty-nine markers showed significant association with *EGFR* mutations in tumors in the multivariate analysis adjusting by patients' sex and smoking status, and tumor stage. Additionally, 47 markers showed significant differences in the level of expression comparing patients' smoking status, including current, former and never smokers.

Conclusion. Our findings indicate that there are multiple molecular differences between lung adenocarcinomas arising in never and ever smokers, suggesting that they are different entities. These findings have implications for the selection of molecular targets for developing novel therapy in patients with lung adenocarcinoma based on their smoking history (Supported by grant VITAL W81XWH-04-1-0142 and UT-Lung SPORE P50CA070907).

[American Association for Cancer Research](#)
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

EZH2 expression is an early event in the pathogenesis of non-small cell lung cancer (NSCLC) and correlates with tumor progression.

Author Block: Carmen Behrens¹, Ping Yuan², Luisa Solis¹, Pierre Saintigny¹, Humam Kadara¹, Junya Fujimoto¹, Cesar Moran¹, Stephen G. Swisher¹, John V. Heymach¹, Ignacio I. Wistuba¹.

¹UT M.D. Anderson Cancer Ctr., Houston, TX; ²Md Anderson, Houston, TX

Abstract:

Background. The molecular events associated with NSCLC pathogenesis and tumor progression need to be better elucidated. The enhancer of zeste homolog 2 (EZH2) is a DNA methyl transferase involved in malignant transformation and tumor progression of several human carcinomas, including lung. We investigated EZH2 expression by immunohistochemistry (IHC) in the early pathogenesis of NSCLC and progression in a large series of clinically well-annotated tissue specimens.

Methods. We examined by IHC nuclear EZH2 expression using formalin-fixed and paraffin-embedded tissue specimens obtained from surgically resected tumors in tissue microarrays (TMAs) including: a) stage I-III NSCLC tumors (SCCs, n=272; adenocarcinomas, n=456); b) paired primary tumors and brain metastases (n=70); and, c) bronchial preneoplastic squamous lesions (n=51) and mildly abnormal/normal bronchial epithelia (n=203). In stage I-III tumors, we correlated EZH2 expression with clinico-pathological features, including patients' recurrence-free survival (RFS), and overall survival (OS), in a subset of these tumors, with IHC expression of 80 proteins and *EGFR* and *KRAS* mutation status.

Results. EZH2 expression was significantly ($P<0.0001$) higher in SCC (mean score=128.6) compared to adenocarcinoma (mean score=56.8). In adenocarcinoma, higher EZH2 expression significantly correlated with ever-smoking status ($P<0.0001$) and less differentiated histology features (solid histology pattern; $P<0.0001$). In multivariate analysis, for adenocarcinoma patients, higher EZH2 expression, as a continuous variable, associated with significantly worse RFS (HR 1.006 95%CI 1.0-1.011; $P=0.03$) and OS (HR 1.004 95%CI 1.0-1.009; $P=0.03$). In publicly available array datasets of lung adenocarcinoma patients, high *EZH2* mRNA correlated with worse RFS and OS. NSCLC brain metastases showed significantly ($P=0.0004$) higher EZH2 expression than corresponding primary tumors. In bronchial epithelia, normal and hyperplastic cells demonstrated low levels of EZH2 expression; significantly higher expression was associated with increasing severity of squamous dysplastic changes ($P<0.0001$). In NSCLC tumors, EZH2 expression positively correlated ($P<0.0001$) with IHC expression of Ki67, FEN1, and UBE2C. In lung adenocarcinomas, *EGFR*-mutant tumors showed significantly lower EZH2 expression than wild-type tumors.

Conclusions. Our findings indicate that EZH2 is frequently expressed in NSCLC, particularly in poorly differentiated adenocarcinomas. In adenocarcinomas, EZH2 associates with worse patient outcomes. These data suggest that EZH2 expression represents an early event in NSCLC pathogenesis and associates with tumor progression and metastasis, representing a novel target for chemoprevention and therapeutic strategies. Supported by DoD grants W81XWH-04-1-0142 and DoD W81XWH-07-1-0306.

Implication of GPRC5A loss in lung carcinogenesis in patients with and without chronic obstructive pulmonary disease

Junya Fujimoto¹, Humam Kadara¹, Carmen Behrens¹, Diane Liu², J Jack Lee², Luisa M Solis³, Edward S Kim¹, Amir Shalafkhane⁴, Ignacio I Wistuba^{3,1} and Reuben Lotan¹

UT MD Anderson Cancer Center, Houston, TX, ¹ THNMO, ² Biostatistics and Applied Mathematics, ³ Pathology, ⁴ Baylor College of Medicine, Houston, TX Department of Medicine

Abstract

Lung cancer is often associated with inflammation induced by cigarette smoke. In addition, chronic obstructive pulmonary disease (COPD), typically associated with inflammation, is a leading cause of morbidity and mortality in the United States and presents an increased risk of lung cancer development compared to patients without COPD. Mice with knockout of both alleles of the G-protein coupled receptor, family C, group 5, member A (Gprc5a) gene develop spontaneous lung adenomas and adenocarcinomas at a much higher incidence than their wild-type littermates, indicating that this gene is a novel lung-specific tumor suppressor gene. Interestingly, the majority of tumors in the Gprc5a knockout mice are associated with inflammatory cell infiltration, possibly due to increased NF- κ B activation in mouse lung epithelial cells and tissues. Furthermore, the human GPRC5A can suppress NF- κ B activation in human lung adenocarcinoma cells. Therefore, in the present study we investigated the expression patterns of GPRC5A in clinical specimens, including normal bronchial epithelia (NBE) of COPD patients with and without overt lung cancer as well as in non-small cell lung cancer (NSCLC) tumors by immunohistochemical (IHC) analysis. We performed IHC analysis of a tissue microarray (TMA) comprised of 311 lung adenocarcinomas and 166 squamous cell carcinomas (SCCs), as well as of normal bronchial epithelial specimens from 50 patients with COPD, which included 24 cancer-free cases and 26 cases with NSCLC (12, adenocarcinoma; 12, SCC; 2, bronchioalveolar carcinoma). Kruskal-Wallis test was used to compare GPRC5A levels among histology levels. All statistical tests were two-sided, and p values of 0.05 or less were considered to be statistically significant. Cytoplasmic GPRC5A expression was significantly higher in lung adenocarcinomas compared to SCCs ($p < 0.001$). Moreover, GPRC5A expression exhibited a positive correlation with never-smoking status ($p = 0.005$). Interestingly, we noted a statistically significant inverse correlation between the expression of GPRC5A and that of NF- κ B ($p < 0.001$), which we had previously found to be activated and elevated following loss of the GPRC5A tumor suppressor. Furthermore, analysis of two NBE obtained from each of 50 COPD patients demonstrated statistically significant decreased expression of GPRC5A in NBE of COPD patients with NSCLC compared to NBE from NSCLC-free COPD patients ($p < 0.001$). Our findings demonstrate that decreased GPRC5A expression may be associated with development of lung malignancies, especially in individuals with chronic lung inflammation, and pinpoints GPRC5A's potential suppressive effects on a lung tumor-promoting microenvironment. Assessment of GPRC5A's potential use as a risk factor for NSCLC development in

COPD patients is warranted. Supported by the Samuel Waxman Cancer Research Foundation and by W81XWH-04-1-0142.

[Print this Page](#)

Presentation Abstract

Abstract
Number: 4817

Presentation Title: A five-gene signature predictive of survival in lung adenocarcinoma but not in squamous cell carcinoma

Presentation Time: Tuesday, Apr 20, 2010, 3:10 PM - 3:25 PM

Location: Room 204, Washington Convention Center

Author Block: Humam N. Kadara, Waun K. Hong, Ignacio I. Wistuba, Reuben Lotan. UT M.D. Anderson Cancer Ctr., Houston, TX

Abstract Body: Identification of novel molecular markers for lung carcinogenesis, outcome and response to therapy is expected to improve the clinical management of lung cancer such as non-small cell lung cancer (NSCLC). Previously, we have derived gene expression signatures indicative of differential gene expression among cells constituting an in vitro model of human lung carcinogenesis and relevant to survival in NSCLC. In this study we assess the prognostic efficacy of our previously described six genes by using several prediction algorithms and a leave-one-out-cross-validation (LOOCV) strategy as well as risk-score prediction models. The NCI Director's Challenge datasets (n=443) were used as a training set and gene expression data of adenocarcinomas from the Duke and Harvard cohorts (DH cohort; n=183) were pooled as a validation set. In addition, two independent published datasets comprised of 130 and 58 lung squamous cell carcinomas (SCCs) served as a SCC validation cohort. A Five-gene in vitro lung carcinogenesis model signature (FILM) classifier was derived and found to be superior in prediction as the lowest specificity or sensitivity of the six prediction algorithms was 0.943. Importantly, all six prediction algorithms showed that the overall survival of all-stage or stage-I only human lung adenocarcinoma patients in the DH validation cohort that expressed FILM was significantly poorer than that of patients predicted to lack the signature. Moreover, no differences in overall survival between lung SCCs predicted to express or lack FILM were observed demonstrating the prognostic specificity of this classifier for lung adenocarcinomas. We then developed a risk score-prediction model for lung adenocarcinoma based on the Cox regression coefficients and expression of FILM genes. Lung adenocarcinoma patients identified to be at high risk based on the FILM risk model exhibited significantly worse survival ($p=5.4 \times 10^{-7}$, 100 months follow-up) than patients at low risk. For validation of the FILM risk model, Cox regression coefficients and the

dichotomization cut-off threshold generated from the training cohort (n=443) were directly applied to the DH validation cohort (n=183). All stages or stage-I only lung adenocarcinoma patients in the validation cohort and predicted to be at high risk displayed significantly worse survival than patients predicted to be at low risk by the FILM risk model ($p=0.0006$ and $p=0.0005$ of the log-rank test, respectively). Our findings highlight a novel five-gene signature which, although derived originally from an in vitro cell model, is highly effective in predicting survival of lung adenocarcinoma patients. Studies to validate the effectiveness of FILM in predicting, in particular, the response of lung adenocarcinoma patients to various therapies are highly warranted. Supported by DOD grant W81XWH-04-1-0142 and NCI lung cancer SPORE (P50 CA70907).

American Association for Cancer Research
615 Chestnut St. 17th Floor
Philadelphia, PA 19106

Gene expression analysis of field of cancerization in early stage NSCLC patients towards development of biomarkers for personalized prevention

Humam Kadara, Pierre Saintigny, Youhong Fan, Chi-Wan Chow, ZouMing Chu, Wenhua Lang, Carmen Behrens, Kathryn Gold, Diane Liu, J. Jack Lee, Li Mao, Edward Kim, Waun Ki Hong, Ignacio I. Wistuba

Background: The identification of early stage non-small cell lung cancer (ES NSCLC) patients (pts) at higher risk for recurrence or second primary tumor (SPT) development is vital to personalizing prevention and therapy. We sought to decipher spatial and temporal patterns of gene expression in the airway field of ever-smoker ES NSCLC pts to better understand lung cancer pathogenesis and predict recurrence or SPT development.

Methods: Pts on the prospective Vanguard study had definitively treated ES (I/II) NSCLC, were current/former smokers, and had bronchoscopies with brushings obtained from the main carina (MC) at baseline, 12, and 24 months following resective surgery and from different anatomical regions at baseline. Expression profiling is ongoing for all eligible pts (41 pts, 326 samples). To query temporal and spatial airway expression profiles, two sets of six pts were selected based on complete processed time point and baseline airway site (3 different sites per pt) arrays (Affymetrix Human Gene 1.0 ST), respectively. Temporally and spatially differentially expressed genes were independently identified based on a $p < 0.01$ of a univariate t-test with estimation of the false discovery rate (FDR), studied by hierarchical clustering and principal component analysis (PCA), and functionally analyzed using network analysis.

Results: 871 gene features were differentially expressed among MCs of six NSCLC pts at baseline, 12, and 24 months and were shown to separately group the MCs as evident in both cluster and PC analyses. Moreover, pathways analysis of the temporally modulated genes showed that a gene-network mediated by extracellular regulated kinase (ERK1/2) was most significantly elevated ($p < 0.001$) in function between MCs at 24 months versus baseline. 763 and 931 gene features were differentially expressed between MCs and adjacent-to-resected tumors (ADJ) airways and between MC, ADJ, and non-adjacent (distant-to-resected tumor) (NON-ADJ) airways, respectively. Moreover, pathway analysis of the spatially modulated genes revealed that gene-networks mediated by nuclear factor- κ B (NF- κ B) and ERK1/2-mediated were most significantly elevated ($p < 0.001$) in function in ADJ airway samples versus MCs. Furthermore, PCA revealed that while ADJ airway samples grouped separately and closely together, one MC and 3 NON-ADJ airway samples resided closely with ADJ samples, which were then found to originate from 3 pts with evidence of recurrence, SPT, or suspicion of recurrence.

Conclusions: Our findings highlight expression signatures and pathways (ERK1/2 and NF- κ B) in a “cancerization field” that may drive lung cancer pathogenesis and be associated with recurrence or SPT development in ES NSCLC pts, and thus useful for derivation of biomarkers to guide personalized prevention strategies. Supported by DoD grants W81XWH-04-1-0142 and W81XWH-10-1-1007.

[Print this Page](#)

Presentation Abstract

Abstract
Number: 4127

Presentation
Title: ***EGFR and K-Ras mutations and resistance of lung cancer to the IGF-1R tyrosine kinase inhibitors***

Presentation
Time: Tuesday, Apr 20, 2010, 2:00 PM - 5:00 PM

Location: Exhibit Hall A-C, Poster Section 11

Poster
Section: 11

Poster
Board
Number: 1

Author
Block: Woo-Young Kim, Quanri Jin, Ludmilla Prudkin, Jin-Soo Kim, Floriana Morgillo, Lei Feng, Edward S. Kim, Bryan Hennessy, Ju-Seog Lee, Gordon Mills, J. J. Lee, Bonnie Glisson, Scott M. Lippman, Ignacio I. Wistuba, Ho-Young Lee. UT M.D. Anderson Cancer Ctr., Houston, TX

Abstract
Body: **Background:** The majority of patients with non-small cell lung cancer (NSCLC) has responded poorly to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs). We investigated (1) the involvement of insulin-like growth factor 1 receptor (IGF-1R) signaling in primary resistance to EGFR TKIs and (2) the molecular determinants of resistance to IGF-1R TKIs. **Methods:** Phosphorylated IGF-1R/insulin receptor (pIGF-1R/IR) was immunohistochemically evaluated in NSCLC tissue microarrays. The antitumor effects of IGF-1R TKIs (PQIP, OSI906), either alone or in combination with small-molecular inhibitors or siRNA targeting K-Ras or MAPK/extracellular signal-regulated kinase (MEK) were analyzed in vitro and in vivo in 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK)-transformed human bronchial epithelial (HBE) cells and in NSCLC cells with variable histologic features and mutations in EGFR or K-Ras. **Results:** pIGF-1R/IR expression in NSCLC specimens was positively correlated with presence of a history of tobacco smoking, squamous cell carcinoma, mutant (mut) K-Ras, and wild-type (wt) EGFR, all of which have been strongly associated with poor response to EGFR TKIs. IGF-1R TKIs exhibited significant antitumor

activity in NNK-transformed HBE cells and in NSCLC cells harboring wt EGFR and wt K-Ras, but not those with mutations in these genes. Introduction of mut K-Ras attenuated the effects of IGF-1R TKIs on wt K-Ras-expressing NSCLC cells. Conversely, inactivation of MEK restored sensitivity to IGF-1R TKI in cells carrying mut K-Ras. **Conclusions:** The mutation status of both EGFR and K-Ras could be a predictive marker for response to IGF-1R TKIs. Also, MEK antagonism can abrogate primary resistance of NSCLC to IGF-1R TKIs. This work was supported by NIH grants R01 CA-109520-01 and CA-100816 (all to H-YL.) and in part by DOD grant W81XWH-04-1-0142 VITAL and W8XWH-06-1-0303 BATTLE (W-K H.)

[American Association for Cancer Research](#)

**615 Chestnut St. 17th Floor
Philadelphia, PA 19106**

Control/Tracking Number: 11-A-6897-AACR

Activity: Abstract Submission

Current Date/Time: 11/16/2010 3:11:20 PM

CXCR2 expression in tumor cells is associated with an adverse outcome in a large set of non-small-cell lung cancer (NSCLC)

Short Title:

CXCR2 in lung cancer

Author Block Pierre Saintigny, Diane Liu, J. Jack Lee, Yuan Ping, Carmen Behrens, Luisa M. Solis Soto, John V. Heymach, Edward S. Kim, Waun K. Hong, Jonathan M. Kurie, Ignacio I. Wistuba, Ja Seok Koo. UT M.D. Anderson Cancer Ctr., Houston, TX

Abstract: Background: CXCR2 plays an important role in inflammation, and stimulation of CXCR2-expressing endothelial cells by ELR+ CXC chemokines promotes angiogenesis. Our goal was to study the expression of CXCR2 by tumor cells and its impact on prognosis in NSCLC. Material and Methods: CXCR2 expression was determined using immunohistochemistry and a large set of tissue microarray including 458 NSCLC. The association between cytoplasmic CXCR2 (cCXCR2) expression in tumor cells and clinico-pathological factors as well as survival was analyzed. Distribution of CXCR2 and its ligands (IL8, CXCL1, CXCL2, CXCL3, CXCL5, CXCL6 and CXCL7) gene expression was studied using publicly available gene expression profiles from 52 NSCLC cell lines (GSE4824) and 444 lung adenocarcinomas (adc) (NCI Director's Challenge). To summarize the effect of CXCR2/CXCR2 ligands biological axis, Principal Component Analysis (PCA) and unsupervised hierarchical clustering were performed using CXCR2 and its ligands gene expression in both cell lines and lung adc. The first Principal Component (PC1) was correlated (Pearson) with the whole genome in 52 NSCLC cell lines. All genes were ranked according to their correlation with PC1, and used for Gene Set Enrichment Analysis (GSEA) "pre-ranked" analysis. Results: Using the median of expression to dichotomize the patients in a high versus low expression group, 238 (52.1%) tumors expressed high cCXCR2. No association was observed with gender, race, smoking habits, histology, and stage. High cCXCR2 was associated with overall survival (Hazard ratio (HR) 1.5696; confidence interval (CI)=1.176-2.096, p-value=0.002) and recurrence-free survival (HR 1.321; CI=1.027-1.698, p-value=0.030) in a univariate Cox proportional hazards (CPH) model. High cCXCR2 remained significant for overall in a multivariate CPH after adjusting for age, gender, histology, stage, neoadjuvant chemotherapy for overall survival (HR 1.465; CI=1.088-1.972, p-value=0.012) and a trend was observed for recurrence-free survival (HR 1.261; CI=0.973-1.633, p-value=0.080). Gene expression distribution of CXCR2 and its ligands were strikingly similar in cell lines and lung adc. In both cases, hierarchical clustering showed a cluster mostly driven by CXCR2, CXCL5, and CXCL7, representing 20% of the samples. PC1 accounted for 48.25 and 46.15% of the variation of the PCA in cell lines and lung adc respectively. KRAS and NFkB oncogenic pathways were the top 2 gene sets associated with PC1. Using the median as a cutoff, PC1 was associated with a worse overall survival in 444 lung adc (Log-rank P=0.006). Conclusion: cCXCR2 expression in NSCLC tumor cells is frequent and associated with an adverse outcome. CXCR2/CXCR2 ligands biological axis may be associated with an activation of KRAS and NFkB pathways, and a poor prognosis in lung adc. Funding Source: Department of Defense-VITAL.

[Print this Page](#)

Presentation Abstract

Abstract
Number: 5166

Presentation
Title: **Sex determining region Y-box 2 (SOX2) is a potential cell-lineage gene highly expressed in the pathogenesis of squamous cell carcinomas of the lung**

Presentation
Time: Wednesday, Apr 21, 2010, 8:00 AM -11:00 AM

Location: Exhibit Hall A-C, Poster Section 12

Poster
Section: 12

Poster
Board
Number: 25

Author
Block: Ping Yuan¹, Humam Kadara¹, Carmen Behrens¹, Ximing Tang¹, Denise Woods¹, Luisa M. Solis¹, Jiaoti Huang², Monica Spinola³, Wenli Dong¹, Guosheng Yin¹, Junya Fujimoto¹, Edward Kim¹, Yang Xie³, Luc Girard³, Cesar Moran¹, Waun Ki Hong¹, John D. Minna³, Ignacio Ivan Wistuba¹. ¹UT M.D. Anderson Cancer Ctr., Houston, TX; ²David Geffen School of Medicine, Los Angeles, CA; ³The University of Texas Southwestern Medical Center, Dallas, TX

Abstract
Body: Non-small cell lung cancer (NSCLC) represents the majority (85%) of lung cancers and is comprised mainly of adenocarcinomas and squamous cell carcinomas (SCCs). The sequential pathogenesis of lung adenocarcinomas and SCCs occurs through dissimilar phases as the former tumors typically arise in the lung periphery whereas the latter normally arise near the central airway. We assessed the expression of SOX2, an embryonic stem cell transcriptional factor that also plays important roles in the proliferation of basal tracheal cells and whose expression is restricted to the main and central airways and bronchioles of the developing and adult mouse lung, in NSCLC by various methodologies. Here, we found that SOX2 mRNA levels, from various published datasets, were significantly elevated in lung SCCs compared to adenocarcinomas (all $p < 0.001$). Moreover, a previously characterized *OCT4/SOX2/NANOG* signature effectively separated lung SCCs from adenocarcinomas following integration with two independent publicly available gene expression microarray datasets and which correlated with increased SOX2 mRNA in SCCs. Immunohistochemical analysis of various histological lung tissue specimens demonstrated marked nuclear SOX2 protein expression in all normal bronchial epithelia, alveolar bronchiolization structures and premalignant lesions in SCC development (hyperplasia, dysplasia and carcinoma *in situ*) and absence of expression in all normal alveoli and atypical adenomatous hyperplasias. Moreover, SOX2 protein expression was greatly higher in lung SCCs compared to adenocarcinomas following analyses in two independent large tissue microarray (TMA) sets (TMA set I, $n=287$; TMA set II, $n=511$ both $p < 0.001$). Furthermore, amplification of SOX2 DNA was detected in 20% of lung SCCs tested ($n=40$) and in none of the adenocarcinomas ($n=17$). Our findings highlight a cell-lineage gene expression pattern for the stem cell transcriptional factor SOX2 in the pathogenesis of lung SCCs and raise the intriguing possibility of the growth addition of lung SCCs specifically to SOX2-dependent pathways. Supported in part by grants from the Department of Defense (W81XWH-04-1-0142 and W81XWH-07-1-03060), and the Specialized Program of Research Excellence in Lung Cancer grant P50CA70907.
Running title: SOX2 abnormalities in NSCLC

[American Association for Cancer Research](#)
615 Chestnut St. 17th Floor
Philadelphia, PA 19106