Inflammation-Related Genetic Variations and Survival in Patients With Advanced Non–Small Cell Lung Cancer Receiving First-Line Chemotherapy

X Pu1, MAT Hildebrandt1, C Lu2, JA Roth3, DJ Stewart4, Y Zhao5, RS Heist6, Y Ye1, DW Chang1, L Su5, JD Minna7, SM Lippman8, MR Spitz9, DC Christiani5,6 and X Wu1

Accurate prognostic prediction is challenging for patients with advanced-stage non–small cell lung cancer (NSCLC). We systematically investigated genetic variants within inflammation pathways as potential prognostic markers for advanced-stage NSCLC patients treated with first-line chemotherapy. A discovery phase in 502 patients and an internal validation phase in 335 patients were completed at the MD Anderson Cancer Center. External validation was performed in 371 patients at Harvard University. A missense single-nucleotide polymorphism (SNP) in the gene encoding the major histocompatibility complex class II, DO-β chain (HLA-DOB: rs2071554), predicted to influence protein function, was significantly associated with poor survival in the discovery (hazard ratio (HR): 1.46; 95% confidence interval (CI): 1.02–2.09), internal validation (HR: 1.51; 95% CI: 1.02–2.25), and external validation (HR: 1.52; 95% CI: 1.01–2.29) populations. KLRK1: rs2900420 was associated with reduced risk in the discovery (HR: 0.76; 95% CI: 0.60–0.96), internal validation (HR: 0.77; 95% CI: 0.61–0.99), and external validation (HR: 0.80; 95% CI: 0.63–1.02) populations. A strong cumulative effect on overall survival was observed for these SNPs. Genetic variations in inflammation-related genes could have potential to complement prediction of prognosis.

Lung cancer is a highly lethal disease and was responsible for an estimated 160,000 deaths in 2013 in the United States.1 Patients are typically diagnosed at an advanced stage (stage III/IV), with a dismal 5-year survival rate.2 Combination chemotherapy is the standard of care for stage IV non–small cell lung cancer (NSCLC), whereas combined platinum-based chemoradiation or chemoradiation/surgery is the standard therapy for stage III NSCLC.3,4 Although some patients benefit from standard of care, others do not. Prognostic biomarkers that improve the accuracy of outcome prediction for individual patients could be useful clinically. Inflammation is estimated to contribute to 15% of all cancer-related deaths.5 The lung is a frequent site of inflammation due to environmental exposures, and inflammatory diseases of the lung, such as chronic obstructive pulmonary disease, have been shown to be related to increased incidence of and a poor prognosis for lung cancer.6,7 Evidence has shown that inflammatory molecules and effectors are independently associated with tumor progression and survival in advanced-stage lung cancer patients.8,9 Moreover, chemotherapeutic agents are known to induce cellular damage, which could trigger an acute inflammatory response.10,11 Uncontrolled inflammation can attenuate treatment effectiveness and lead to the development of chemoresistance12 or toxicities, both of which worsen prognosis.13 Taken together, these findings suggest that genetic markers of inflammation might be promising prognostic biomarkers for patients with advanced lung cancer.

The first two authors contributed equally to this work; the last three authors contributed equally to this work.

1Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA; 2Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA; 3Department of Thoracic and Cardiovascular Surgery, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA; 4Division of Medical Oncology, The Ottawa Hospital Cancer Center, Ottawa, Ontario, Canada; 5Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA; 6Department of Medicine, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, USA; 7Hamon Center for Therapeutic Oncology Research, The University of Texas Southwestern Medical Center, Dallas, Texas, USA; 8Rebecca and John Moores Cancer Center, The University of California, San Diego, San Diego, California, USA; 9The Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, USA. Correspondence: X Wu (xwu@mdanderson.org)

Received 24 February 2014; accepted 15 April 2014; advance online publication 4 June 2014. doi:10.1038/clpt.2014.89
Genome-wide association studies (GWASs) have recently been used to detect genetic loci as potential biomarkers of risk and outcomes for various diseases, including lung cancer.14–19 However, some genomic regions have relatively low coverage due to weak linkage disequilibrium relationships and the design of the GWAS arrays. Thus, a comprehensive evaluation of genetic regions of interest based on previous knowledge of the disease biology using pathway-based or gene-based approaches are needed to complement GWAS findings.20 Toward this end, we conducted a multiphase, pathway-based study to evaluate single-nucleotide polymorphisms (SNPs) in major inflammation genes for their effect on overall survival in patients with advanced NSCLC treated with first-line primary chemotherapy (either alone or in combination with radiotherapy), with the goal of identifying potential prognostic biomarkers that will benefit this group of patients.

RESULTS
Patient characteristics
A total of 837 (discovery group: 502, validation group: 335) patients from MD Anderson Cancer Center and 371 patients from Harvard University were included in the analysis (Table 1). MD Anderson populations had a relatively longer median survival time (MST) (discovery: 16.5 months, validation: 16.8 months) as compared with the Harvard University population (12.2 months). The median follow-up time of patients in the MD Anderson discovery phase was relatively short (30.5 months), which is probably due to the higher percentage of stage IV patients. All the patients were non-Hispanic whites with stage III or IV NSCLC, and age was not significantly different between patients who had died and those who were alive.

Association of inflammation-related SNPs with overall survival
A total of 11,930 SNPs from 904 genes were genotyped, of which 11,689 passed quality control measures and were included in the MD Anderson discovery analysis (Figure 1). 1,123 SNPs were significantly associated with overall survival in this group (P < 0.05). Among these SNPs, genotyping data from a previously published GWAS19 were available for 267 SNPs. After removing 413 SNPs that had insignificant (P < 0.2) proxy SNPs (r² > 0.8) on the GWAS chip, we genotyped an additional 443 SNPs using a custom-designed iSelect BeadChip.

After completing quality control, 657 SNPs (390 genotyped and 267 using existing genotype data) were selected for analysis in the internal validation step. We validated the association with overall survival for 49 SNPs (hazard ratios (HRs) consistent, and P < 0.05 for both phases). We then performed an external validation of 32 of the 49 SNPs (those that had data available from previously published GWASs) in the Harvard University population.17 Seventeen SNPs were found to have consistent effects on overall survival in all three populations, with two being significant (or borderline significant) in all three phases (Table 2).

A missense variation in the first exon of HLA-DOB (the gene encoding the major histocompatibility complex class II, DO-β chain), rs2071554, was associated with increased risk of death in all three populations (Figure 2a). In the MD Anderson discovery population (HR: 1.46; 95% confidence interval (CI): 1.02–2.09; P = 0.040), patients carrying at least one variant allele (AG or AA) had a significant survival disparity of 6 months, a decrease from 17 to 11 months, as compared with those who were homozygous for the common allele (GG, P for log-rank test = 0.009, Figure 3a). In the MD Anderson internal validation population, rs2071554 was also associated with increased risk of death (HR: 1.51; 95% CI: 1.02–2.25; P = 0.041) and with a nonsignificant, but appreciable, shortened (by 7 months) MST (Figure 3b). A similar effect was observed in the Harvard University external validation population. The variant allele was associated with shortened overall survival (HR: 1.52; 95% CI: 1.01–2.29; P = 0.045); patients carrying at least one copy of the variant allele had a shorter MST than patients who were homozygous for the common allele (P for log-rank test = 0.007; Figure 3c). Meta-analysis of the association of rs2071554 with overall survival under the fixed-effects model showed a P value of 4.3 × 10^-4 (HR: 1.49; 95% CI: 1.19–1.87; P for heterogeneity = 0.988; Figure 2a).

KLRK1:rs2900420, which is located in the 3′-flanking region of the KLRK1 gene (encoding killer cell lectin-like receptor subfamily K, member 1), a component of the natural killer cell signaling pathway, was associated with reduced risk in the MD Anderson discovery population (HR: 0.76; 95% CI: 0.60–0.96; P = 0.021) and in the
In the analysis of the cumulative effects, we observed a significant “SNP-dosage” effect of these SNPs on overall survival: the more risk genotypes a patient carried, the greater the deleterious effects on overall survival (Figure 2c). As compared with individuals without any unfavorable genotypes (UFGs), patients carrying one UFG had a combined 31% increased risk of death (MD Anderson discovery group: HR: 1.37, 95% CI: 1.07–1.76, \(P = 0.013\); MD Anderson internal validation group: HR: 1.32, 95% CI: 1.03–1.71, \(P = 0.031\); and Harvard University external validation group: HR: 1.25, 95% CI: 0.98–1.60, \(P = 0.073\)). This rose to an 83% increase in risk in the overall population for those with two UFGs (MD Anderson discovery group: HR: 1.83, 95% CI: 1.14–2.94, \(P = 0.012\); MD Anderson internal validation group: HR: 1.96, 95% CI: 1.17–3.30, \(P = 0.011\); and Harvard University external validation group: HR: 1.75, 95% CI: 1.07–2.85, \(P = 0.025\)) and significantly decreased MST (Figure 4).

**In silico function analysis of HLA-DOB: rs2071554**

To determine the potential consequences of this variant and to explore the underlying mechanism, we applied bioinformatics tools to the *in silico* evaluation of the effects on protein structure and function. rs2071554 is a missense variation that results in an arginine to glutamine substitution in the first exon of HLA-DOB. Polyphen-2 analysis suggested that this amino acid change may potentially damage protein function (Polyphen-2: 0.923, sensitivity: 0.80, specificity: 0.94). Similarly, SIFT predicted this SNP to be deleterious (SIFT score: 0.02). Both tools provide additional evidence in support of the potential importance of this SNP on protein function.

### Table 1 Characteristics of the study populations at the time of analysis

| Variable               | MD Anderson discovery population | MD Anderson validation population | Harvard University validation population |
|------------------------|----------------------------------|-----------------------------------|------------------------------------------|
|                       | Deceased (%) | Alive (%) | \(P\) | Deceased (%) | Alive (%) | \(P\) | Deceased (%) | Alive (%) | \(P\) |
| MST (months)           | 16.5          | 16.8      |       | 12.2          |           |       |
| MFT (months)           | 30.5          | 89.6      |       | 60.0          |           |       |
| Age, mean (SD)         | 60.7 (11.2)   | 62.4 (10.5) | 0.099 | 59.3 (10.4)   | 57.5 (9.0) | 0.374 | 63.58 (10.55) | 60.45 (10.76) | 0.053 |
| Sex                    |               |           |       |               |           |       |
| Male                   | 166 (67)      | 80 (33)   | 0.907 | 196 (94)      | 12 (6)    | 0.016 | 171 (89)     | 22 (11)    | 0.098 |
| Female                 | 174 (68)      | 82 (32)   |       | 110 (87)      | 17 (13)   |       | 147 (83)     | 31 (17)    |       |
| Smoking status         |               |           |       |               |           |       |
| Never                  | 129 (76)      | 41 (24)   |       | 4 (100)       | 0 (0)     |       | 25 (76)      | 8 (24)     |       |
| Former                 | 117 (61)      | 74 (39)   |       | 145 (92)      | 13 (8)    |       | 154 (87)     | 23 (13)    |       |
| Current and RQ         | 94 (67)       | 47 (33)   | 0.012 | 157 (91)      | 16 (9)    | 0.782 | 139 (86)     | 22 (14)    | 0.227 |
| Clinical stage         |               |           |       |               |           |       |
| Stage III              | 99 (58)       | 72 (42)   | 0.001 | 164 (95)      | 9 (5)     | 0.020 | 200 (87)     | 31 (13)    | 0.540 |
| Stage IV               | 241 (73)      | 90 (27)   |       |               |           |       |
| Chemotherapy regimens  |               |           |       |               |           |       |
| Platinum-based + other agent\(^a\) | 276 (67) | 136 (33) | 0.449 | 253 (90)      | 27 (10)   | 0.148 | 264 (85)     | 48 (15)    | 0.251 |
| Non–platinum-based     | 64 (71)       | 26 (29)   |       | 53 (96)       | 2 (4)     |       | 35 (90)      | 4 (10)     |       |
| Missing                | 19 (95)       | 1 (5)     |       |               |           |       |
| Total                  | 340           | 162       | 306    | 29            | 318       | 53    |

*Current, current smoker; RQ, recent quitter; MFT, median follow-up time; MST, median survival time.*

*\(^a\)Including taxanes, pemetrexed, gemcitabine, bevacizumab, and erlotinib.*
We systematically evaluated the effects of SNPs from major inflammation genes on overall survival of advanced NSCLC patients who received first-line chemotherapy. In our three-phase pathway-based association study, we identified two potential prognostic biomarkers: HLA-DOB: rs2071554 and KLRK1: rs2900420. The HLA-DOB variant increased risk, with a corresponding decrease in MST, whereas the KLRK1 SNP was protective and prolonged overall survival. Moreover, the HLA-DOB variant was predicted to alter function through in silico analysis, consistent with the observed association of increased risk of death and shortened MST.

HLA-DOB is the \( \beta \) subunit of the HLA-DO class II paralogs. It functions as a negative regulator of major histocompatibility complex class II molecules by inhibiting HLA-DM molecules in a pH-dependent manner. The DO:DM ratio dictates major histocompatibility complex class II–restricted antigen-presentation efficiency.\(^{21}\) Evidence has shown that dysregulation of the antigen presentation pathway is involved in cancer development.\(^{22}\) Moreover, major histocompatibility complex class II molecules are key immune response molecules that have been reported to have a positive relationship with prognosis in various cancers.\(^{23,24}\) In our study, we determined that this missense SNP may alter protein structure and function, and we identified a robust adverse effect on survival across all three populations. Currently, no studies have implicated this gene as playing a role in lung cancer risk or clinical outcomes. Our results suggest a potential predictive role of this locus, making it worthy of future deep sequencing (to identify the causal variant) and functional analysis in vitro (to elucidate the mechanisms responsible).

KLRK1 (coding for member 1 of the killer cell lectin-like receptor subfamily K) encodes a transmembrane protein that interacts with various ligands to activate natural killer and T cells, leading...
Table 2  Seventeen inflammation-related SNPs with consistent effects on overall survival across three analytical phases

| SNP     | Gene       | Discovery MD Anderson population | Internal validation MD Anderson population | External validation Harvard population | Combined validation population | Combined overall population |
|---------|------------|---------------------------------|-------------------------------------------|---------------------------------------|--------------------------------|----------------------------|
|         |            | HR (95% CI)a                     | HR (95% CI)a                               | HR (95% CI)b                          | HR (95% CI)b                    | HR (95% CI)b               |
| rs2071554 | HLA-DOB    | 1.46 (1.02–2.09)                 | 1.51 (1.02–2.25)                           | 1.52 (1.01–2.29)                      | 1.52 (1.14–2.02)              | 1.49 (1.19–1.87)          |
| rs2900420 | KLRK1      | 0.76 (0.60–0.96)                 | 0.77 (0.61–0.99)                           | 0.80 (0.63–1.02)                      | 0.79 (0.66–0.93)               | 0.78 (0.68–0.89)          |
| rs12141265 | FAF1       | 0.75 (0.57–0.97)                 | 0.71 (0.52–0.97)                           | 0.87 (0.66–1.33)                      | 0.80 (0.65–0.98)               | 0.78 (0.66–0.91)          |
| rs1986649 | FOXO1A     | 0.76 (0.60–0.96)                 | 0.75 (0.59–0.95)                           | 0.88 (0.69–1.33)                      | 0.81 (0.68–0.96)               | 0.79 (0.69–0.91)          |
| rs7972757 | KLRK1      | 0.73 (0.55–0.98)                 | 0.67 (0.49–0.92)                           | 0.87 (0.66–1.15)                      | 0.78 (0.63–0.95)               | 0.76 (0.64–0.90)          |
| rs17466614 | FOXO1A     | 0.72 (0.56–0.93)                 | 0.69 (0.53–0.90)                           | 0.89 (0.68–1.16)                      | 0.78 (0.65–0.94)               | 0.76 (0.65–0.88)          |
| rs216136  | CSF1R      | 1.21 (1.03–1.42)                 | 1.17 (1.00–1.37)                           | 1.07 (0.91–1.25)                      | 1.12 (1.00–1.25)               | 1.07 (1.05–1.25)          |
| rs2189521 | IL21R      | 1.41 (1.03–1.94)                 | 1.43 (1.08–1.89)                           | 1.13 (0.85–1.50)                      | 1.27 (1.04–1.55)               | 1.31 (1.10–1.55)          |
| rs1509    | CAPN10     | 0.83 (0.69–0.99)                 | 0.83 (0.68–1.00)                           | 0.93 (0.78–1.11)                      | 0.88 (0.77–1.00)               | 0.86 (0.78–0.96)          |
| rs10964912 | IFNA14     | 1.49 (1.01–2.19)                 | 2.00 (1.26–3.17)                           | 1.16 (0.78–1.72)                      | 1.46 (1.08–1.97)               | 1.47 (1.16–1.86)          |
| rs971768  | IL17RA     | 1.47 (1.09–1.98)                 | 1.46 (1.00–2.12)                           | 1.16 (0.78–1.74)                      | 1.31 (1.00–1.73)               | 1.38 (1.13–1.86)          |
| rs10000856 | IFR2       | 1.26 (1.07–1.50)                 | 1.22 (1.03–1.44)                           | 1.06 (0.90–1.25)                      | 1.13 (1.01–1.28)               | 1.18 (1.07–1.29)          |
| rs2133092 | TNL2       | 1.30 (1.04–1.63)                 | 1.30 (1.03–1.64)                           | 1.08 (0.84–1.38)                      | 1.19 (1.01–1.41)               | 1.23 (1.07–1.41)          |
| rs11903566 | PRKCE      | 1.60 (1.15–2.24)                 | 1.45 (1.04–2.03)                           | 1.11 (0.74–1.67)                      | 1.30 (1.01–1.69)               | 1.41 (1.15–1.73)          |
| rs908742  | PRKcz      | 1.28 (1.03–1.60)                 | 1.33 (1.06–1.67)                           | 1.03 (0.82–1.29)                      | 1.17 (0.99–1.37)               | 1.21 (1.06–1.37)          |
| rs3749166 | CAPN10     | 1.41 (1.04–1.92)                 | 1.42 (1.02–1.99)                           | 1.00 (0.71–1.41)                      | 1.20 (0.94–1.52)               | 1.27 (1.06–1.54)          |

Boldface indicates P < 0.1.

ADD, additive model; CI, confidence interval; DOM, dominant model; HR, hazard ratio; P-het, P for heterogeneity test; REC, recessive model; SNP, single-nucleotide polymorphism.

*aAdjusted for age, sex, smoking status, clinical stage, and treatment regimen.  **Combined (meta-analysis) is based on the fixed-effects model.

to lysis of targeted cells, including tumor cells. This gene has been previously shown to be involved in chemoresistance for osteosarcoma, and ligands binding to KLRK1 have been found to prevent cisplatin-induced cytotoxic lymphocyte killing. Studies have also reported that lung adenocarcinoma cells were able to escape the innate immune response of natural killer cells by expressing heterogeneous ligands for KLRK1. Furthermore, this gene has been identified as a promising target for cancer immunotherapy. However, similar to HLA-DOB, no previous studies have linked KLRK1 to lung cancer risk or clinical outcomes, highlighting the ability of targeted approaches to identify novel predictors. KLRK1:rs2900420 is located 3 kb 3’ to the KLRK1 gene. In our study, it was associated with prolonged overall survival in the MD Anderson populations, and its association with prolonged overall survival was nearly significant in the Harvard University external validation population. It is very likely that with increased sample size, the results would reach statistical significance. An additional KLRK1 variant (rs7972757) was significant in the MD Anderson discovery and internal validation populations, but this finding was not replicated in the Harvard University external validation group, providing additional support to the potential importance of this gene in lung cancer. Further exploration of the potential underlying biological mechanism(s) of this association would increase our understanding of this relationship and solidify the role of KLRK1 in lung cancer prognosis.

To minimize differences in tumor characteristics and treatment regimens between the two study sites (MD Anderson and Harvard University), we followed strict inclusion criteria based...
Articles on cancer stage and treatment. For example, a majority (>80%) of the patients in all three study populations were treated with platinum-based chemotherapy (Table 1), most commonly with the addition of a taxane, although other agents included pemetrexed, gemcitabine, bevacizumab, and erlotinib. However, even with these measures in place there are always subtle, often unidentifiable, differences in the patient populations among different hospitals that could result in differences in survival times such as those we observed between the MD Anderson and Harvard University cohorts. For example, patients of the Harvard University cohort who died were of a slightly older age (63.6 years in the Harvard University cohort, as compared

![Forest plot for meta-analysis of the association of the single-nucleotide polymorphisms (a) HLA-DOB rs2071554 and (b) KLRK1 rs2900420, as well as (c) their cumulative effect, with overall survival in the discovery and internal validation populations from MD Anderson Cancer Center and in the external validation population from Harvard University. CI, confidence interval; HR, hazard ratio; NSCLCs, number of patients with non–small cell lung cancer; UFG, unfavorable genotype.](image)

**Figure 2** Forest plot for meta-analysis of the association of the single-nucleotide polymorphisms (a) HLA-DOB rs2071554 and (b) KLRK1 rs2900420, as well as (c) their cumulative effect, with overall survival in the discovery and internal validation populations from MD Anderson Cancer Center and in the external validation population from Harvard University. CI, confidence interval; HR, hazard ratio; NSCLCs, number of patients with non–small cell lung cancer; UFG, unfavorable genotype.
Articles with 60.7 years in the MD Anderson discovery cohort and 59.3 years in the MD Anderson validation cohort). These slight differences among the populations underscore the potential impact of the two validated SNPs—the effects are stronger than any differences among the study populations, making the findings more transferable across the general population of lung cancer patients and not study-site specific. Several other genetic variants in inflammation genes were significant in the MD Anderson discovery and validation populations but did not reach significance in the Harvard University external validation group. Ten variants did become significant in the validation meta-analysis (Table 2), suggesting that they may indeed be additional predictors of overall survival. These candidate variants are located in several well-known inflammation genes, including those encoding the receptors for several circulating cytokines (CSF1R, IL21R, and IL17RA), cytokines (IRF2 and IFNA14), and cellular signaling molecules (PRKCE and PRKCZ). Further analysis of these genetic variants and genes would be of interest to definitively establish or abolish a relationship with overall survival in advanced lung cancer patients.

To our knowledge, this is the first study to systematically investigate the effects of inflammation-related genetic variations on the survival of patients with advanced NSCLC. The major strength of this study was the three-phase screening and validation approach using two independent patient populations that were drawn from the largest lung cancer pharmacogenetic clinical outcome studies in the United States. All patients were at advanced stages and were treated using first-line chemotherapy with or without radiotherapy. In addition, we developed a comprehensive panel of inflammation-related genetic variations that covered major cellular processes involved in inflammation responses and regulatory processes. With this extensive coverage, our results provide a broad overview of the role of genetic variation within the overall inflammation network in modulating patients’ clinical outcomes.

In conclusion, we identified and validated two potential genetic markers within the inflammation pathway, and these markers may affect overall survival in patients with advanced NSCLC treated with first-line chemotherapy. Considering the important role of inflammation throughout the cancer continuum, these genetic markers may be promising prognostic markers to help in treatment decision making in the clinic.

METHODS

Study populations and data collection

MD Anderson discovery and validation populations. Patients from the University of Texas MD Anderson Cancer Center included in this study are part of an ongoing lung cancer study that has been recruiting since 1995. All patients were non-Hispanic white, had histologically confirmed advanced-stage (stage III or IV, American Joint Committee on Cancer v6.)
NSCLC, did not undergo surgery, and received first-line chemotherapy with or without radiotherapy at MD Anderson. A total of 502 patients were included in the discovery population, with an additional 335 in the validation analysis. A structured questionnaire was used to collect epidemiological and demographic data during an in-person interview. In addition, genomic DNA was extracted from peripheral blood samples using the QIAamp DNA extraction kit (Qiagen, Valencia, CA). Clinical and follow-up data were obtained from medical records. Each patient signed informed consent, and this study was approved by the MD Anderson institutional review board.

**Harvard University external validation.** The details of the Harvard University lung cancer population have been described in detail previously. In brief, participants were non-Hispanic white patients newly diagnosed with histologically confirmed lung cancer. From this population, we selected patients with advanced NSCLC who had received first-line chemotherapy with or without radiation therapy and had not undergone surgery for inclusion in the external validation population. A total of 371 patients met these criteria. An interviewer-administered questionnaire was used to collect epidemiological data. Peripheral blood was drawn for DNA extraction. Informed consent was signed by each study participant, and the Harvard University institutional review board approved this study.

**Genotyping and quality control**

**MD Anderson discovery population.** A custom Illumina iSelect genotyping BeadChip (Illumina, San Diego, CA) was designed to genotype genetic variants in inflammation-related genes (study design detailed in Figure 1). Genes involved in inflammatory responses and regulation were retrieved using the T1Dbase (http://www.t1dbase.org; University of Cambridge), which focuses on diabetes-related and inflammation-related genes. Additional gene information was obtained from the Wkinflam panel. Tagging SNPs for each gene were selected from within a 10-kb flanking region using Utah residents with ancestry from northern and western Europe (CEU) data from the HapMap Project (http://www.hapmap.org), based on the National Center for Biotechnology Information B36 assembly and the Single Nucleotide Polymorphism database b126 using the Tagger Pairwise method ($r^2 > 0.8$ and minor allele frequency ≥ 0.05). Candidate SNPs were then submitted to Illumina and tested for designability using the Assay Design Tool. SNPs with a score >0.6 were considered qualified for the creation of the iSelect BeadChip.

Detailed genotyping and quality control methods used in the discovery phase have been previously described. Briefly, genotyping was performed according to the standard Infinium II assay protocol for the iSelect HD BeadChips. Quality control measures were applied to the data sets, excluding any DNA samples or SNPs with a call rate (percentage of data available for all SNPs or samples) <95%. For patients with direct relatives also enrolled in the study, only one patient within the relationship—the one whose DNA sample had a higher SNP call rate—was included in the final analysis. SNPs with a minor allele frequency <0.01 were excluded.

**MD Anderson internal validation.** Genotyping for SNPs selected for the validation phase was conducted either through the design of a custom iSelect BeadChip or using existing HumanHap300/HumanHap317/HumanHap660 genotyping data. Quality control for the iSelect BeadChip was performed on the basis of sample and SNP call rates; we removed any samples or SNPs with a call rate <95%. Detailed quality control measures for the HumanHap300/HumanHap317/HumanHap660 BeadChip have been described previously; these were also based on genotyping call rate (call rate >95% for all samples and SNPs included). SNPs with a minor allele frequency <0.01 were also excluded.

---

**Kaplan–Meier estimates of unfavorable genotypes (UFGs) and overall survival in advanced NSCLC patients treated with chemotherapy:**

**(a) MD Anderson discovery population; (b) MD Anderson internal validation population; and (c) Harvard University external validation population. N = A/B, A: number of patients who died, B: total number of patients. MST, median survival time; NSCLC, non–small cell lung cancer.**

---

**Figure 4** Kaplan–Meier estimates of unfavorable genotypes (UFGs) and overall survival in advanced NSCLC patients treated with chemotherapy: (a) MD Anderson discovery population; (b) MD Anderson internal validation population; and (c) Harvard University external validation population. N = A/B, A: number of patients who died, B: total number of patients. MST, median survival time; NSCLC, non–small cell lung cancer.

---

**Graphs a, b, and c show Kaplan–Meier survival estimates for patients with advanced NSCLC treated with chemotherapy: (a) MD Anderson discovery population, with unfavorable genotypes (UFG) = 0, 1, or 2; (b) MD Anderson internal validation population, with UFG = 0, 1, or 2; and (c) Harvard University external validation population, with UFG = 0, 1, or 2.**

| UFG | MST (months) | N = A/B |
|-----|-------------|---------|
| 0   | 20          | 91/157  |
| 1   | 15          | 226/315 |
| 2   | 13          | 23/28   |

**Analysis time (months)**

**Survival probability (%)**

| UFG | MST (months) | N = A/B |
|-----|-------------|---------|
| 0   | 18          | 18/20   |
| 1   | 15          | 197/210 |
| 2   | 11          | 19/126  |

**Survival probability (%)**

| UFG | MST (months) | N = A/B |
|-----|-------------|---------|
| 0   | 20          | 91/157  |
| 1   | 15          | 195/223 |
| 2   | 12          | 20/21   |

**Survival probability (%)**

| UFG | MST (months) | N = A/B |
|-----|-------------|---------|
| 0   | 18          | 18/20   |
| 1   | 15          | 197/210 |
| 2   | 12          | 20/21   |

---

**Log-rank P-values:**

- MD Anderson discovery: $P = 0.001$
- MD Anderson internal validation: $P = 0.050$
- Harvard University external validation: $P = 0.078$
Harvard external validation. Genotypes for external validation were obtained from the Illumina HumanHap610-Quad chip following standard protocol, as previously described.18 Quality control measures were similar to those used in the MD Anderson populations: only SNPs and samples with a genotyping call rate >95% and SNPs with a minor allele frequency >0.01 were included in the analysis.

Statistical analyses. For each phase, multivariable Cox proportional-hazards regression models, with corresponding HRs and 95% CIs, were used to estimate the effect of a single SNP on overall survival (the time between diagnosis and death or last follow-up), adjusting for age at diagnosis, sex, smoking status (current, former, or never), clinical stage (stage III or IV), and treatment regimen (chemotherapy and/or radiotherapy). Patients who had smoked fewer than 100 cigarettes over their lifetime were defined as never-smokers; ever-smokers were defined as patients who had smoked ≥100 cigarettes over their lifetime, including former smokers (those who had quit smoking >1 year before diagnosis), current smokers, and recent quitters (those who had quit smoking within a year before diagnosis).

Kaplan–Meier survival curves and corresponding log-rank tests were used to test the survival difference between the genotypes of each SNP. Meta-analysis was performed to obtain summary HRs and 95% CIs. Heterogeneity was tested with χ²-based Q-statistics. A fixed-effects model was used when heterogeneity was absent (P for heterogeneity >0.05).

The cumulative effect of the top two validated SNPs within each population was determined by counting the number of risk genotypes each patient carried and by using patients without any risk genotypes as a reference group. Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/index.shtml)19 and SIFT (http://sift.bii.a-star.edu.sg/)20 were used in silico to predict the influence of the SNP on protein function.

The potential effect of population stratification was evaluated using quantile–quantile plots of the test statistics in the MD Anderson discovery population. We calculated the inflation factor (λ) by dividing the observed median of test statistics by expected median (from χ² distribution with 1 degree of freedom) value. The obtained λ was close to 1 (0.92), indicating that population substructure has no substantial effect on the test statistics in the discovery stage analysis.

ACKNOWLEDGMENTS

This work was supported by National Cancer Institute grants R01 CA111646, P50 CA070907, R01 CA127219, R01 CA074386, R01 CA092824, and P50 CA090578, as well as by the National Institutes of Health through MD Anderson's Cancer Center Support grant CA016672.

AUTHOR CONTRIBUTIONS

X.P., M.A.T.H., C.L., J.A.R., D.J.S., Y.Z., R.S.H., Y.Y., D.W.C., L.S., J.D.M., S.M.L., M.R.S., D.C.C., and X.W. wrote the manuscript. X.P., M.A.T.H., Y.Y., D.W.C., M.R.S., D.C.C., and X.W. designed the research. X.P., M.A.T.H., C.L., J.A.R., D.J.S., Y.Z., R.S.H., Y.Y., D.W.C., L.S., J.D.M., S.M.L., M.R.S., D.C.C., and X.W. performed the research. X.P., M.A.T.H., Y.Z., Y.Y., D.C.C., and X.W. analyzed the data.

CONFLICT OF INTEREST

The authors declared no conflict of interest.

© 2014 American Society for Clinical Pharmacology and Therapeutics

1. Siegel, R., Naishadham, D. & Jemal, A. Cancer statistics, 2013. CA Cancer J. Clin. 63, 11–30 (2013).
2. Jemal, A., Siegel, R., Xu, J. & Ward, E. Cancer statistics, 2010. CA Cancer J. Clin. 60, 277–300 (2010).
3. Ardizzoni, A. et al.; CISCA (CISplatin versus Carboplatin) Meta-analysis Group. CISplatin–versus carboplatin-based chemotherapy in first-line treatment of advanced non-small-cell lung cancer: an individual patient data meta-analysis. J. Natl. Cancer Inst. 99, 847–857 (2007).
4. Ung, Y.C., Yu, E., Falkson, C., Haynes, A.E., Stys-Norman, D. & Evans, W.K.; Lung Cancer Disease Site Group Of Cancer Care Ontario’s Program In Evidence-Based Care. The role of high-dose-rate brachytherapy in the palliation of symptoms in patients with non-small-cell lung cancer: a systematic review. Brachytherapy 5, 189–202 (2006).
5. Ramos-Nino, M.E. The role of chronic inflammation in obesity-associated cancers. ISRN Oncol. 2013, 697521 (2013).
6. Zaynagetdinov, R. et al. Epithelial nuclear factor-κB signaling promotes lung carcinogenesis via recruitment of regulatory T lymphocytes. Oncogene 31, 3164–3176 (2012).
7. Vlachostergios, P.J. et al. Baseline insulin-like growth factor-I plasma levels, systemic inflammation, weight loss and clinical outcome in metastatic non-small cell lung cancer patients. Oncology 81, 113–118 (2011).
8. Ulrich, C.M., Bigler, J. & Potter, J.D. Non-steroidal anti-inflammatory drugs for cancer prevention: promise, perils and pharmacogenetics. Nat. Rev. Cancer 6, 130–140 (2006).
9. Kasymjanova, G. et al. The predictive value of pre-treatment inflammatory markers in advanced non-small-cell lung cancer. Curr. Oncol. 17, 52–58 (2010).
10. Mukhopadhyay, P. et al. Poly(ADP-ribose) polymerase-1 is a key mediator of cisplatin-induced kidney inflammation and injury. Free Radic. Biol. Med. 51, 1774–1788 (2011).
11. Nakano, S. & Niho, Y.S-Fluorouracil and cisplatin sequential chemotherapy; mechanism of action and clinical application in unresectable non-small cell lung cancer. Intern. Med. 34, 288–291 (1995).
12. Mahon, K.L., Henshall, S.M., Sutherland, R.L. & Horvath, L.G. Pathways of chemotherapy resistance in castration-resistant prostate cancer. Endocr. Relat. Cancer 18, R103–R123 (2011).
13. Yeh, I., George, E. & Fleckman, P. Efficine hirudinidin sin neutrophilis: a toxic response to chemotherapy. J. Cutan. Pathol. 38, 905–910 (2011).
14. Amos, C.I. et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25. 1 Nat. Genet. 40, 616–622 (2008).
15. Hung, R.J. et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 452, 633–637 (2008).
16. Thorgersson, T.E. et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 452, 638–642 (2008).
17. Huang, Y.T. et al. Genome-wide analysis of survival in early-stage non-small-cell lung cancer. J. Clin. Oncol. 27, 2660–2667 (2009).
18. Li, Y. et al. Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. Lancet Oncol. 11, 321–330 (2010).
19. Wu, X. et al. Genome-wide association study of survival in non-small cell lung cancer patients receiving platinum-based chemotherapy. J. Natl. Cancer Inst. 103, 817–825 (2011).
20. Li, M., Li, C. & Guan, W. Evaluation of coverage variation of SNP chips for genome-wide association studies. Eur. J. Hum. Genet. 16, 635–643 (2008).
21. Denzin, L.K., Sant’Angelo, D.B., Hammond, C., Surman, M.J. & Cresswell, P. Negative regulation by HLA-DO of MHC class II-restricted antigen processing. Science 278, 106–109 (1997).
22. Seliger, B. Different regulation of MHC class I antigen processing components in human tumors. J. Immunotoxicol. 5, 361–367 (2008).
23. Chamuleau, M.E., Ossenkoppele, G.J. & van de Loosdrecht, A.A. MHC class II negative regulation by HLA-DO of MHC class II-restricted antigen processing. Eur. J. Immunol. 38, 527–534 (2008).
24. Buddingh, E.P. et al. Chemotherapy-resistant osteosarcoma is highly susceptible to IL-1S-activated allogeneic and autologous NK cells. Cancer Immunol. Immunother. 60, 575–586 (2011).
25. Conejo-Garcia, J.R. et al. Ovarian carcinoma expresses the NKG2D ligand Letal and promotes the survival and expansion of CD28-antitumor T cells. Cancer Res. 64, 2175–2182 (2004).
26. Le Maux Chansac, B. et al. Potentiation of NK cell-mediated cytotoxicity in human lung adenocarcinoma: role of NKG2D-dependent pathway. Int. Immunol. 20, 801–810 (2008).
27. Zhang, T. & Sentman, C.L. Cancer immunotherapy using a bispecific NK receptor fusion protein that engages both T cells and tumor cells. Cancer Res. 71, 2066–2076 (2011).
28. Park, M.J., Bae, J.H., Chung, J.S., Kim, S.H. & Kang, C.D. Induction of NKG2D ligands and increased sensitivity of tumor cells to NK cell-mediated cytotoxicity by hematopoietin-based photodynamic therapy. Immunol. Invest. 40, 367–382 (2011).
30. Asomaning, K. et al. Second hand smoke, age of exposure and lung cancer risk. Lung Cancer 61, 13–20 (2008).
31. Loza, M.J., McCall, C.E., Li, L., Isaacs, W.B., Xu, J. & Chang, B.L. Assembly of inflammation-related genes for pathway-focused genetic analysis. PLoS ONE 2, e1035 (2007).
32. de Bakker, P.I., Yelensky, R., Pe’er, I., Gabriel, S.B., Daly, M.J. & Altshuler, D. Efficiency and power in genetic association studies. Nat. Genet. 37, 1217–1223 (2005).
33. Spitz, M.R. et al. Variants in inflammation genes are implicated in risk of lung cancer in never smokers exposed to second-hand smoke. Cancer Discov. 1, 420–429 (2011).
34. Adzhubei, I.A. et al. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249 (2010).
35. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4, 1073–1081 (2009).