



Novel Variants of *ELP2* and *PIAS1* in the Interferon Gamma Signaling Pathway Are Associated with Non-Small Cell Lung Cancer Survival

Yu Chen Zhao^{1,2}, Dongfang Tang^{1,2}, Sen Yang^{1,2}, Hongliang Liu^{1,2}, Sheng Luo³, Thomas E. Stinchcombe^{1,4}, Carolyn Glass^{1,5}, Li Su⁶, Sipeng Shen⁶, David C. Christiani^{6,7}, and Qingyi Wei^{1,2,4}

ABSTRACT

Background: IFN γ is a pleiotropic cytokine that plays critical immunomodulatory roles in intercellular communication in innate and adaptive immune responses. Despite recognition of IFN γ signaling effects on host defense against viral infection and its utility in immunotherapy and tumor progression, the roles of genetic variants of the IFN γ signaling pathway genes in survival of patients with cancer remain unknown.

Methods: We used a discovery genotyping dataset from the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial ($n = 1,185$) and a replication genotyping dataset from the Harvard Lung Cancer Susceptibility Study ($n = 984$) to evaluate associations between 14,553 genetic variants in 150 IFN γ pathway genes and survival of non-small cell lung cancer (NSCLC).

Results: The combined analysis identified two independent potentially functional SNPs, *ELP2* rs7242481G>A and *PIAS1* rs1049493T>C, to be significantly associated with NSCLC sur-

vival, with a combined HR of 0.85 (95% confidence interval, 0.78–0.92; $P < 0.0001$) and 0.87 (0.81–0.93; $P < 0.0001$), respectively. Expression quantitative trait loci analyses showed that the survival-associated *ELP2* rs7242481A allele was significantly associated with increased mRNA expression levels of elongator acetyltransferase complex subunit 2 (*ELP2*) in 373 lymphoblastoid cell lines and 369 whole-blood samples. The *PIAS1* rs1049493C allele was significantly associated with decreased mRNA expression levels of *PIAS1* in 383 normal lung tissues and 369 whole-blood samples.

Conclusions: Genetic variants of IFN γ signaling genes are potential prognostic markers for NSCLC survival, likely through modulating the expression of key genes involved in host immune response.

Impact: Once validated, these variants could be useful predictors of NSCLC survival.

Introduction

Lung cancer is one of the most common malignancies both in the United States and worldwide, with 228,150 new cases of lung cancer and approximately 142,670 deaths from this disease in the United States in 2019 (1). Non-small cell lung cancer (NSCLC), mostly adenocarcinoma and squamous cell carcinoma, are the most common histologic subtypes, accounting for around 85% of all patients with lung cancer (2). Despite devoted efforts in the treatment over the past decades, lung cancer remains the cause for the highest cancer-related mortality worldwide, with an underwhelming 5-year survival rate of 18.6% between 2008 and 2014 in

the United States (3). Conventional treatments for NSCLC include surgery, chemotherapy, and radiotherapy for its early stages, but the responses to these treatments are heterogeneous (4), likely due to genetic variation among the patients, such as single-nucleotide polymorphisms (SNP), the most common genetic variants that could affect both short-term response and long-term prognosis of patients with cancer; thus, identifying the role of these genetic factors for NSCLC survival may lead to a better understanding of the variability in treatment outcomes (5).

Recently, utilization of the immune system to halt cancer development and tumor progression has been widely recognized (6, 7) and immunotherapy is now considered the “fourth pillar” alongside the three conventional treatments (8). Theoretically, immunotherapy either assists the ability of the immune system to target cancer cells directly or stimulate the immune system in a more general matter (9). In NSCLC treatment, programmed death-ligand 1 (PD-L1) inhibitors, such as pembrolizumab, are often applied to patients with a high PD-L1 expression in tumors, in combination with chemotherapy to improve therapeutic results (10). Despite recent advances, not all patients respond to immunotherapy (11). Therefore, it is important to identify survival-related biomarkers for immunotherapy.

To date, genome-wide association studies (GWAS) investigating millions of SNPs at the same time have identified few SNPs that are associated with cancer survival, because a hypothesis-free GWAS focuses on the most significant SNPs or genes with a significant P value after the stringent multiple testing correction. In reported GWASs on cancer survival, most identified SNPs lack functional annotations, which limits clinical application of these results (12, 13). As a promising hypothesis-driven method in the post-GWAS era, a biological pathway-based approach has been applied to reanalyze

¹Duke Cancer Institute, Duke University Medical Center, Durham, North Carolina.

²Department of Population Health Sciences, Duke University School of Medicine, Durham, North Carolina. ³Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, North Carolina. ⁴Department of Medicine, Duke University Medical Center, Durham, North Carolina. ⁵Department of Pathology, Duke University School of Medicine, Durham, North Carolina. ⁶Department of Environmental Health and Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts. ⁷Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

Corresponding Author: Qingyi Wei, Duke University Medical Center, Department of Population Health Sciences and Department of Medicine, Duke University School of Medicine, 905 S. LaSalle Street, Durham, NC 27710. Phone: 919-660-0562; Fax: 919-660-0178; E-mail: qingyi.wei@duke.edu

Cancer Epidemiol Biomarkers Prev 2020;XX:XX-XX

doi: 10.1158/1055-9965.EPI-19-1450

©2020 American Association for Cancer Research.

published GWAS datasets and to evaluate the cumulative effect of SNPs across multiple genes in the same biological pathway (14). Because much fewer SNPs in the candidate genes of a particular biological pathway of interest will be included in the analysis, it avoids unnecessary multiple tests for SNPs that may have no apparent biological significance, which improves the overall study power to identify statistically significant and biologically important associations (14).

IFNs are a group of pleiotropic cytokines that play immunomodulatory roles during intercellular communication in innate and adaptive immune responses (15). IFNs are divided into three types, of which the type II IFN (i.e., IFN γ) in humans is the most extensively studied (15). IFN γ is a signaling protein released by cytotoxic T cells and type I Th cells in response to inflammatory or immune stimuli (16). IFN γ primarily modulates the activation of IFN response factor 1 through the JAK–STAT pathway, leading to the activation of a group of secondary responsive genes that upregulate pathogen recognition, antigen presentation, and inhibit cellular proliferation (17, 18). Despite IFN γ signaling being crucial for activating the immune system, it remains to be determined whether IFN γ has a role in assisting tumor immune evasion, especially in the context of clinical trials (19). Furthermore, the roles of genetic variants of candidate IFN γ signaling genes and their biological functions in tumor growth or suppression remain unknown.

In the present study, therefore, we hypothesize that genetic variants in genes related to the IFN γ signaling pathway, a critical cascade in the activation of both innate and adaptive immune responses, are associated with survival of patients with NSCLC. We used the existing genotyping data from two previously published GWASs to test this hypothesis.

Materials and Methods

Study populations

We utilized a GWAS dataset of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial as the discovery; PLCO is a multicenter dataset of a randomized controlled study conducted between 1993 and 2011 (20). In addition to data on genotyping and survival, the PLCO dataset of 1,185 patients with NSCLC also included blood samples and personal information about smoking status, histologic diagnosis, tumor stage, treatment method, and family history collected at enrollment (21). Blood DNA samples of the participants were genotyped with Illumina HumanHap240Sv1.0 and HumanHap550v3.0 [database of Genotypes and Phenotypes (dbGaP) accession numbers: phs000093.v2.p2 and phs000336.v1.p1; refs. 22, 23]. Each institutional review board of the participating institutions had approved the PLCO Trial with a written informed consent from each of the participants permitting use of the collected data.

For replication, we used genotyping data extracted from another GWAS dataset of 984 Caucasian patients with histologically confirmed NSCLC from the Harvard Lung Cancer Susceptibility (HLCS) Study, which began in 1991 (24), where whole-blood samples and personal information were collected after diagnosis, and DNA extracted from the blood samples were genotyped using the Illumina Humanhap610-Quad array. The genotyped data were utilized for imputation with the MACH3 software based on sequencing data from the 1000 Genomes Project (24).

The use of these two GWAS datasets was approved by the Internal Review Board of Duke University School of Medicine

(Durham, NC, #Pro00054575) and the dbGaP database administration for the PLCO dataset (#6404 with dbGaP accession numbers: phs000093.v2.p2 and phs000336.v1.p1).

Gene and SNP selection

The genes involved in the IFN γ signaling pathway were selected using the molecular signatures database (<http://software.broadinstitute.org/gsea/msigdb/index.jsp>) with the keyword “interferon and gamma.” With the removal of 133 duplicated genes, two pseudogenes, one withdrawn gene, and one gene located on X chromosome, 150 genes remained as candidates for further analysis (Supplementary Table S1). Imputation with IMPUTE2 and the 1000 Genomes Project data (phase III) was performed for SNPs within ± 500 kb flanking regions of these candidate genes. SNPs within the genes and their ± 2 kb flanking regions were then extracted according to the following criteria: an imputation info score ≥ 0.8 (Supplementary Fig. S1), a genotyping rate $\geq 95\%$, a minor allelic frequency (MAF) $\geq 5\%$, and a Hardy–Weinberg equilibrium (HWE) $\geq 1 \times 10^{-5}$, which resulted in 14,553 (1,053 genotyped and 13,500 imputed) SNPs for further analyses.

Statistical analyses

The endpoints for analysis included overall survival (OS) and disease-specific survival (DSS). In the single-locus analysis, we used multivariate Cox proportional hazards regression analysis to analyze the association between each of these SNPs and NSCLC survival in an additive model using the PLCO dataset, with adjustment for various covariates including age, sex, smoking status, histology, tumor stage, chemotherapy, radiotherapy, surgery, and the first four principal components (Supplementary Table S2) by using the GenABEL package of R software (25). We then used Bayesian false discovery probability (BFDP) with a cut-off value of 0.80 for multiple testing correction to decrease the probability of potentially false-positive results as recommended for SNPs in high linkage disequilibrium (LD) as a result of imputation (26, 27). We assigned a prior probability of 0.10 and detected an upper boundary hazards ratio (HR) of 3.0 for an association with variant genotypes or minor alleles of the SNPs with $P < 0.05$. The chosen SNPs were then validated afterwards by using the HLCS genotyping dataset. Next, we performed an inverse variance weighted meta-analysis to combine the results of both discovery and replication datasets, followed by a multivariate stepwise Cox model to identify independent SNPs through adjustments for covariates, the top four principal components, demographic characteristics, as well as 15 previously published SNPs (28). We used Manhattan plots and regional association plots to visualize the locations and LD of the selected SNPs.

We then used the combined genotypes or alleles of the identified SNPs to evaluate their cumulative effects and the Kaplan–Meier (KM) survival curves to show their associations with survival probability, constructed the receiver operating characteristic (ROC) curve and time-dependent area under the curve (AUC) to illustrate prediction accuracy for NSCLC survival (29), and performed the expression quantitative trait loci (eQTL) analyses with the genomic data from the 1000 Genomes Project and the Genotype–Tissue Expression (GTEx) project (30, 31). The bioinformatics functional prediction for the tagging SNPs was then performed with SNPinfo (ref. 32; <https://snpinfo.niehs.nih.gov>), RegulomeDB (ref. 33; <http://www.regulomedb.org>), and HaploReg (ref. 34; <http://archive.broadinstitute.org/mammals/haploreg/haploreg.php>). Finally, the differences in mRNA expression levels were examined in 111 pairs of lung cancer tissues and

adjacent normal tissues from The Cancer Genome Atlas (TCGA) database through a paired Student *t*-test. We also assessed the differences in mRNA expression levels in a larger, but not paired, dataset from TCGA (<http://ualcan.path.uab.edu>), and the KM survival analysis was performed to evaluate the association between the mRNA expression levels and survival probability (<http://kmplot.com/analysis/index.php?p=service&cancer=lung>). All statistical analyses were performed with a statistical significance level of $P < 0.05$ by using the SAS Software (version 9.4; SAS Institute), unless otherwise indicated.

Results

Associations between SNPs in the IFN γ signaling pathway genes and NSCLC survival

The overall flowchart of this study is shown in **Fig. 1**. Basic characteristics of 1,185 patients with NSCLC from the PLCO Trial and 984 patients with NSCLC from the HLCS Study have been described in Supplementary Table S3 (28). After multiple testing correction by BFDP ≤ 0.80 , we identified 340 SNPs that were significantly associated with NSCLC OS ($P < 0.05$), of which 48 SNPs in two genes remained significant after further replication by the HLCS genotyping dataset. Subsequently, we performed a combined analysis of the PLCO and HLCS datasets for these newly identified SNPs. As shown in **Table 1**, three representative SNPs were determined after considering the LD among the SNPs, one SNP in elongator acetyltransferase complex subunit 2 (*ELP2*) and two SNPs in *PIAS1*, were found to be associated with a better survival, without heterogeneity observed between the two datasets.

Other 45 SNPs in high LD (>0.80) with these three SNPs in the same genes are presented in Supplementary Table S4; these 45 SNPs also feature the same directionality in terms of survival (Supplementary Table S5).

Independent SNPs associated with NSCLC survival in the PLCO dataset

When the three validated SNPs were included in the multivariate stepwise Cox model for the PLCO dataset only (because the HLCS Study dataset did not have individual genotyping data), two SNPs remained independently and significantly associated with survival, even after adjustment for additional 15 previously reported SNPs significantly associated with NSCLC survival from the same PLCO GWAS dataset (**Table 2**). The results of selected SNPs from PLCO and HLCS are summarized in two separate Manhattan plots (Supplementary Fig. S2A and S2B), respectively, and the regional association plot (<http://locuszoom.org/>) of each SNP is shown in Supplementary Fig. S3 (35).

In the PLCO dataset with complete adjustment for available covariates, patients with the protective *ELP2* rs7242481 A (i.e., GA+AA) allele or *PIAS1* rs1049493 C (i.e., TC+CC) allele had a better OS and DSS ($P_{\text{trend}} = 0.004$ and $P_{\text{trend}} = 0.009$ for *ELP2* rs7242481 A, respectively; $P_{\text{trend}} = 0.006$ and $P_{\text{trend}} = 0.023$ *PIAS1* rs1049493 C, respectively; **Table 3**). In comparison with the GG risk genotype, the *ELP2* rs7242481 GA genotype was associated with a decreased risk of death [HR, 0.85; 95% confidence interval (CI), 0.73–0.99; $P = 0.033$ for OS and HR, 0.84; 95% CI, 0.72–0.99; $P = 0.033$ for DSS], and the *ELP2* rs7242481 AA genotype was also associated with a decreased risk of

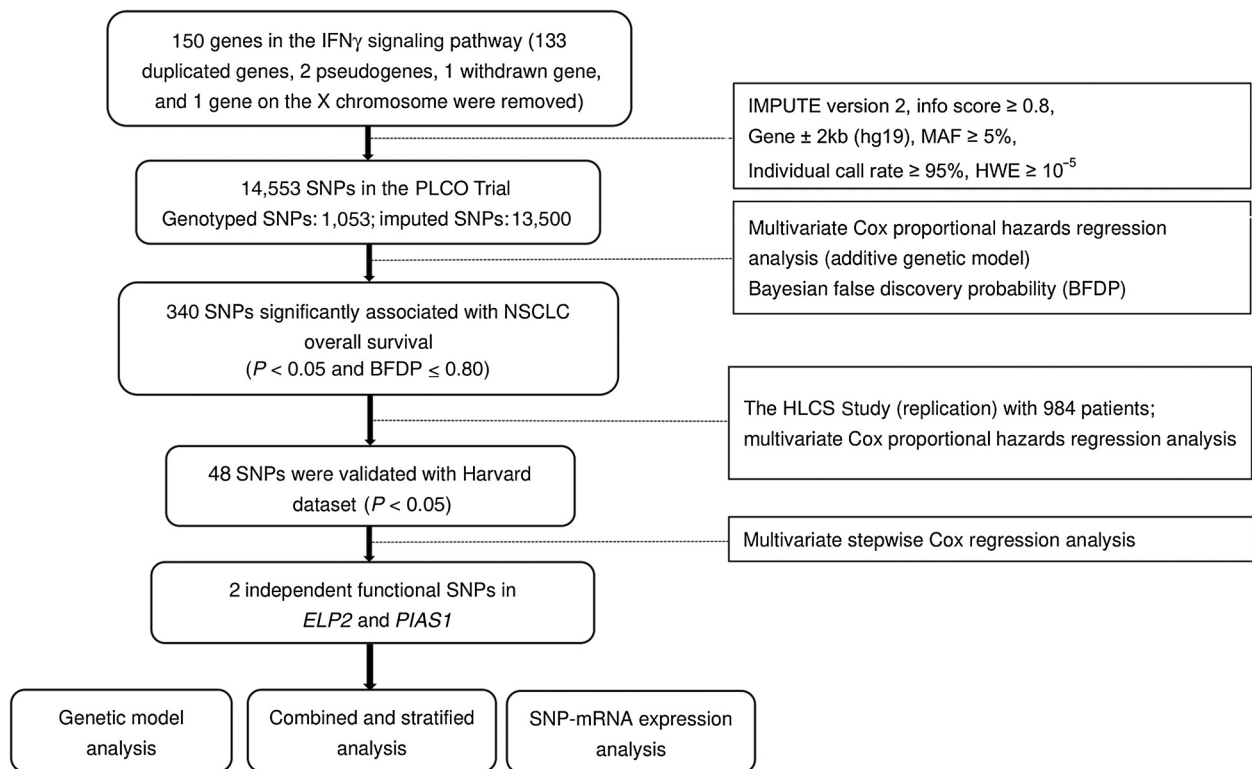


Figure 1.

Study flowchart. The overall procedures of this study. HLCS, Harvard Lung Cancer Susceptibility Study; HWE, Hardy–Weinberg equilibrium; MAF, minor allelic frequency; NSCLC, non-small cell lung cancer; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; SNP, single-nucleotide polymorphism.

Table 1. Associations of three significant SNPs with OS of patients with NSCLC in both discovery and validation datasets from two previously published GWASs.

SNPs ^a	Allele ^b	Gene	PLCO (n = 1,185)					HLCS (n = 984)			Combined analysis			
			FDR	BFDP	EAF	HR (95% CI) ^c	P ^c	EAF	HR (95% CI) ^d	P ^d	P _{het} ^e	I ²	HR (95% CI) ^f	P ^f
rs7242481	G>A	<i>ELP2</i>	0.50	0.56	0.35	0.86 (0.77–0.95)	0.004	0.36	0.83 (0.74–0.93)	0.002	0.658	0	0.85 (0.78–0.92)	3.02 × 10 ⁻⁵
rs11071978	T>A	<i>PIAS1</i>	0.50	0.72	0.36	0.86 (0.78–0.96)	0.006	0.34	0.82 (0.72–0.92)	0.001	0.529	0	0.84 (0.77–0.91)	2.34 × 10 ⁻⁵
rs1049493	T>C	<i>PIAS1</i>	0.50	0.69	0.45	0.87 (0.79–0.96)	0.006	0.45	0.87 (0.78–0.96)	0.009	0.989	0	0.87 (0.81–0.93)	1.24 × 10 ⁻⁴

Abbreviations: BFDP, Bayesian false discovery probability; CI, confidence interval; EAF, effect allele frequency; FDR, false discovery rate; GWAS, genome-wide association study; HLCS, Harvard Lung Cancer Susceptibility Study; HR, hazards ratio; NSCLC, non-small cell lung cancer; OS, overall survival; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; SNP, single-nucleotide polymorphism.

^aThe other 45 SNPs in high linkage disequilibrium ($r^2 > 0.8$) with these three SNPs are presented in Supplementary Table S4 and Supplementary Table S5.

^bEffect/reference allele.

^cAdjusted for age, sex, stage, histology, smoking status, chemotherapy, radiotherapy, surgery, PC1, PC2, PC3, and PC4 (PC = principal component).

^dAdjusted for age, sex, stage, histology, smoking status, chemotherapy, radiotherapy, surgery, PC1, PC2, and PC3.

^eP_{het}: P value for heterogeneity by Cochran Q test.

^fMeta-analysis in the fix-effects model.

death (HR, 0.74; 95% CI, 0.59–0.94; $P = 0.012$ for OS and HR, 0.76; 95% CI, 0.59–0.97; $P = 0.026$ for DSS). Similarly, in comparison with the TT risk genotype, the *PIAS1* rs1049493 TC genotype was associated with a nonsignificant better OS and DSS (HR, 0.92; 95% CI, 0.78–1.0; $P = 0.303$ and HR, 0.91; 95% CI, 0.76–1.08; $P = 0.285$, respectively), and the *PIAS1* rs1049493 CC genotype was associated with a significantly better OS and DSS (HR, 0.75; 95% CI, 0.61–0.91; $P = 0.005$ and HR, 0.78; 95% CI, 0.63–0.97; $P = 0.022$; respectively; **Table 3**).

Combined effects of the two independent SNPs in the PLCO dataset

We first combined the significant protective genotypes (i.e., *ELP2* rs7242481 TA+AA and *PIAS1* rs1049493 TC+CC) into a genetic

score as the number of protective genotypes (NPG). As shown in **Table 3**, the increased genetic score of the NPGs was associated with a better survival in the multivariate analysis in the PLCO dataset ($P_{\text{trend}} < 0.0004$ for OS and $P_{\text{trend}} = 0.002$ for DSS). When we dichotomized all the patients into groups with 0–1 and 2 NPGs, the 2-score group had a significantly better survival (HR, 0.64; 95% CI, 0.52–0.80; $P < 0.0001$ for OS and HR, 0.66; 95% CI, 0.53–0.83; $P = 0.0004$ for DSS), in comparison with the 0–1 score group. As shown in **Table 3**, the increased genetic score of the number of protective alleles (NPA) was associated with a better survival in the multivariate analysis in the PLCO dataset ($P_{\text{trend}} < 0.0001$ for OS and $P_{\text{trend}} = 0.0006$ for DSS). When we dichotomized all the patients into groups with 0–1 and 2–4 NPAs, the 2–4 score group had a significantly better survival

Table 2. Two independent SNPs in a multivariate Cox proportional hazards regression analysis with adjustment for other covariates and 15 previously published SNPs for NSCLC in the PLCO Trial.

Variables	Category	Frequency	HR (95% CI) ^a	P ^a	HR (95% CI) ^b	P ^b
Age	Continuous	1,185	1.03 (1.02–1.05)	<0.0001	1.04 (1.02–1.05)	<0.0001
Sex	Male	698	1.00		1.00	
	Female	487	0.80 (0.69–0.93)	0.004	0.78 (0.67–0.91)	0.002
Smoking status	Never	115	1.00		1.00	
	Current	423	1.67 (1.26–2.26)	0.0004	1.94 (1.44–2.62)	<0.0001
	Former	647	1.64 (1.25–2.16)	0.0004	1.90 (1.42–2.52)	<0.0001
Histology	Adenocarcinoma	577	1.00		1.00	
	Squamous cell	285	1.20 (0.99–1.45)	0.057	1.25 (1.03–1.51)	0.025
	Others	323	1.32 (1.11–1.56)	0.002	1.37 (1.15–1.63)	0.0006
Tumor stage	I–IIIA	655	1.00		1.00	
	IIIB–IV	528	2.94 (2.42–3.58)	<0.0001	3.11 (2.55–3.79)	<0.0001
Chemotherapy	No	639	1.00		1.00	
	Yes	538	0.58 (0.48–0.69)	<0.0001	0.58 (0.48–0.70)	<0.0001
Radiotherapy	No	762	1.00		1.00	
	Yes	415	0.95 (0.81–1.12)	0.526	0.94 (0.80–1.12)	0.497
Surgery	No	637	1.00		1.00	
	Yes	540	0.21 (0.17–0.28)	<0.0001	0.20 (0.15–0.26)	<0.0001
<i>ELP2</i> rs7242481 G>A	GG/GA/AA	495/544/146	0.86 (0.78–0.96)	0.007	0.86 (0.77–0.96)	0.007
<i>PIAS1</i> rs1049493 T>C	TT/TC/CC	368/579/248	0.87 (0.79–0.97)	0.009	0.89 (0.80–0.99)	0.024

Abbreviations: CI, confidence interval; HR, hazards ratio; NSCLC, non-small cell lung cancer; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; SNP, single-nucleotide polymorphism.

^aAdjusted for age, sex, tumor stage, histology, smoking status, chemotherapy, radiotherapy, surgery, and PC1, PC2, PC3, and PC4 (PC = principal component).

^bAnother 15 published SNPs were included for further adjustment: five SNPs in PMID: 27557513; one SNP in PMID: 29978465; two SNPs in PMID: 30259978; two SNPs in PMID: 26757251; three SNPs in PMID: 30650190; and two SNPs in PMID: 30989732.

Table 3. Associations between the NPAs of two independent SNPs with OS and DSS of NSCLC in the PLCO Trial.

Genotypes/alleles	Frequency ^a	OS ^b			DSS ^b		
		Death (%)	HR (95% CI)	P	Death (%)	HR (95% CI)	P
<i>ELP2</i> rs7242481 G>A							
GG	491	336 (68.43)	1.00		306 (62.32)	1.00	
GA	538	361 (67.10)	0.85 (0.73–0.99)	0.033	318 (59.11)	0.84 (0.72–0.99)	0.033
AA	146	92 (63.01)	0.74 (0.59–0.94)	0.012	85 (58.22)	0.76 (0.59–0.97)	0.026
Trend test				0.004			0.009
Dominant							
GG	491	336 (68.43)	1.00		306 (62.32)	1.00	
GA+AA	684	453 (66.23)	0.82 (0.71–0.95)	0.008	403 (58.92)	0.82 (0.81–0.96)	0.011
<i>PIAS1</i> rs1049493 T>C							
TT	356	250 (70.22)	1.00		225 (63.20)	1.00	
TC	571	377 (66.02)	0.92 (0.78–1.08)	0.303	334 (58.49)	0.91 (0.76–1.08)	0.285
CC	248	162 (65.32)	0.75 (0.61–0.91)	0.005	150 (60.48)	0.78 (0.63–0.97)	0.022
Trend test				0.006			0.023
Dominant							
TT	356	250 (70.22)	1.00		225 (63.20)	1.00	
TC+CC	819	539 (65.81)	0.86 (0.74–1.00)	0.053	484 (59.10)	0.87 (0.74–1.02)	0.082
NPG ^c							
0	398	274 (68.84)	1.00		249 (62.06)	1.00	
1	622	415 (66.72)	0.94 (0.81–1.10)	0.424	371 (59.65)	0.95 (0.81–1.12)	0.566
2	155	100 (64.52)	0.62 (0.49–0.79)	<0.0001	91 (58.71)	0.64 (0.50–0.83)	0.001
Trend test				0.0004			0.002
Dichotomized NPG							
0–1	1,020	689 (67.6)	1.00		618 (60.59)	1.00	
2	155	100 (64.5)	0.64 (0.52–0.80)	<0.0001	91 (58.71)	0.66 (0.53–0.83)	0.0004
NPA ^d							
0	140	100 (71.43)	1.00		88 (62.86)	1.00	
1	431	295 (68.45)	1.03 (0.82–1.30)	0.796	269 (62.41)	1.08 (0.84–1.37)	0.561
2	379	248 (65.44)	0.85 (0.68–1.08)	0.189	222 (58.58)	0.89 (0.69–1.14)	0.355
3	192	120 (62.50)	0.74 (0.56–0.97)	0.028	110 (57.29)	0.79 (0.59–1.05)	0.103
4	33	21 (63.64)	0.51 (0.31–0.82)	0.005	19 (57.58)	0.53 (0.32–0.81)	0.013
Trend test				<0.0001			0.0006
Dichotomized NPA							
0–1	571	395 (69.18)	1.00		357 (62.52)	1.00	
2–4	604	389 (64.40)	0.77 (0.67–0.89)	0.0004	351 (58.11)	0.79 (0.68–0.91)	0.002

Abbreviations: CI, confidence interval; DSS, disease-specific survival; HR, hazards ratio; NPA, number of protective alleles; NPG, number of protective genotypes; NSCLC, non-small cell lung cancer; OS, overall survival; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; SNP, single-nucleotide polymorphism.

^aTen missing data were excluded.

^bAdjusted for age, sex, smoking status, histology, tumor stage, chemotherapy, surgery, radiotherapy, and principal components.

^cProtective genotypes were *ELP2* rs7242481 GA+AA and *PIAS1* rs1049493 TC+CC.

^dProtective alleles were *ELP2* rs7242481_A and *PIAS1* rs1049493_C.

(HR, 0.77; 95% CI, 0.67–0.89; $P = 0.0004$ for OS and HR, 0.79; 95% CI, 0.68–0.91; $P = 0.002$ for DSS), in comparison with the 0–1 score group. Because the NPAs was better than NPGs to evenly dichotomize the patients, we used NPAs to facilitate the stratification analysis. We further presented KM survival curves to depict these associations between protective alleles and NSCLC OS and DSS (Fig. 2A–D).

Stratified analysis for associations between NPAs and NSCLC survival

In the stratified analysis by age, sex, smoking status, histology, tumor stage, chemotherapy, radiotherapy, and surgery in the PLCO dataset, there were no obvious differences in survival, or interactions, between the strata of these covariates observed in either NSCLC OS or DSS ($P > 0.05$ for all strata; Supplementary Table S6).

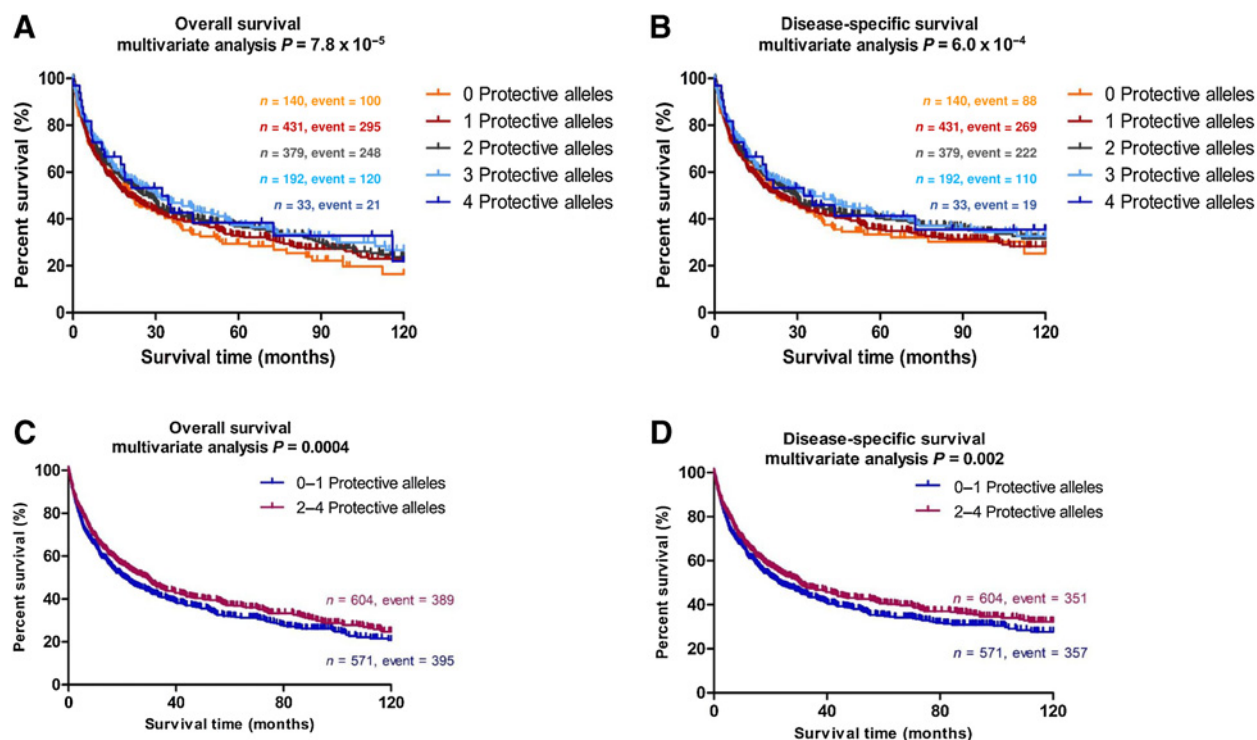
The ROC curves and time-dependent AUC

We assessed the predictive value of the two SNPs with time-dependent AUC and ROC curves at the 60th month (or 5-year

survival) in the PLCO dataset. Compared with the model of covariates including age, sex, smoking status, histology, tumor stage, chemotherapy, radiotherapy, surgery, and the first four principal components, the time-dependent AUC plot with addition of the two independent SNPs did not improve prediction performance of the model at the 60th month. On the other hand, when we performed the time-dependent AUC and ROC curves for survival of 120 months by combining the two independent SNPs with the 15 previously published SNPs using the same dataset, the prediction performance of the model was improved significantly for both OS and DSS. The AUCs changed from 87.42% to 89.81% ($P = 0.024$) for OS and from 87.95% to 90.46% ($P = 0.014$) for DSS (Supplementary Fig. S4A and S4B).

The eQTL analysis

In the RNA-sequencing data of lymphoblastoid cell lines from 373 European descendants available from the 1000 Genomes Project, the *ELP2* rs7242481 A allele showed a significant correlation

**Figure 2.**

Prediction of survival with combined protective alleles. KM survival curves for patients with NSCLC by NPA in the PLCO Trial. By 0, 1, 2, 3, and 4 protective alleles in OS (A) and in DSS (B). Dichotomized groups of the NPA into 0-1 and 2-4 in OS (C) and in DSS (D) from the PLCO Trial. DSS, disease-specific survival; KM, Kaplan-Meier; NPA, number of protective alleles; NSCLC, non-small cell lung cancer; OS, overall survival; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial.

with increased mRNA expression levels of *ELP2* in all additive, dominant, and recessive models ($P = 9.8 \times 10^{-5}$, $P = 0.005$, and $P < 2 \times 10^{-4}$, respectively; Fig. 3A; Supplementary Fig. S5A and S5B); however, there was no significant correlation between the *PIAS1* rs1049493 C allele and mRNA expression levels of *PIAS1* in all three genetic models (Fig. 3D; Supplementary Fig. S5C and S5D). Then, we performed eQTL by using the data of 369 whole-blood samples and 383 normal lung tissue from the GTEx project and found that the rs7242481 A allele remained significantly correlated with a higher expression level of *ELP2* in whole-blood samples but not in normal lung tissues ($P = 2.63 \times 10^{-15}$ and $P = 0.141$, respectively; Fig. 3C and B). The rs1049493 C allele was correlated with a lower expression level of *PIAS1* in both normal lung tissues and whole blood ($P = 0.008$ and $P = 0.0002$; Fig. 3E and F). Finally, we performed functional prediction for these two independent SNPs utilizing three online bioinformatics tools: SNPinfo, RegulomeDB, and HaploReg to predict their biological functions, as summarized in Supplementary Fig. S6 and Supplementary Table S7.

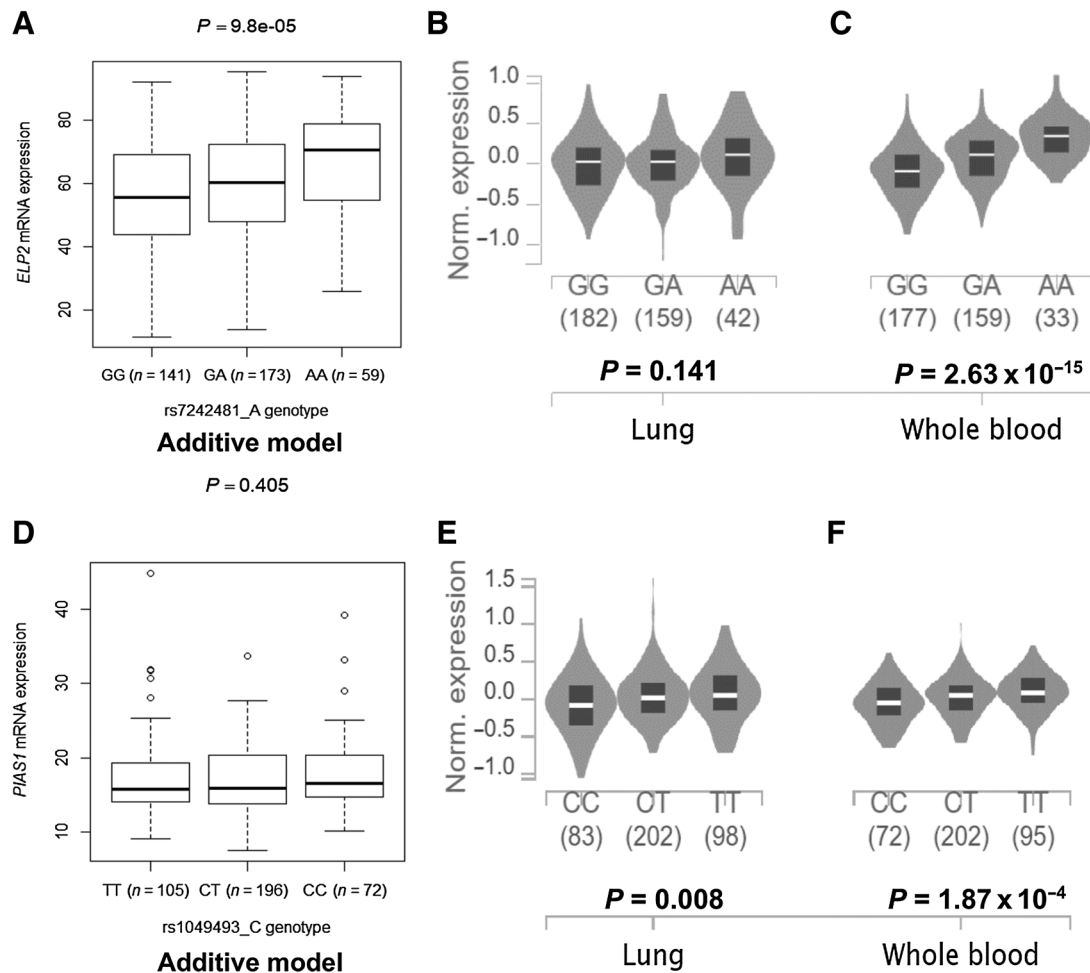
Differential mRNA expression analysis in target tissues

As shown in Supplementary Fig. S7A–S7C, in comparison with adjacent normal tissues, tumor tissues had a higher mRNA expression level of *ELP2* in lung adenocarcinoma, lung squamous cell carcinoma, and lung adenocarcinoma + lung squamous cell carcinoma samples combined ($P < 0.001$, $P < 0.001$, and $P < 0.001$, respectively). In the UALCAN (<http://ualcan.path.uab.edu>) database, the mRNA expression levels of *ELP2* were

also significantly higher in tumor tissues in lung adenocarcinoma ($P < 0.001$) and in lung squamous cell carcinoma ($P < 0.001$; Supplementary Fig. S7D and S7E). A higher expression level of *ELP2* was associated with improved NSCLC survival as shown by a KM survival curve of lung cancer constructed online (<http://kmplot.com/analysis/index.php?p=service&cancer=lung>). On the other hand, as shown in Supplementary Fig. S8, compared with adjacent normal tissues, tumor tissues had a lower mRNA expression level of *PIAS1* in lung adenocarcinoma ($P < 0.0001$), lung squamous cell carcinoma ($P < 0.001$), and lung adenocarcinoma + lung squamous cell carcinoma combined ($P < 0.001$). In the UALCAN (<http://ualcan.path.uab.edu>) database, the results were also similar, in which the mRNA expression levels of *PIAS1* in tumor tissues were lower in both lung adenocarcinoma tissues ($P < 0.001$) and lung squamous cell carcinoma tissues ($P < 0.001$) in comparison with normal tissues. However, a higher expression level of *PIAS1* mRNA was associated with a better NSCLC survival in the TCGA database.

Discussion

In the present study, we identified and validated two potentially functional and independent SNPs (i.e., *ELP2* rs7242481 and *PIAS1* rs1049493) that were significantly associated with the survival of patients with NSCLC in Caucasian populations. We also found that *ELP2* rs7242481 A allele was significantly associated with an increased mRNA expression level of *ELP2* in 373 lymphoblastoid cell lines from the 1000 Genomes Project and an increased mRNA

**Figure 3.**

Correlation of genotypes with the mRNA expression levels of the corresponding genes. eQTL analysis of the independent and significant SNPs associated with NSCLC survival. rs7242481 A allele is associated with a higher mRNA expression of *ELP2* in 373 European samples from the 1000 Genomes Project (A) and in 369 genotyped whole-blood samples (C), but not in 383 genotyped normal lung tissues from the GTEx project (B). rs1049493 C allele was not significantly associated with an altered mRNA expression of *PIAS1* in 373 European samples from the 1000 Genomes Project (D) but is significantly associated with a decreased mRNA expression of *PIAS1* in 383 normal lung tissue (E) and 369 whole-blood samples (F) from the GTEx project. eQTL, expression quantitative trait loci; GTEx, Genotype-Tissue Expression; NSCLC, non-small cell lung cancer; SNP, single-nucleotide polymorphism.

expression level in 369 whole-blood samples from the GTEx project. In addition, we also found that *PIAS1* rs1049493 C allele was significantly associated with a lower mRNA expression level of *PIAS1* in 383 normal lung tissues and 369 whole-blood samples from the GTEx project. These results were consistent with those of the gene expression analysis between paired tumor and adjacent normal tissue samples and survival analysis in the TCGA database. It is worth noting that a significantly higher mRNA expression level of *ELP2* was found in tumor tissues in the TCGA database, yet a higher expression level of *ELP2* was associated with a better survival using the same data.

ELP2, or *STAT3*-interacting protein 1 (*STATI1*), is located on chromosome 18 and is responsible for encoding the protein *ELP2* (elongator protein 2) that makes up the core subunit of the elongator complex, a histone acetyltransferase complex that is highly associated with RNA polymerase II and various cellular activities (36, 37). In *Homo sapiens*, *ELP2* is mostly known as *STATI1* or *StI1*, a well-known *STAT3* interactor that is involved in regulating cytokine signal

transduction, and an overexpression of *STATI1* inhibits *STAT3* activation and nuclear translocation (38). Studies in the past have also identified *STAT3* as a converging point of various oncogenic pathways, and the activated *STAT3* may trigger tumor progression by directly regulating oncogenic gene expression as well as promoting cancer growth via immunosuppression through the inhibition of immune mediators such as proinflammatory cytokines (39, 40).

In the present study, we found that the rs7242481 A allele was significantly associated with an increased mRNA expression level of *ELP2* in lymphoblastoid cell lines and whole-blood tissues; this finding is consistent with the results from other studies, suggesting that the novel genetic variant rs7242481 A allele may affect survival of patients with NSCLC by increasing *ELP2* mRNA expression, likely inhibiting the oncogenic effects of the activated *STAT3* (39, 40). In addition, we found that mRNA expression levels of *ELP2* in tumor tissues from the TCGA database were higher than in adjacent normal tissues in both paired and nonpaired samples and were associated with a better survival of NSCLC. These

suggest that overexpression of *ELP2* may have occurred as part of our innate immune response to alleviate STAT3 inhibitory effect on immune-stimulating cytokines, resulting in a more rapid activation of the innate immune system (38, 39). According to the ENCODE database, *ELP2* rs7242481G>A is located in a DNase I hypersensitive site with highly observable levels of histone modifications in H3K4Me3 and H3K27Ac acetylation, suggesting that the *ELP2* rs7242481 G>A SNP may lead to significantly enhanced transcriptional activities of *ELP2*.

PIAS1, or protein inhibitor of the activated STAT1, is located on chromosome 15 and is responsible for encoding the E3 SUMO-protein ligase enzyme, which plays a central role as a transcriptional coregulator of STAT1 (41). Recent studies suggest that the deregulation in SUMO (small ubiquitin-like modifier)-related pathways may lead to oncogenic transformation of tumor suppressors such as promyelocytic leukemia protein (PML; refs. 41, 42). An interesting finding is that most of the published molecular and functional studies have proposed or identified *PIAS1* as an oncogene in NSCLC and other malignant tumors, demonstrating that *PIAS1* plays an essential role in degrading PML and that the expression of PML and *PIAS1* are inversely correlated in NSCLC cell lines (42). However, results obtained from the TCGA database suggest that *PIAS1* may be a potential suppressor gene in NSCLC, because its expression was higher in adjacent normal tissues than in tumor tissues in both paired and nonpaired samples with additional evidence that higher expression levels of *PIAS1* to be associated with a better survival. Therefore, additional mechanistic studies are needed to unravel these discrepancies in the association between *PIAS1* expression levels and NSCLC survival. A potential explanation to the difference in results between TCGA analysis and evidence from studies that support the oncogenic role of *PIAS1* (42) might be because that the tumor tissues outlined in TCGA also contain fibroblast cells and immune cells, which may alter the resulting observation.

In the present study, we showed that the *PIAS1* rs1049493 C allele was associated with lower mRNA expression levels of *PIAS1* in whole blood and normal lung tissues, suggesting that the novel rs1049493 C variant allele may affect survival of NSCLC by lowering *PIAS1* mRNA expression levels; however, it must be noted that the directionality of *PIAS1* rs1049493 C allele yielded mixed results between mRNA levels of *PIAS1* and NSCLC survival. Mixed observations on *PIAS1* and cancer survival were also observed in other literatures as well as in the KM plotter for combined histology of lung cancer (refs. 42, 43; Supplementary Fig. S8F). As a result, the directionality and mechanisms of *PIAS1* in carcinogenesis and patient survival warrant additional molecular validation. According to the ENCODE database, *PIAS1* rs1049493 T>C is located in a DNase I hypersensitive site with observable levels of histone modifications in H3K4Me1 acetylation, suggesting that the *PIAS1* rs1049493 T>C SNP may lead to altered transcriptional activities of *PIAS1*.

Currently, there are conflicting outcomes in the IFN γ signaling: to increase an innate immune cell recruitment and type I Th-cell activation but to induce an increased tolerant molecule (i.e., PD-L1) expression in tumor cells to promote immune evasion (11, 44, 45). Also, further functional investigation into the mechanisms of the *ELP2* rs7242481 and *PIAS1* rs1049493 SNPs are warranted. Furthermore, because both the discovery and replication datasets were from Caucasian populations, our findings may not be generalizable to other ethnic populations. It is also worth noting that the discovery dataset and replication dataset had some differences in distribution of baseline characteristics, leading to fewer significant SNPs being validated (46).

Although the PLCO discovery dataset had a relatively large number of patients, the number of patients in some subgroups were still relatively small, which could cause a reduced statistical power in detecting potential weak effects of other SNPs. Furthermore, genetic mutation detection and tumor mutation burden (47) have been considered important for developing targeted therapies as the first-line treatments for cancer (i.e., PD-1 or PD-L1 inhibitors) in precision medicine as well as in the prediction of primary resistance to immunotherapy, but such treatment information was not available for further analysis in the present study. For the translational significance, any novel biomarkers for prognosis, such as the SNPs identified in the present study, would be clinically informative, if their effects on tumor phenotypes of NSCLC, such as tumor microenvironment (48), could be evaluated. This suggests that future studies need to collect such important clinical information in the study design, which could establish the associations between these IFN γ -related SNPs and immunotherapy-related NSCLC survival.

The observed directionality of mRNA levels of *PIAS1* and NSCLC survival appears to be inconsistent in the biological plausibility, likely due to the diversity of tissues and different patients used in the analysis, as observed in previously published studies (42, 43); therefore, the expression data should be obtained from appropriate and relevant tissue types of the same study populations to avoid discrepant results. Our findings may not have immediate clinical utility, because the magnitude of the HR for clinical significance needs to be much higher and well-verified compared with those from GWAS studies of survival. Once further validated, however, these findings may allow researchers to explore potentially functional SNPs in more rigorous experimental investigation for their biological plausibility and clinical utility.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: Y.C. Zhao, D.C. Christiani, Q. Wei

Development of methodology: Q. Wei

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H. Liu, D.C. Christiani, Q. Wei

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y.C. Zhao, D. Tang, S. Yang, H. Liu, S. Shen, D.C. Christiani, Q. Wei

Writing, review, and/or revision of the manuscript: Y.C. Zhao, S. Yang, S. Luo, T.E. Stinchcombe, C. Glass, D.C. Christiani, Q. Wei

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y.C. Zhao, S. Yang, L. Su, D.C. Christiani, Q. Wei

Study supervision: D.C. Christiani, Q. Wei

Disclaimer

The statements contained herein are solely those of the authors and do not represent or imply concurrence or endorsement by the NCI.

Acknowledgments

Q. Wei was supported by the V Foundation for Cancer Research (D201719) and also partly supported by the Duke Cancer Institute as part of the P30 Cancer Center Support Grant (grant ID: NIH/NCI CA014236). The Harvard Lung Cancer Susceptibility Study was supported by NIH grants U01CA209414, CA092824, CA074386, and CA090578 to D.C. Christiani. The authors thank all the participants of the PLCO Cancer Screening Trial. The authors also thank the NCI for providing access to the data collected by the PLCO Trial. The authors also acknowledge the dbGaP repository for providing cancer genotyping datasets. The authors wish to thank all the investigators and funding agencies that enabled the deposition of data in dbGaP and PLCO that we used in this study. The datasets used for the analyses described in this study were obtained from dbGaP at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession numbers phs000336.v1.p1 and phs000093.v2.p2. Principal investigators

are as follows: Maria Teresa Landi, Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, NCI, NIH (Bethesda, MD); and Neil E. Caporaso, Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, NCI, NIH (Bethesda, MD). Funding support for the GWAS of Lung Cancer and Smoking was provided through the NIH Genes, Environment and Health Initiative (GEI; Z01 CP 010200). The human subjects participating in the GWAS derive from The Environment and Genetics in Lung Cancer Etiology (EAGLE) case-control study and the PLCO Screening Trial, and these studies were supported by intramural resources of the NCI. Assistance with phenotype harmonization and genotype cleaning as well as with general study coordination was provided by the Gene Environment Association Studies, GENEVA Coordinating Center (U01HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by the NIH

GEI (U01HG004438). PLCO was also supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics and by contracts from the Division of Cancer Prevention, NCI, NIH, U.S. Department of Health and Human Services. The authors thank PLCO screening center investigators and staff, and the staff of Information Management Services Inc. and Westat Inc. Most importantly, the authors acknowledge trial participants for their contributions that made this study possible.

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 November 27, 2019; revised March 12, 2020; accepted May 29, 2020; published first June 3, 2020.

References

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin* 2019;69:7–34.
- Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc* 2008;83:584–94.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin* 2017;67:7–30.
- Lemjabbar-Alaoui H, Hassan OU, Yang YW, Buchanan P. Lung cancer: biology and treatment options. *Biochim Biophys Acta* 2015;1856:189–210.
- Zienolddiny S, Skaug V. Single nucleotide polymorphisms as susceptibility, prognostic, and therapeutic markers of nonsmall cell lung cancer. *Lung Cancer* 2012;3:1–14.
- Koury J, Lucero M, Cato C, Chang L, Geiger J, Henry D, et al. Immunotherapies: exploiting the immune system for cancer treatment. *J Immunol Res* 2018;2018:9585614.
- Pandya PH, Murray ME, Pollok KE, Renbarger JL. The immune system in cancer pathogenesis: potential therapeutic approaches. *J Immunol Res* 2016;2016:4273943.
- Mizukoshi E, Kaneko S. Telomerase-targeted cancer immunotherapy. *Int J Mol Sci* 2019;20:1823.
- Wang RF, Wang HY. Immune targets and neoantigens for cancer immunotherapy and precision medicine. *Cell Res* 2017;27:11–37.
- Gong J, Chehrizi-Raffle A, Reddi S, Salgia R. Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations. *J Immunother Cancer* 2018;6:8.
- Borcoman E, Kanjanapan Y, Champiat S, Kato S, Servois V, Kurzrock R, et al. Novel patterns of response under immunotherapy. *Ann Oncol* 2019;30:385–96.
- Fridley BL, Biernacka JM. Gene set analysis of SNP data: benefits, challenges, and future directions. *Eur J Hum Genet* 2011;19:837–43.
- Witte JS. Genome-wide association studies and beyond. *Annu Rev Public Health* 2010;31:9–20.
- Gallagher MD, Chen-Plotkin AS. The post-GWAS era: from association to function. *Am J Hum Genet* 2018;102:717–30.
- Zaidi MR, Merlino G. The two faces of interferon- γ in cancer. *Clin Cancer Res* 2011;17:6118–24.
- Lloyd CM, Hessel EM. Functions of T cells in asthma: more than just T(H)2 cells. *Nat Rev Immunol* 2010;10:838–48.
- Ivashkiv LB, Donlin LT. Regulation of type I interferon responses. *Nat Rev Immunol* 2014;14:36–49.
- Tau G, Rothman P. Biologic functions of the IFN- γ receptors. *Allergy* 1999;54:1233–51.
- Castro F, Cardoso AP, Goncalves RM, Serre K, Oliveira MJ. Interferon- γ at the crossroads of tumor immune surveillance or evasion. *Front Immunol* 2018;9:847.
- Hocking WG, Hu P, Oken MM, Winslow SD, Kvale PA, Prorok PC, et al. Lung cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. *J Natl Cancer Inst* 2010;102:722–31.
- Oken MM, Marcus PM, Hu P, Beck TM, Hocking W, Kvale PA, et al. Baseline chest radiograph for lung cancer detection in the randomized Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. *J Natl Cancer Inst* 2005;97:1832–9.
- Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, et al. The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet* 2007;39:1181–6.
- Tryka KA, Hao L, Sturcke A, Jin Y, Wang ZY, Ziyabari L, et al. NCBI's database of Genotypes and Phenotypes: dbGaP. *Nucleic Acids Res* 2014;42:975–9.
- Zhai R, Yu X, Wei Y, Su L, Christiani DC. Smoking and smoking cessation in relation to the development of co-existing non-small cell lung cancer with chronic obstructive pulmonary disease. *Int J Cancer* 2014;134:961–70.
- Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 2007;23:1294–6.
- Park JH, Geum DI, Eisenhut M, van der Vliet HJ, Shin JI. Bayesian statistical methods in genetic association studies: empirical examination of statistically non-significant Genome wide association study (GWAS) meta-analyses in cancers: a systematic review. *Gene* 2019;685:170–8.
- Wakefield J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet* 2007;81:208–27.
- Tang D, Zhao YC, Qian D, Liu H, Luo S, Patz EF, et al. Novel genetic variants in HDAC2 and PPARGC1A of the CREB-binding protein pathway predict survival of nonsmall cell lung cancer. *Mol Carcinog* 2020;59:104–15.
- Chambless LE, Diao G. Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Stat Med* 2006;25:3474–86.
- GTEX Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015;348:648–60.
- Lappalainen T, Sammeth M, Friedlander MR, Hoen PA, Monlong J, Rivas MA, et al. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* 2013;501:506–11.
- Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res* 2009;37:600–5.
- Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012;22:1790–7.
- Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res* 2016;44:877–81.
- Priem R, Welch R. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336–7.
- Dong C, Lin Z, Diao W, Li D, Chu X, Wang Z, et al. The Elp2 subunit is essential for elongator complex assembly and functional regulation. *Structure* 2015;23:1078–86.
- Winkler GS, Petrakis TG, Ethelberg S, Tokunaga M, Erdjument-Bromage H, Tempst P, et al. RNA polymerase II elongator holoenzyme is composed of two discrete subcomplexes. *J Biol Chem* 2001;276:32743–9.
- Collum RG, Brutsaert S, Lee G, Schindler C. A Stat3-interacting protein (StIP1) regulates cytokine signal transduction. *Proc Natl Acad Sci U S A* 2000;97:10120–5.
- Wang T, Niu G, Kortylewski M, Burdelya L, Shain K, Zhang S, et al. Regulation of the innate and adaptive immune responses by Stat-3 signaling in tumor cells. *Nat Med* 2004;10:48–54.

Zhao et al.

40. Wang Y, Shen Y, Wang S, Shen Q, Zhou X. The role of STAT3 in leading the crosstalk between human cancers and the immune system. *Cancer Lett* 2018;415:117–28.
41. Lee JS, Choi HJ, Baek SH. Sumoylation and its contribution to cancer. *Adv Exp Med Biol* 2017;963:283–98.
42. Rabellino A, Carter B, Konstantinidou G, Wu SY, Rimessi A, Byers LA, et al. The SUMO E3-ligase PIAS1 regulates the tumor suppressor PML and its oncogenic counterpart PML-RARA. *Cancer Res* 2012;72:2275–84.
43. Chanda A, Chan A, Deng L, Kornaga EN, Enwere EK, Morris DG, et al. Identification of the SUMO E3 ligase PIAS1 as a potential survival biomarker in breast cancer. *PLoS One* 2017;12:e0177639.
44. Ni L, Lu J. Interferon gamma in cancer immunotherapy. *Cancer Med* 2018;7:4509–16.
45. Qing Y, Stark GR. Alternative activation of STAT1 and STAT3 in response to interferon-gamma. *J Biol Chem* 2004;279:41679–85.
46. Wang Y, Liu H, Ready NE, Su L, Wei Y, Christiani DC, et al. Genetic variants in ABCG1 are associated with survival of nonsmall-cell lung cancer patients. *Int J Cancer* 2016;138:2592–601.
47. Berland L, Heeke S, Humbert O, Macoco A, Long-Mira E, Lasalle S, et al. Current views on tumor mutation burden in patients with non-small cell lung cancer treated by immune checkpoint inhibitors. *J Thoracic Dis* 2019;11:71–80.
48. Graves E, Maity A, Le Q. The tumor microenvironment in non-small cell lung cancer. *Semin Radiat Oncol* 2010;20:156–63.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association
for Cancer Research

Novel Variants of *ELP2* and *PIAS1* in the Interferon Gamma Signaling Pathway Are Associated with Non–Small Cell Lung Cancer Survival

Yu Chen Zhao, Dongfang Tang, Sen Yang, et al.

Cancer Epidemiol Biomarkers Prev Published OnlineFirst June 3, 2020.

Updated version	Access the most recent version of this article at: doi: 10.1158/1055-9965.EPI-19-1450
Supplementary Material	Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2020/06/03/1055-9965.EPI-19-1450.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, use this link http://cebp.aacrjournals.org/content/early/2020/07/01/1055-9965.EPI-19-1450 . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.