

Role for High-Affinity IgE Receptor in Prognosis of Lung Adenocarcinoma Patients

Dalam Ly^{1,2}, Chang-Qi Zhu³, Michael Cabanero³, Ming-Sound Tsao^{3,4}, and Li Zhang^{1,2,4}



Abstract

Cancer development and biology is influenced by the host immune system. Emerging data indicate that the context of immune cell infiltrates may contribute to cancer prognosis. However, the types of infiltrating immune cells that are critical for cancer development remain controversial. In attempts to gain insights into the immune networks that regulate and/or predict tumor progression, gene expression analysis was conducted on microarray datasets of resected tumor samples from 128 early-stage non-small cell lung cancer (NSCLC) adenocarcinoma patients. By limiting analysis to immune-related genes, we identified a 9-gene signature using MAXimizing R Square Algorithm that selected for the greatest separation between favorable and adverse prognostic patient subgroups. The prognostic value of this 9-gene signature was validated in

10 additional independently published microarray datasets of lung adenocarcinoma [$n = 1,097$; overall survival hazard ratio (HR), 2.05; 95% confidence interval, 1.64–2.56; $P < 0.0001$] and was found to be an independent prognostic indicator relative to tumor stage (overall survival HR, 2.09, 95% confidence interval, 1.65–2.66; $P < 0.0001$). Network analysis revealed that genes associated with Fcε complex (*FCER1*, *MS4A2*) formed the largest and most significant pathway of the signature. Using immunohistochemistry, we validated that *MS4A2*, the β subunit of the IgE receptor expressed on mast cells, is a favorable prognostic indicator and show that *MS4A2* gene expression is an independent prognostic marker for early-stage lung cancer patient survival. *Cancer Immunol Res*; 5(9); 1–9. ©2017 AACR.

Introduction

Lung cancer is a leading cause of cancer death, with approximately 85% being non-small cell lung cancer (NSCLC). Adenocarcinoma (ADC) is the most common form of NSCLC. Complete surgical resection is the standard treatment for early-stage I and II NSCLC patients, but up to 50% of these patients experience post-surgical recurrence (1). Cancer development is influenced by its microenvironment through interactions between cancer cells, stroma, vasculature and immune infiltrates (2, 3). Histologic evidence indicates that the presence of intratumoral lymphoid structures and immune cell infiltration are prognostic markers for NSCLC (4, 5). The number and types of tumor-infiltrating leukocytes (TIL) present within the tumor microenvironment influence tumor growth and metastasis. Indeed, the majority of

immune cell subsets can be found within lung tumors and many influence NSCLC prognosis (6–9).

High densities of CD8⁺ T cells, dendritic cells, and Th1 CD4⁺ T cells are correlated with beneficial prognosis, whereas regulatory T cells (Treg) have been shown to have detrimental prognostic effects (6, 7, 9). Though granulocytic populations, such as neutrophils, mast cells, and eosinophils make up a large part of resident innate immune cells in the lung, their role in NSCLC prognosis remains unclear or contradictory. High densities of neutrophils have been reported to have detrimental effects or no prognostic impact (7), whereas reports on mast cell densities have been contradictory, with studies showing both positive and negative prognostic roles (10, 11). With the clinical success of therapies that target the immune system, an understanding of how immune cells contribute to the microenvironment would complement staging and treatment options (12, 13).

Gene expression analysis has been used to reveal altered pathways and regulators of NSCLC tumorigenesis, with some signatures found to include immune-associated genes involved in recruitment and survival of immune cell subsets (14, 15). However, few signatures derived from resected tumor have described immune-related genes associated with the prognosis of NSCLC patients (14). Using an analysis that limited microarray gene expression data to curated immune-related genes, we identified a prognostic 9-gene signature that reveals a favorable prognostic role for expression of the high-affinity Fcε receptor.

Materials and Methods

Training and validation datasets

The immune-related gene signature was derived using the microarray dataset of 128 resected samples of lung adenocarcinoma patients who underwent surgical resection at University

¹Toronto General Research Institute, University Health Network, Toronto, Ontario, Canada. ²Department of Immunology, University of Toronto, Toronto, Ontario, Canada. ³Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada. ⁴Departments of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada.

Note: Supplementary data for this article are available at Cancer Immunology Research Online (<http://cancerimmunolres.aacrjournals.org/>).

D. Ly and C.Q. Zhu contributed equally to this article.

Corresponding Authors: Li Zhang, University of Toronto, MaRS Centre, Toronto Medical Discovery Tower, 2nd Floor Room 2-807, Toronto, Ontario M5G 1L7, Canada. Phone: 416-581-7521; Fax: 416-581-7515; E-mail: lzhang@uhnresearch.ca; and Ming-Sound Tsao, Princess Margaret Cancer Centre, Toronto Medical Discovery Tower, 101 College Street, Rm 14-401, Toronto, Ontario, M5G 1L7. Phone: 416-634-8722; E-mail: ming.tsao@uhn.ca

doi: 10.1158/2326-6066.CIR-16-0392

©2017 American Association for Cancer Research.

Health Network (UHN, Toronto, ON) between 1995 and 2005 (see Table 1 for demographics and analysis). This represents a subset of 181 stage I–II NSCLC patients whose gene expression data were used in a previously published prognostic signature validation study and have been deposited with GEO (GSE50081; refs. 16, 17). None of the patients received adjuvant therapy (see Table 1 for demographics and analysis). The profiling was performed using the Affymetrix U133 2.0 Plus arrays. The data were normalized using robust multiarray averaging. Patient samples were collected and analysis done under protocol 05-0221-T approved by the UHN Research Ethics Board. Signature validation was performed using 10 additional published microarray datasets (18–26) of NSCLC resected samples for which overall and disease-free survival data were available ($n = 1524$; Supplementary Table S1 for demographics).

Signature selection and validation

The "immune-related" gene list was obtained from the NIAID ImmPort Comprehensive List of Immune-Related Genes (version 2011_08_01) and mapped onto Affymetrix U133 Plus 2 array by querying Netaffy (<https://www.affymetrix.com/analysis/netaffy/>) using gene symbol (Supplementary Table S2). The mapped gene list contained 3,690 genes (6,437 probe sets) that were compiled using immune-related Gene Ontology (GO) keywords and included CD antigens, cytokines, chemokines, and their ligands

(27). Using this gene list, unsupervised hierarchical clustering using Genesis (28) showed two distinguishable clusters in the adenocarcinoma samples of the UHN training dataset. Probe sets that were significantly different between the two clusters were further introduced to a Cox proportional regression model for their associations with 5-year survival. A $P < 0.005$ was considered as significantly associated. For signature optimization, we used the Maximizing R Square Algorithm (MARSA; ref. 17) to select a minimum set of genes that had the highest independent ability to classify patients into high- and low-risk subgroups. The MARSA started with an exclusion procedure, which took all preselected probe-sets (ps) significantly associated with univariate 5-year survival (191 ps; Supplementary Table S3). An exclusion (backward) and an inclusion (forward) selection procedure were applied consecutively, the probe set against R-square value was plotted and a minimum number of probe sets (9 ps) having the largest c-statistic was identified.

Gene annotation and functional categorization

The Uniprot database was used to assign all gene annotations and functions. For GO and functional interaction networks, the 191 ps significantly associated with univariate 5-year survival were analyzed using DAVID Bioinformatics Resources v.6.8 or with the Reactome FI network plugin in Cytoscape. Networks were considered when at least two interacting partners were determined

Table 1. Clinical-pathologic markers and association to the 9-gene signature

	Total (%)	9-gene signature		P
		Low risk (n = 64)	High risk (n = 64)	
Age				
<65	40 (31.2)	68.4 ± 10.4	69.0 ± 8.9	0.23
≥65	88 (68.7)			
Gender				
Male	65 (50.7)	29 (45.3)	36 (56.2)	0.216
Female	63 (49.2)	35 (54.7)	28 (43.8)	
Smoking				
Never	23 (17.9)	16 (25.0)	7 (10.9)	0.064
Ex-smoker	56 (43.7)	30 (46.9)	26 (40.6)	
Current smoker	36 (28.1)	13 (20.3)	23 (35.9)	
Unknown	13 (10.1)	5 (7.8)	8 (12.5)	
Stage				
IA	36 (28.1)	27 (42.2)	9 (14.1)	0.0003
IB	56 (43.7)	25 (39.1)	31 (48.4)	
IIA	7 (5.4)	5 (7.8)	2 (3.1)	
IIB	29 (22.6)	7 (10.9)	22 (34.4)	
EGFR				
Wild-type	100 (78.1)	42 (65.6)	58 (90.6)	<0.0001
Mutant	26 (20.3)	22 (34.4)	4 (6.3)	
Unknown	2 (1.5)	0	2 (3.1)	
KRAS				
Wild-type	87 (67.9)	47 (73.4)	40 (62.5)	0.228
Mutant	40 (31.2)	17 (26.6)	23 (35.9)	
Unknown	1 (0.8)	0	1 (1.6)	
Histological pattern				
Acinar	28 (21.9)	16 (25.0)	12 (18.7)	0.002
Lepidic	5 (3.9)	3 (4.7)	2 (3.1)	
Papillary	32 (25.0)	23 (35.9)	9 (14.1)	
Micropapillary	4 (3.1)	2 (3.1)	2 (3.1)	
Solid	27 (21.1)	5 (7.8)	22 (34.4)	
Unknown	32 (25)	15 (23.4)	17 (26.6)	
Resection type				
Segmentectomy	5 (3.9)	3 (2.3)	2 (1.6)	0.357
Bisegmentectomy	1 (0.78)	0 (0)	1 (0.78)	
Lobectomy	120 (93.7)	61 (47.6)	59 (46.1)	
Pneumonectomy	2 (1.6)	0 (0)	2 (1.6)	

and related modules formed. CIBERSORT (29) analysis and resources (<https://cibersort.stanford.edu/>) were used to predict the fraction of immune cells within UHN adenocarcinoma datasets and Pearson correlation was done between the immune cell fraction and the 9-gene signature risk score to identify immune subset signatures responsible for prognosis in the dataset. For correlation to basophil or mast cell signatures, mouse gene lists identified as specific to basophil or mast cells (30) were mapped to Affymetrix U133 Plus 2 array probes through entrez gene ID (Supplementary Table S4). Correlations of the 9-gene signature with either basophil or mast cell scores were evaluated by Spearman correlation.

Immunohistochemistry and immunofluorescence

Immunohistochemistry (IHC) was performed at Princess Margaret Cancer Center Advanced Molecular Profiling Lab (AMPL) core facility, using the BenchMark XT automated stainer (Ventana Medical System) with antigen retrieval (CC1, Tris/Borate/EDTA pH8.0, #950-124) for 64 minutes. The anti-MS4A2 (clone: HPA059967, Atlas Antibodies) dilution was 1:200 and incubation was for 32 minutes. The Ventana's OptiView Detection Kit (#860-099) was utilized and the slides were counterstained with Gill modified hematoxylin. For immunofluorescence (IF) staining, a serial section was deparaffinized through successive incubations in xylene and decreasing concentrations of ethanol and antigens were retrieved in 10 mmol/L citrate buffer, pH 6.0, using the 2100 Retriever (Aptum Biologics, Ltd.). The sections were permeabilized with 0.1% Triton X-100/PBS, blocked with 3% bovine serum albumin (BSA)/0.1% Triton X-100/PBS for 1 hour at room temperature, and then incubated with the same MS4A2 antibody at 1:100 and anti-mast cell tryptase (clone AA1; Dako, 1:50) antibody diluted in 3% BSA/0.1% Triton X-100/PBS overnight at 4°C. After primary antibody incubation, cells were washed three times with 0.1% Triton X-100/PBS and incubated with secondary antibody in 3% BSA/0.1% Triton X-100/PBS for 2 hours at room temperature. Secondary antibodies were Alexafluor 488 goat anti-rabbit (Invitrogen, 1:500) and Alexafluor 568 goat anti-mouse (Invitrogen, 1:500). After secondary antibody staining, cells were washed three times with 0.1% Triton X-100/PBS and mounted in Vectashield mounting media containing DAPI (Vector Laboratories). Images were taken with a Zeiss Axioimager Z1 using an AxioCam MRm and Axiovision software.

Histologic staining and quantification of MS4A2

The hematoxylin and eosin (H&E)-stained sections of the training dataset tumors were evaluated for inflammatory pathology. The grade of inflammatory severity was semiquantitatively assessed on scale from 1 to 5: 1 = minimal, 2 = minimal to moderate, 3 = moderate, 4 = moderate to severe, 5 = severe. The pattern of lymphoplasmacytic infiltration was also assessed and categorized to be predominantly localized to tumor and/or stromal compartments. No cases were observed that could be categorized with high tumor but low stromal infiltrate. For MS4A2-positive counts, the average number of MS4A2-positive cells per high-power field (HPF) was determined. Five to 10 HPFs of random intratumoral inflammatory areas were assessed at a magnification of 400×. The results were assessed by a pathologist (MC) and expressed as the mean value of mast cells per HPF (objective 40×) per case.

Statistical analysis

Five-year overall survival (OS) and disease-free survival (DFS) were calculated from the date of surgery until death/recurrence or last follow-up. Survival curves were estimated using Kaplan-Meier method and compared using the log-rank test. Univariate and multivariate survival analyses were estimated by Cox proportional hazard regression using SAS v9.4 (SAS Institute). Graphs were constructed in GraphPad Prism 5.

Results

Identification of an immune-related prognostic gene signature

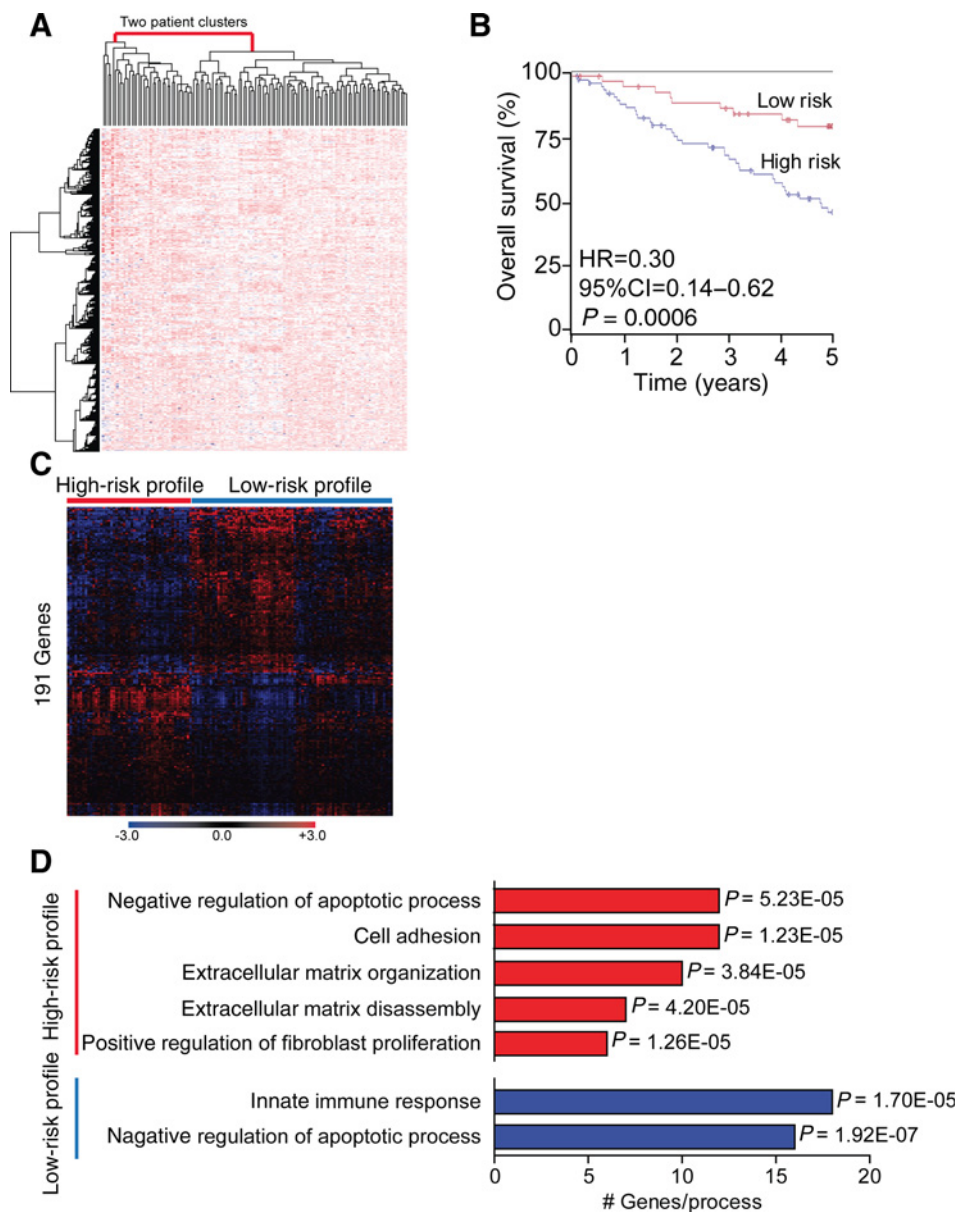
We hypothesized that by limiting initial analysis of microarray datasets to immune-related genes, we would identify prognostic immunological biomarkers in lung adenocarcinoma. Using a curated immune-related gene list, unsupervised hierarchical clustering identified two patient clusters within the UHN adenocarcinoma training set (Table 1; Fig. 1A), which differ significantly in predicting patient survival (Fig. 1B; HR, 0.30; 95% CI, 0.14–0.62, log rank $P = 0.0006$). Within the two ADC patient clusters, 2,782 probe sets, representing 1,216 genes as identified by the Jetset v1.6 (31) were differentially ($P < 0.001$) expressed. When further filtered by their univariate association with OS, 191 genes segregated patients into two clusters that differentiated patients into high or low risk for survival (Fig. 1C).

To determine whether genes associated with OS were enriched in immune-related pathways, we applied GO enrichment analysis using DAVID Bioinformatics Resources (32). From the 191-gene list, 89 and 102 genes were overexpressed in high- and low-risk subgroups, respectively (Fig. 1C; Supplementary Table S3). Of note, 86% (77/89) of the overexpressed genes within the high-risk subgroup were enriched within 63 GO categories at $P < 0.05$. After correcting for false-discovery rate (Benjamini FDR < 0.05), five distinct biological process categories emerged showing enrichment of terms consistent with cellular regulation (Fig. 1D). Amongst the low-risk (favorable) prognostic subgroup, 71% (72/102) of genes enriched amongst 55 categories at $P < 0.05$ were found, and after Benjamini FDR < 0.05 correction, two biological process categories remained, including a broadly enriched immunological process, namely, *innate immune response* (Fig. 1D).

To focus on relevant genes and remove noise, we focused on 191 genes that show univariate association with survival. MARSA analysis of the 191 genes reveals a 9-gene risk score based on the expression of genes *MS4A2*, *DAPK2*, *MKKS*, *CTSL2*, *LAIR2*, *ALOX15B*, *APITD1*, *CEBPA*, and *PSME2* (Supplementary Table S5). This 9-gene signature stratified ADC patients into high- or low-risk OS ($P < 0.0001$, Fig. 2A) and DFS groups ($P < 0.0001$, Fig. 2B). Clinicopathologic associations were analyzed between patients within high- or low-risk subgroups and associations with stage, EGFR mutational status, and histological patterns were observed (Table 1). After adjustment for clinicopathologic indicators, multivariate analysis indicate that the 9-gene signature remain an independent prognostic indicator within the training dataset (Supplementary Table S6).

Validation of immune-related 9-gene signature

The 9-gene signature was validated using 10 publicly available NSCLC datasets (Supplementary Table S1), for which patients' OS and/or DFS were available (17–26). Cumulative analysis of the validation cohort found the 9-gene signature could significantly predict OS ($n = 1386$; HR, 1.78; 95% CI, 1.51–2.09; $P < 0.0001$)

**Figure 1.**

Identification of prognostic immune-related genes in early-stage lung adenocarcinoma. **A**, Hierarchical clustering of the UHN adenocarcinoma NSCLC dataset ($n = 128$) with the immune-related gene list. **B**, Kaplan-Meier OS using the immune-related gene list. **C**, Heatmap of genes filtered for univariate association with OS at $P < 0.01$ using a Cox proportional regression model. **D**, GO enrichment analysis of genes associated with univariate survival. Data display enriched ontologies with Benjamini FDR < 0.05 .

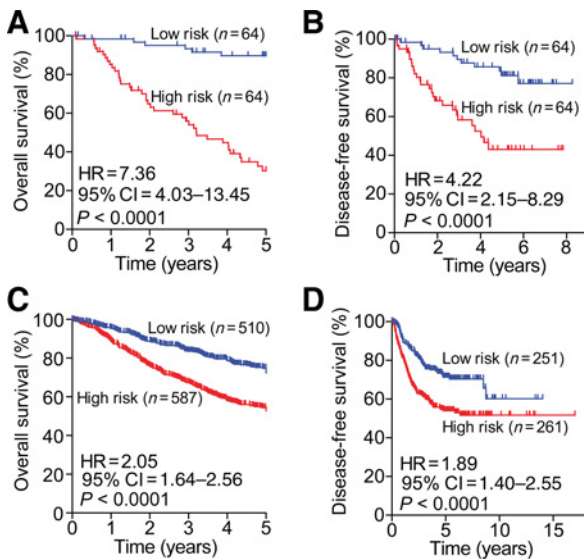
and DFS ($n = 650$; HR, 1.57; 95% CI, 1.27–1.95; $P < 0.0001$) in validation datasets when all patient subtypes were analyzed. However, when the validation cohort was stratified by histologic subtypes, the 9-gene signature was only significantly prognostic in ADC but not squamous cell carcinoma or other NSCLC subtypes ($n = 1097$; HR, 2.05; 95% CI, 1.64–2.56; $P < 0.0001$; Supplementary Table S7). Therefore, consistent with the training cohort, the 9-gene signature was able to separate ADC patients into high or low risk for OS and DFS (Fig. 2C and D).

The 9-gene signature remained significant after adjustment for disease stage, both for OS (HR, 2.09; 95% CI, 1.65–2.66; $P < 0.0001$) and DFS (HR, 1.62; 95% CI, 1.19–2.21; $P = 0.0021$) in the validation cohort (Table 2). As not all validation datasets had patients' age, gender, and smoking history available, these factors were not introduced in the multivariate model. To gain insight into the signature's performance, separate validation with ADC

patients of individual dataset were performed. The signature significantly ($P < 0.05$) predicted OS in 5 of 9 datasets and DFS in 2 of 4 datasets (Supplementary Table S8). Individual validation datasets that did not show significance had fewer than 100 patients.

Correlation with stromal lymphoplasmacytic localization

We further sought to determine if this signature correlated with the presence and location of tumor infiltrating leukocytes. Analysis of the degree of immune cell infiltration with the 9-gene risk score showed a statistically significant correlation with increasing infiltration score for the tumors of the training set (Spearman correlation coefficient $r = 0.24$, $P = 0.019$), suggesting a direct relationship between the 9-gene risk score and extent of lymphoplasmacytic infiltration. As location of immune subsets may have impact on their prognostic utility (6, 7), the location of the

**Figure 2.**

Validation of the 9-gene signature. **A**, Kaplan-Meier survival plots showing OS and **B**, DFS of patients in training set classified into high- or low-risk survival groups by 9-gene immune-related signature. **C**, Kaplan-Meier survival plots showing OS and **D**, DFS of patients within validation datasets classified into high- or low-risk survival groups by 9-gene immune-related signature. The number of patients in analysis is shown (*n*).

lymphoplasmacytic infiltrate was also scored. Three types of infiltration were observed: a diffuse lymphoplasmacytic infiltration confined to the tumor stroma, infiltration into tumor epithelia and stroma, and samples with minimal infiltration in either location. Samples with high 9-gene risk scores had greater tumoral infiltration, whereas samples with low risk scores were associated with stromal leukocyte infiltration (Fig. 3A). These results suggest that immune infiltrates associated with favorable patient prognosis are predominantly located in the tumor stroma.

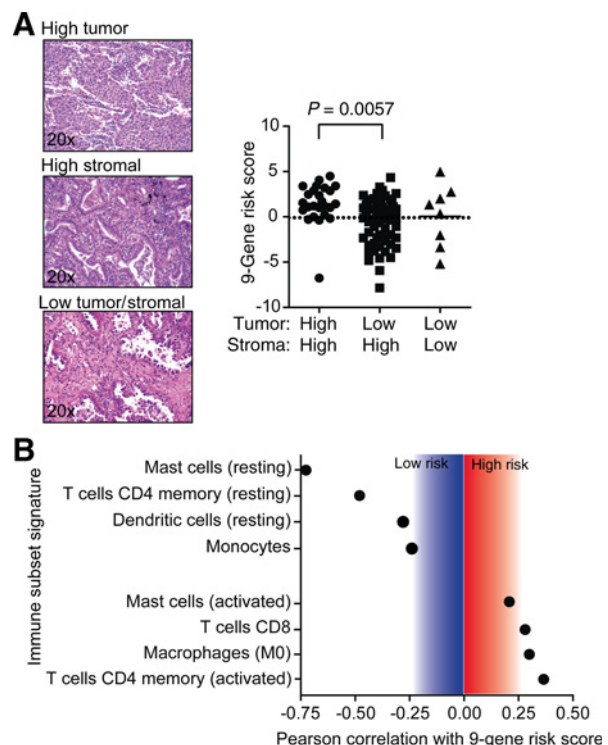
Immune cell subset genes enriched in the 9-gene signature

To investigate which immune cell subsets might correlate to the 9-gene signature, we took advantage of leukocyte-specific gene sets comprising of 547 genes that distinguish 22 human hematopoietic cell phenotypes and CIBERSORT algorithm to estimate immune cells within a reference immune cell dataset (29, 33). Applying Pearson correlation to the 9-gene risk score, we found that 8 immune cell subset genes significantly correlated with the 9-gene signature at $P < 0.05$. $CD4^+$ resting memory T cells, DCs, and monocytes are correlated with low risk scores whereas activated $CD4^+$ memory, $CD8^+$ T cells, and macrophages correlated with high risk scores albeit with Pearson coefficients at <0.5 . The

highest relation to the 9-gene risk score were genes enriched for "resting mast cells" that associated with low risk scores (Pearson coefficient of -0.73); genes enriched for "activated mast cells" associated with higher risk scores, but with a low Pearson coefficient of 0.2 (Fig. 3B). These results suggest that the 9-gene risk score may reveal a role for mast cells in the prognosis of lung adenocarcinoma patients.

IgE receptor pathway enrichment among favorably prognostic genes

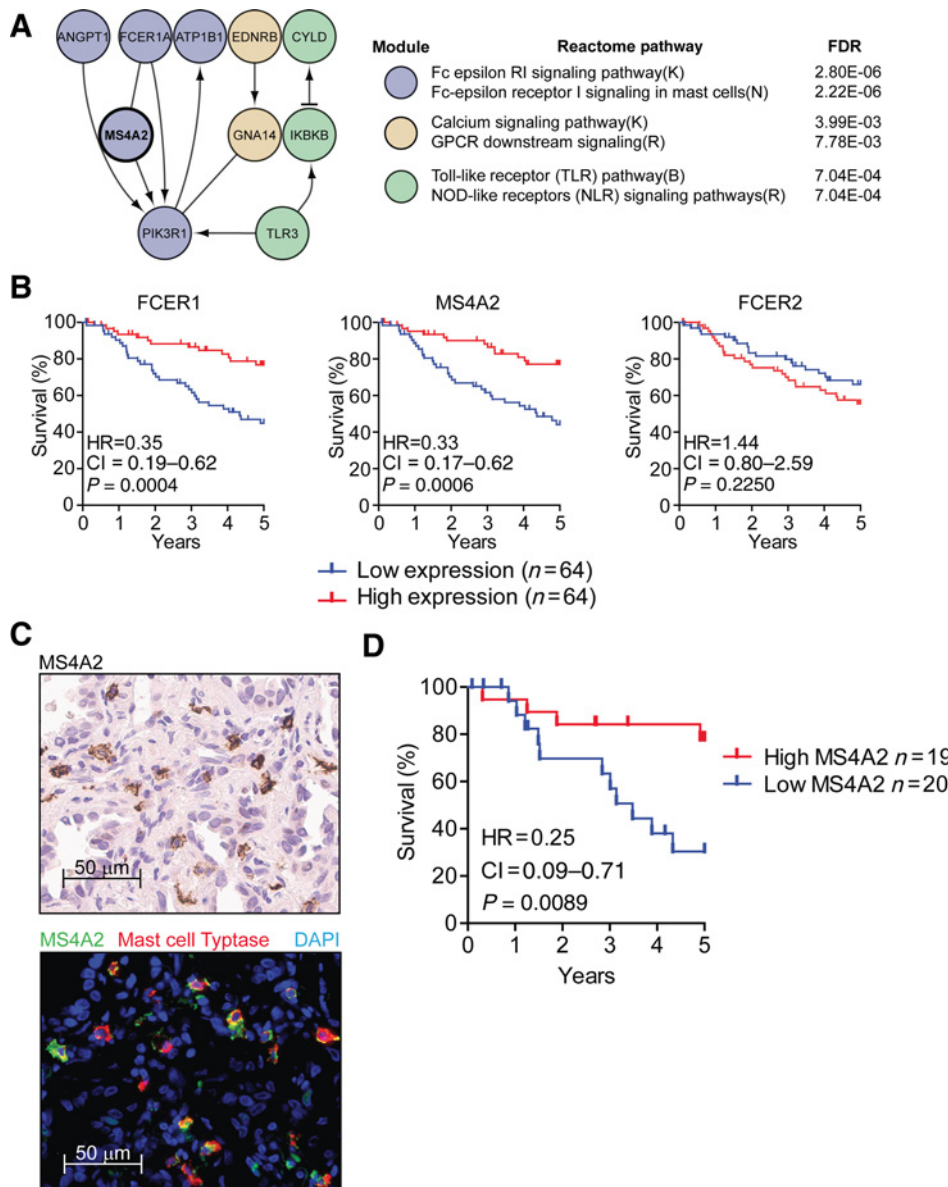
With the finding that mast cell-enriched genes correlated well with the 9-gene risk score, and that GO analysis revealed innate immune response as the most enriched ontology (Fig. 1E), we further explored potential pathways using the Reactome Functional Interaction (FI) network (34). Among genes associated with univariate favorable survival (Fig. 1C), the largest functional network formed consisted of 3 modules (with 10 interacting partners) enriched in innate immune signaling pathways, including *Fc epsilon RI signaling pathway* (FDR = $2.80E-06$), *calcium signaling pathway* (FDR = $3.99E-03$), and *Toll-Like Receptor 3 signaling* (FDR = $7.04E-04$; Fig. 4A). Consistent with innate immune responses, the pathways identified in the FI network are

**Figure 3.**

Immune-related gene signature correlate with lymphoplasmacytic infiltrate location and immune cell subset genes. **A**, Representative H&E stains of stromal or tumoral location of lymphoplasmacytic infiltrate within patient tumor. The 9-gene signature risk score correlates with predominant location of infiltrate. Of note, no cases were observed that could be categorized with high tumor but low stromal infiltrate. Ninety-six cases with available H&E-stained slides were analyzed. Significance was calculated using one-way ANOVA with Bonferroni posttest correction. **B**, CIBERSORT analysis with Pearson correlation of immune cell subsets to the 9-gene signature risk score, only subsets with significant correlations ($P < 0.05$) are displayed.

Table 2. Multivariate analysis showing the effect of prognostic factors on 9-gene signature within the validation dataset

Independent variable	HR (95% CI)	P
Overall survival		
9-gene signature	2.09 (1.65–2.66)	<0.0001
Stage (stage I vs. II)	1.46 (1.33–1.60)	<0.0001
DFS		
9-gene signature	1.62 (1.19–2.21)	0.0021
Stage (Stage I vs. II)	1.90 (1.54–2.35)	<0.0001

**Figure 4.**

MS4A2 expression is a prognostic biomarker for lung adenocarcinoma. **A**, Reactome functional interaction network generated from genes filtered for favorable univariate association with OS. Genes forming the largest network of at least two interacting partners with FDR < 0.01 are displayed. Genes are colored by their enrichment in the Reactome category and bolded genes represent genes from 9-gene signature. **B**, Univariate Kaplan-Meier survival plots for expression of IgE receptor complex genes. **C**, Immunohistochemical and immunofluorescence staining for MS4A2 and mast cell tryptase. **D**, Kaplan-Meier survival plot of patients expressing high and low MS4A2 cells. Representative IHC and IF stains amongst the number of patients analyzed are shown (n).

known to function within mast cells and basophils, specifically *Fc epsilon RI* (high-affinity IgE receptor) signaling (35–37).

As the CIBERSORT leukocyte reference gene set did not include basophil-enriched genes, we sought to determine which high-affinity IgE receptor expressing immune subset most correlated with the 9-gene signature. Utilizing murine basophil and mast cell-specific gene signatures (30) we mapped 98 of 99 mast cell-specific genes and 58 of 61 basophil-specific genes to the human Affymetrix U133A microarray and determined their correlation to the 9-gene signature. Consistent with CIBERSORT analysis, the 9-gene risk score significantly correlated with mast cell-enriched genes ($P < 0.0001$, Spearman correlation coefficient $r = -0.48$), but even with shared expression of the IgE receptor, the basophil-enriched gene signature was not correlated to the 9-gene risk score ($P = 0.65$, Spearman correlation coefficient $r = 0.04$).

MS4A2-expressing mast cells are favorably prognostic in lung adenocarcinoma

With the finding that FI network enrichment of *Fc epsilon RI* signaling pathway likely derived from a mast cells signature, we sought to determine whether individual mast cell genes may contribute to prognosis. The high-affinity IgE receptor is expressed as a tetrameric complex consisting of an α (*FCER1*), β (*MS4A2*), and two γ (*FCER1G*) subunits (35, 37). As genes encoding the α and β subunit formed the largest functional interaction network and as *MS4A2* is part of the 9-gene signature, this receptor likely has prognostic roles. Univariate analysis confirmed that high-affinity IgE receptor genes (*FCER1*, *MS4A2*) alone were prognostic within the ADC dataset, whereas the low-affinity IgE receptor (*FCER2*) and shared IgE γ (*FCER1G*) subunit were not (Fig. 4B and Supplementary Fig. S1A). Similarly, univariate analysis of *TLR3* expression, which formed a smaller node related to the IgE

Table 3. Multivariate analysis showing the effect of prognostic factors on MS4A2 expression

Independent variable	HR (95% CI)	P
MS4A2 (high vs. low risk)	0.24 (0.1–0.59)	0.002
Age (over 65 vs. below)	1.18 (0.51–2.77)	0.699
Gender (male vs. female)	0.57 (0.25–1.28)	0.173
Smoking (ever vs. never)	1.48 (0.40–5.51)	0.558
Stage (Stage I vs. II)	2.57 (1.15–5.74)	0.022
Histological pattern (acinar/papillar/ lepidic vs. micropapillary/solid)	0.57 (0.24–1.34)	0.197

receptor signaling, was also prognostic (Supplementary Fig. S1A). Mast cell–specific protease genes (*TSPAB1*, *CPA3*, *CMA1*), although not forming an FI network, are important effector components (38). Amongst three mast cell–enriched protease genes, only *CPA3* (carboxypeptidase 3) expression was prognostic. *TSPAB1* (tryptase α/β 1) did not reach statistical significance, and *CMA1* (chymase 1) was not prognostic (Supplementary Fig. S1A).

With the finding that subunits of the high-affinity IgE receptor were prognostic and that *MS4A2* was a member of the 9-gene signature, a subgroup of high and low genomic expressing *MS4A2* tumors from the training dataset was studied by immunohistochemistry for *MS4A2*-positive cells. The number of *MS4A2* positive cells ranged from 0.6 to 42.8 cells per average HPF ($40\times$) and significantly correlated with the *MS4A2* gene expression ($P < 0.0001$, Spearman correlation coefficient $r = 0.65$, Supplementary Fig. S1B). Furthermore, amongst the samples analyzed *MS4A2*-positive cells were localized in the tumor stroma, were morphologically consistent with mast cells, and costained with the mast cell–specific protease, mast cell tryptase (ref. 38; Fig. 4C). Consistent with gene expression analysis (Fig. 4B), the number of *MS4A2*-positive cells was prognostic for patient survival (Fig. 4D). Given high correlation between *MS4A2* gene expression and IHC cell counts, we explored whether *MS4A2* gene expression associated with other clinical factors. Univariate analysis shows *MS4A2* expression associated with smoking and histological subtype (Supplementary Table S9), and in multivariate analysis, *MS4A2* expression alone emerged as an independent prognostic marker (Table 3). Taken together, these data indicate that the high-affinity IgE receptor gene *MS4A2*, expressed by tumor-infiltrating mast cells, is a favorable prognostic indicator for lung adenocarcinoma patients.

Discussion

By limiting our analysis of lung adenocarcinoma gene expression data to immune-related genes, we have identified a 9-gene signature that was prognostic. We validated the signature in a combined cohort of 10 independent microarray datasets. After adjustment for stage, the signature remained an independent prognostic indicator within the validation dataset. The signature was correlated to lymphoplasmacytic infiltrates, with favorable prognosis correlating with higher stromal localization of leukocytes. GO revealed that genes associated with favorable prognosis correlated with innate immune responses and were enriched in genes found in mast cells, with significance in pathways involving high-affinity IgE receptor signaling. Amongst the 9-gene signature, *MS4A2*, the β subunit of the IgE receptor was found to be highly expressed in patients with favorable prognosis and alone was an independent prognostic indicator. By immunohistochemistry, we

validated that tumors with high numbers of *MS4A2*-positive cells correlated with the 9-gene signature prognosis and *MS4A2* expression.

The 9-gene signature presented here differs from previous prognostic NSCLC signatures (17–26) in that we began signature discovery using an immune-related gene list which contained pan-immune markers and immune-related signaling pathways (27). From this initial filtering, we further curated genes associated with univariate survival using the MARS algorithm (17). This method identified prognostic immune-related genes that did not overlap with previous studies (14, 17–26). This is consistent with previous meta-analyses of prognostic NSCLC signatures. Amongst 17 studies reviewed, 84 immune-related genes were found, with only 7 overlapping genes found in at least two signatures (14). The lack of overlap amongst gene signatures may reflect some variation in the patient cohorts, intratumoral heterogeneity, and analysis methods (39). Nonetheless, we found the 9-gene signature to be an independent prognostic indicator cumulatively validated in 10 independently collected patient cohorts.

GO analysis revealed that the genes most associated with favorable prognosis were those enriched in the "innate immune response" category. Specifically, CIBERSORT analysis indicated the 9-gene risk score correlated with genes enriched within mast cells. Indeed, amongst the 547 genes that make up the CIBERSORT immune cell reference dataset (29), three genes (*MS4A2*, *LAIR2*, *DAPK2*) of the 9-gene signature were shared, with *MS4A2* being expressed more by mast cells than other immune cells, with the highest expression amongst "resting" mast cells (29). Though other immune cell subset genes were found to be significantly associated with the 9-gene signature, Pearson coefficients were low (<0.5), suggesting that not all genes enriched with the individual subsets correlated to the 9-gene signature. Additionally, FI network analysis found enrichment for high-affinity IgE receptor signaling pathway amongst the 9-gene signature genes, a pathway known to function within mast cells and basophils (35, 37).

As a multimeric complex the high-affinity IgE receptor exists in tetrameric and trimeric isoforms with an FcR γ chain that is expressed with many Fc receptors (35, 37, 40). The tetrameric IgE isoform contain one α , one β , and two γ subunits, whereas the trimeric isoform contains one α and two γ subunits, but lacks a β subunit (40). In humans, the trimeric isoform can be found on dendritic cells and monocytes, whereas the tetrameric isoform is found on mast cells and basophils. Thus, whereas the α and γ subunits are expressed in different immune cells, the β subunit is enriched in mast cells and basophils (35). As the CIBERSORT immune cell reference dataset did not include basophil-specific genes, we utilized a murine basophil-specific gene set (30) which mapped to $>95\%$ of human basophil genes and found that the 9-gene signature correlated with mast cells but not basophil-enriched genes, suggesting that the FI network enrichment of IgE signaling is most likely mast cell derived. This was further validated by showing that *MS4A2* expression in lung adenocarcinoma colocalized with mast cell tryptase, a mast cell–specific protease (38). Additionally, we found that *MS4A2* gene expression alone was an independent prognostic indicator, and expression correlated with *MS4A2* cell counts in IHC. Finally, our data show that increased *MS4A2*⁺ cell counts were correlated with favorable prognosis in lung adenocarcinoma. These findings support the notion that mast cells have favorable prognostic roles in lung adenocarcinoma.

The impact of mast cells within the tumor microenvironment remains unclear due to the contradictory observations the presence of mast cells have on lung cancer prognosis, reviewed in references 10 and 11. Due to the plethora of mast cell-secreted factors, which include proteases, angiogenic factors, and immunogenic cytokines, studies have shown that high densities of mast cells correlate with either adverse (41–43) or favorable prognosis (44–48). Protumorigenic role for mast cells observed greater mast cell densities within areas of increased microvessel density (MCD) and within vascular endothelial growth factor (VEGF)-positive lung carcinoma, resulting in worse prognosis presumably due to mast cell secretion of tumor angiogenic factors (41–43). Other studies have shown that high total mast cell densities (44, 45, 48) or those localized to tumor islets (46, 47) result in favorable prognostic outcomes. Tumor islet mast cells colocalized with TNF α may be cytolytic to tumor cells, or result in mast cell recruitment of CD8⁺ T cells and macrophages (47). The prognostic differences observed by these studies may be due to whether a study scored for microlocalization or which methods were used for mast cell detection; most such methods rely on the presence of mast cell proteases. Human mast cell subsets have been described that vary in protease expression and activation state; degranulation of proteases may underestimate mast cells counts (38, 44). Our data indicate that the 9-gene signature most correlated with genes enriched for "resting" mast cells and that not all mast cell protease genes were prognostic, suggesting that activation status of mast cells may be a consideration for mast cell identification. As IgE receptor is expressed on mast cells (37), MS4A2 may represent a more consistent prognostic biomarker for mast cell identification.

In summary, by limiting analysis to immune-related gene lists we identified a 9-gene signature that predicts 5-year survival of lung adenocarcinoma. The 9-gene signature correlated with "resting mast cell" genes, with MS4A2 expression on stromal mast cells found to be a favorable prognostic indicator for early-stage lung cancer patient survival, suggesting that mast cells shape lung cancer development.

References

1. Uramoto H, Tanaka F. Recurrence after surgery in patients with NSCLC. *Transl Lung Cancer Res* 2014;3:242–9.
2. Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. *Nat Rev Cancer* 2009;9:239–52.
3. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–74.
4. Brambilla E, Le Teuff G, Marguet S, Lantuejoul S, Dunant A, Graziano S, et al. Prognostic effect of tumor lymphocytic infiltration in resectable non-small-cell lung cancer. *J Clin Oncol* 2016;34:1223–30.
5. Dieu-Nosjean MC, Antoine M, Danel C, Heudes D, Wislez M, Poulot V, et al. Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. *J Clin Oncol* 2008;26:4410–7.
6. Bremnes RM, Busund LT, Kilvaer TL, Andersen S, Richardsen E, Paulsen EE, et al. The role of tumor-infiltrating lymphocytes in development, progression, and prognosis of non-small cell lung cancer. *J Thorac Oncol* 2016;11:789–800.
7. Remark R, Becker C, Gomez JE, Damotte D, Dieu-Nosjean MC, Sautes-Fridman C, et al. The non-small cell lung cancer immune contexture. A major determinant of tumor characteristics and patient outcome. *Am J Respir Crit Care Med* 2015;191:377–90.
8. Fridman WH, Pages F, Sautes-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer* 2012;12:298–306.
9. Suzuki K, Kadota K, Sima CS, Nitadori J, Rusch VW, Travis WD, et al. Clinical impact of immune microenvironment in stage I lung adenocarcinoma: tumor interleukin-12 receptor beta2 (IL-12Rbeta2), IL-7R, and stromal FoxP3/CD3 ratio are independent predictors of recurrence. *J Clin Oncol* 2013;31:490–8.
10. Marichal T, Tsai M, Galli SJ. Mast cells: potential positive and negative roles in tumor biology. *Cancer Immunol Res* 2013;1:269–79.
11. Oldford SA, Marshall JS. Mast cells as targets for immunotherapy of solid tumors. *Mol Immunol* 2015;63:113–24.
12. Donnem T, Kilvaer TK, Andersen S, Richardsen E, Paulsen EE, Hald SM, et al. Strategies for clinical implementation of TNM-Immunoscore in resected non-small-cell lung cancer. *Ann Oncol* 2016;27:225–32.
13. Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer* 2016;16:275–87.
14. Suzuki K, Kachala SS, Kadota K, Shen R, Mo Q, Beer DG, et al. Prognostic immune markers in non-small cell lung cancer. *Clin Cancer Res* 2011;17:5247–56.
15. Zhu C-Q, Tsao M-S. Prognostic markers in lung cancer: is it ready for prime time? *Transl Lung Cancer Res* 2014;3:149–58.
16. Der SD, Sykes J, Pintilie M, Zhu CQ, Strumpf D, Liu N, et al. Validation of a histology-independent prognostic gene signature for early-stage, non-small-cell lung cancer including stage IA patients. *J Thorac Oncol* 2014;9:59–64.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: D. Ly, C.-Q. Zhu, M. Cabanero, M.-S. Tsao, L. Zhang

Development of methodology: D. Ly, C.-Q. Zhu, M. Cabanero, M.-S. Tsao, L. Zhang

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D. Ly, C.-Q. Zhu, M. Cabanero, M.-S. Tsao

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D. Ly, C.-Q. Zhu, M. Cabanero, M.-S. Tsao, L. Zhang

Writing, review, and/or revision of the manuscript: D. Ly, C.-Q. Zhu, M. Cabanero, M.-S. Tsao, L. Zhang

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D. Ly, C.-Q. Zhu, M.-S. Tsao

Study supervision: M.-S. Tsao, L. Zhang

Acknowledgments

The authors would like to acknowledge Jing Xu at Princess Margaret Cancer Center Advanced Molecular Profiling Lab (AMPL) core facility for help in optimization of IHC staining and Dr. Nadeem Moghal for aid in immunofluorescence analysis.

Grant Support

This work was supported by the Canadian Institutes of Health Research grant (to L. Zhang) Canadian Cancer Society IMPACT grant (#704021; to L. Zhang), Canadian Cancer Society IMPACT grant (#701595; to M.S. Tsao), and the Terry Fox Foundation Special Training Initiative in Health Research in Molecular Pathology of Cancer at CIHR (STP 53912; to M.S. Tsao). L. Zhang is the Inaugural Maria H. Bacardi Chair in Transplantation, and M.S. Tsao is the M. Qasim Choksi Chair in Lung Cancer Translational Research.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 28, 2016; revised April 17, 2017; accepted July 25, 2017; published OnlineFirst August 3, 2017.

17. Zhu CQ, Ding K, Strumpf D, Weir BA, Meyerson M, Pennell N, et al. Prognostic and predictive gene signature for adjuvant chemotherapy in resected non-small-cell lung cancer. *J Clin Oncol* 2010; 28:4417–24.
18. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006;439:353–7.
19. Chitale D, Gong Y, Taylor BS, Broderick S, Brennan C, Somwar R, et al. An integrated genomic analysis of lung cancer reveals loss of DUSP4 in EGFR-mutant tumors. *Oncogene* 2009;28:2773–83.
20. Director's Challenge Consortium for the Molecular Classification of Lung A, Shedden K, Taylor JM, Enkemann SA, Tsao MS, Yeatman TJ, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008;14:822–7.
21. Hou J, Aerts J, den Hamer B, van Ijcken W, den Bakker M, Riegman P, et al. Gene expression-based classification of non-small cell lung carcinomas and survival prediction. *PLoS One* 2010;5:e10312.
22. Lee ES, Son DS, Kim SH, Lee J, Jo J, Han J, et al. Prediction of recurrence-free survival in postoperative non-small cell lung cancer patients by using an integrated model of clinical information and gene expression. *Clin Cancer Res* 2008;14:7397–404.
23. Rousseaux S, Debernardi A, Jacquiou B, Vitte AL, Vesin A, Nagy-Mignotte H, et al. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci Transl Med* 2013;5:186ra66.
24. Tang H, Xiao G, Behrens C, Schiller J, Allen J, Chow CW, et al. A 12-gene set predicts survival benefits from adjuvant chemotherapy in non-small cell lung cancer patients. *Clin Cancer Res* 2013;19:1577–86.
25. Wilkerson MD, Yin X, Walter V, Zhao N, Cabanski CR, Hayward MC, et al. Differential pathogenesis of lung adenocarcinoma subtypes involving sequence mutations, copy number, chromosomal instability, and methylation. *PLoS One* 2012;7:e36530.
26. Yamauchi M, Yamaguchi R, Nakata A, Kohno T, Nagasaki M, Shimamura T, et al. Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma. *PLoS One* 2012;7:e43923.
27. Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, et al. ImmPort: disseminating data to the public for the future of immunology. *Immunol Res* 2014;58:234–9.
28. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. *Bioinformatics* 2002;18:207–8.
29. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015;12:453–7.
30. Dwyer DF, Barrett NA, Austen KF. Expression profiling of constitutive mast cells reveals a unique identity within the immune system. *Nat Immunol* 2016;17:878–87.
31. Li Q, Birkbak NJ, Gyorffy B, Szallasi Z, Eklund AC. Jetset: selecting the optimal microarray probe set to represent a gene. *BMC Bioinformatics* 2011;12:474.
32. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003;4:P3.
33. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* 2015;21:938–45.
34. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, et al. The Reactome pathway Knowledgebase. *Nucleic Acids Res* 2016;44(D1):D481–7.
35. Wu LC, Zarrin AA. The production and regulation of IgE by the immune system. *Nat Rev Immunol* 2014;14:247–59.
36. Sandig H, Bulfone-Paus S. TLR signaling in mast cells: common and unique features. *Front Immunol* 2012;3:185.
37. Kraft S, Kinet JP. New developments in FcεRI regulation, function and inhibition. *Nat Rev Immunol* 2007;7:365–78.
38. Wernersson S, Pejler G. Mast cell secretory granules: armed for battle. *Nat Rev Immunol* 2014;14:478–94.
39. Boutros PC, Lau SK, Pintilie M, Liu N, Shepherd FA, Der SD, et al. Prognostic gene signatures for non-small-cell lung cancer. *Proc Natl Acad Sci USA* 2009;106:2824–8.
40. Brandsma AM, Hogarth PM, Nimmerjahn F, Leusen JHW. Clarifying the Confusion between Cytokine and Fc Receptor "Common Gamma Chain". *Immunity* 2016;45:225–6.
41. Imada A, Shijubo N, Kojima H, Abe S. Mast cells correlate with angiogenesis and poor outcome in stage I lung adenocarcinoma. *Eur Respir J* 2000;15:1087–93.
42. Takanami I, Takeuchi K, Naruke M. Mast cell density is associated with angiogenesis and poor prognosis in pulmonary adenocarcinoma. *Cancer* 2000;88:2686–92.
43. Ibaraki T, Muramatsu M, Takai S, Jin D, Maruyama H, Orino T, et al. The relationship of tryptase- and chymase-positive mast cells to angiogenesis in stage I non-small cell lung cancer. *Eur J Cardiothorac Surg* 2005;28:617–21.
44. Carlini MJ, Dalurzo MCL, Lastiri JM, Smith DE, Vassallo BC, Puricelli LI, et al. Mast cell phenotypes and microvessels in non-small cell lung cancer and its prognostic significance. *Hum Pathol* 2010;41:697–705.
45. Kurebayashi Y, Emoto K, Hayashi Y, Kamiyama I, Ohtsuka T, Asamura H, et al. Comprehensive immune profiling of lung adenocarcinomas reveals four immunosubtypes with plasma cell subtype a negative indicator. *Cancer Immunol Res* 2016;4:234–47.
46. Welsh TJ, Green RH, Richardson D, Waller DA, O'Byrne KJ, Bradding P. Macrophage and mast-cell invasion of tumor cell islets confers a marked survival advantage in non-small-cell lung cancer. *J Clin Oncol* 2005;23: 8959–67.
47. Shikotra A, Ohri CM, Green RH, Waller DA, Bradding P. Mast cell phenotype, TNFα expression and degranulation status in non-small cell lung cancer. *Sci Rep* 2016;6:38352.
48. Tomita M, Matsuzaki Y, Onitsuka T. Correlation between mast cells and survival rates in patients with pulmonary adenocarcinoma. *Lung Cancer* 1999;26:103–8.

Cancer Immunology Research

Role for High-Affinity IgE Receptor in Prognosis of Lung Adenocarcinoma Patients

Dalam Ly, Chang-Qi Zhu, Michael Cabanero, et al.

Cancer Immunol Res Published OnlineFirst August 3, 2017.

Updated version	Access the most recent version of this article at: doi: 10.1158/2326-6066.CIR-16-0392
Supplementary Material	Access the most recent supplemental material at: http://cancerimmunolres.aacrjournals.org/content/suppl/2017/08/03/2326-6066.CIR-16-0392.DC1

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.