**Common *TDP1* polymorphisms in relation to survival among small cell lung cancer patients: a multicenter study from the International Lung Cancer Consortium**

Pawadee Lohavanichbutr, Lori C. Sakoda, Christopher I. Amos, Susanne M. Arnold, David C. Christiani, Michael P.A. Davies, John K. Field, Eric B. Haura, Rayjean J. Hung, Takashi Kohno, Maria Teresa Landi, Geoffrey Liu, Yi Liu, Michael W. Marcus, Grainne M. O'Kane, Matthew B. Schabath, Kouya Shiraishi, Stacey A. Slone, Adonina Tardón, Ping Yang, Kazushi Yoshida, Ruyang Zhang, Xuchen Zong, Gary E. Goodman, Noel S. Weiss, Chu Chen

**Author affiliations:**

1Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington. 2Division of Research, Kaiser Permanente, Oakland, California. 3Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire. 4Division of Medical Oncology, University of Kentucky Markey Cancer Center, Lexington, Kentucky. 5Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts. 6Roy Castle Lung Cancer Research Programme, Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool, UK. 7Thoracic Oncology Program, Moffitt Cancer Center, Tampa, Florida. 8Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario, Canada. 9Division of Genome Biology, National Cancer Research Institute, Tokyo, Japan. 10Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland. 11Department of Medical Oncology, Princess Margaret Cancer Center and University Health Network, University of Toronto, Toronto, Ontario, Canada. 12Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota. 13Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida. 14CIBEREpidemiology and Public Health (CIBER-ESP), Health Research Institute Carlos III, Madrid, Spain. 15University Institute of Oncology, University of Oviedo, Oviedo, Asturias, Spain. 16Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington. 17Department of Otolaryngology: Head and Neck Surgery, School of Medicine, University of Washington, Seattle, Washington

**Running title:** TDP1 polymorphism in relation to survival of SCLC patients

**Keywords:** TDP1, polymorphism, survival, SCLC

**Corresponding author:** Chu Chen, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Mailstop M5-C800, 1100 Fairview Ave N, Seattle, WA 98109-1024. Phone: 206-667-6644; Fax: 206-667-2537; E-mail: cchen@fredhutch.org

**Disclosure of Potential Conflicts of Interest**: No potential conflicts of interest were disclosed.

**Word count:** abstract 250 words, text 2,202 words

**Total number of figures and tables:** 2 figures, 3 tables, 3 supplementary tables, and 1 supplementary text

**Abstract**

**Background:** DNA topoisomerase inhibitors are commonly used for treating small cell lung cancer (SCLC). Tyrosyl-DNA phosphodiesterase (TDP1) repairs DNA damage caused by this class of drugs and may therefore influence treatment outcome. In this study, we investigated whether common *TDP1* single nucleotide polymorphisms (SNPs) are associated with overall survival among SCLC patients.

**Methods:** Two *TDP1* SNPs (rs942190 and rs2401863) were analyzed in 890 patients from 10 studies in the International Lung Cancer Consortium (ILCCO). The Kaplan-Meier method and Cox regression analyses were used to evaluate genotype associations with overall mortality at 36 months post-diagnosis, adjusting for age, sex, race, and tumor stage.

**Results:** Patients homozygous for the minor allele (GG) of rs942190 had poorer survival compared to those carrying AA alleles, with a hazard ratio (HR) of 1.36 (95% confidence interval (CI): 1.08-1.72, p-value=0.01), but no association with survival was observed for patients carrying the AG genotype (HR=1.04, 95% CI:0.84-1.29, p-value=0.72). For rs2401863, patients homozygous for the minor allele (CC) tended to have better survival than patients carrying AA alleles (HR=0.79, 95% CI: 0.61-1.02, p-value=0.07). Expression quantitative trait loci (eQTL) analyses from the Genotype-Tissue Expression Project indicated a potential effect of rs942190 on lung tissue, with higher *TDP1* gene expression for GG than AG or AA genotypes.

**Conclusions:** We found the rs942190 GG genotype to be associated with relatively poor survival among SCLC patients. Further investigation is needed to confirm the result and to determine whether this genotype may be a predictive marker for treatment efficacy of DNA topoisomerase inhibitors.

**Introduction**

Small cell lung cancer (SCLC) is the most aggressive form of lung cancer, with a 5-year survival of only 7% (1). Despite rapid advances in cancer therapy, treatment and overall survival of SCLC patients has changed little over the past few decades (2,3). Unlike non-small cell lung cancer (NSCLC), in which several prognostic and predictive biomarkers have been identified and targeted clinically (4), there are relatively few markers to predict survival or to guide treatment selection for SCLC patients (reviewed in (2,5)).

A combination of platinum chemotherapy and a DNA topoisomerase inhibitor is the first-line chemotherapy for treating SCLC patients (6). DNA topoisomerases (TOP1 and TOP2) are important players during DNA replication and transcription as they introduce transient DNA strand breaks (7). TOP1 inhibitors (e.g. Irinotecan, Topotecan) and TOP2 inhibitors (e.g. Etoposide, Teniposide) bind to DNA topoisomerases and generate drug-stabilized DNA cleavage complexes, which eventually result in tumor cell death (8,9). Tyrosyl-DNA phosphodiesterase (TDP1) plays a role in repairing both TOP1- and TOP2-mediated DNA damage (10,11) and it is believed to be responsible for drug resistance to DNA topoisomerase inhibitors (12,13). A study in SCLC cell lines suggests that the TDP1/TOP1 ratio may be an indicator for the response of SCLC to Topotecan (14); however, confirmation in SCLC tissue is lacking. Limited available tissue for such confirmation presents a challenge, since only a small portion of SCLC patients receive surgical resection.

Developing a blood-based marker to predict drug response would be useful to inform appropriate treatment for SCLC patients. Since *TDP1* plays a role in resistance to DNA topoisomerase inhibitors, it is plausible that patients carrying a *TDP1* variant may respond differently to treatment, thus having different survival outcomes. There are very few studies on *TDP1* single nucleotide polymorphisms (SNP) (15-17) and, to the best of our knowledge, none have examined *TDP1* SNPs in relation to SCLC survival. In this study, we investigated whether common *TDP1* SNPs are associated with overall survival among SCLC patients in a multicenter study from the International Lung Cancer Consortium (ILCCO, http://ilcco.iarc.fr).

**Materials and methods**

**Study population**

This study consists of 898 SCLC patients from 10 ILCCO studies that have data on patient survival time and vital status (Table 1). Further details on study population and source of data for each study are provided in the Supplementary Text. All participants provided written informed consent, and each study was approved by its local institutional review board. For the current study, SCLC includes small cell carcinoma, combined small cell cancer, and neuroendocrine carcinoma (ICD-O 8013, 8041, 8042, 8043, 8044, 8045, 8246).

**SNP selection and genotyping**

Tag SNPs for the *TDP1* gene region (± 2.5 kb of the coding sequence) were identified using the Genome Variation Server (http://gvs.gs.washington.edu). SNPs were classified into bins with a pairwise linkage disequilibrium (LD) threshold of r² ≥ 0.8 using the IdSelect algorithm (18). The list of *TDP1* tag SNPs based on HapMap Phase I and II Centre d’Etude du Polymorphism Humain (CEU) population was shown in the Supplementary Table S1. One SNP per bin of the tag SNPs with an average minor allele frequency (MAF) ≥ 5% (a total of six SNPs) was selected by prioritizing on the SNP function class and predicted genotyping success based on Illumina assay design score. Six *TDP1* tag SNPs (rs9488, rs942190, rs1286927, rs2401863, rs4143999, and rs12880397) were genotyped on 1,586 healthy controls and 793 lung cancer cases (including 137 SCLC) from the β-Carotene and Retinol Efficacy Trial (CARET) as part of a study on germ line variation in DNA repair genes and lung cancer risk (19,20). Four of the six SNPs had low MAF among SCLC patients (0.03-0.07) and were excluded from further investigation since a very large sample size would be needed to determine the effect of these SNPs. Thus only two SNPs (rs942190 and rs2401863) were chosen for the current pooled analysis. These two SNPs are partially correlated, especially among individuals of European ancestry with r2of 0.63 (r2(East Asian) = 0.26).

The majority of genotype data for our pooled analysis were obtained from the OncoArray, a custom array manufactured by Illumina which contains approximately 500K SNPs that provide genome-wide coverage of most common genetic variants along with markers of interest for common cancers (21). Genotype data from the Mayo Clinic and part of the genetic data from the Lunenfeld-Tanenbaum Research Institute were from existing genome-wide association studies (GWAS). Samples from CARET participants were genotyped using a custom-designed 384-plex GoldenGate assay (Illumina). Samples from Japan were genotyped using a pre-design (for rs942190) and a custom-design (for rs2401863) TaqMan assay (Applied Biosystems). Race-specific genotype frequencies for both SNPs were in agreement with Hardy-Weinberg equilibrium (Chi-square p-values for rs942190 among White, rs942190 among Asian, rs2401863 among White, and rs2401863 among Asian were 0.41, 0.14, 0.54, and 0.20, respectively).

**Statistical analyses**

Clinical and genotype data were harmonized across studies. Characteristics of all 898 patients by study site are summarized in Supplementary Table S2. Race was imputed as White for the 96 patients of unknown race since their genotype distributions for both SNPs were similar to those of White patients (Supplementary Table S3). Tumor stage was classified as limited stage (LS or stage I-III) and extensive stage (ES or stage IV). The primary outcome was overall mortality as of 36 months post-diagnosis (when deaths are commonly attributed to lung cancer), measured from the date of lung cancer diagnosis until the date of death, last contact, or censoring at 36 months follow-up, whichever occurred first. Disease-specific survival was not examined, since cause of death was missing for 43% of the patients.

Survival analyses were performed using Kaplan-Meier survival plots and Cox proportional hazard regression models with a robust estimator of variance adjusting for age, sex, race (White vs. Asian), and tumor stage. Analyses were conducted to evaluate genotype and haplotype associations with overall mortality at 36 months post-diagnosis. Six patients with no follow-up data and two patients with no genotype data for both SNPs were excluded from survival analyses. Of the remaining 890 patients, six and two did not have genotype data for rs942190 and rs2401863, respectively. Since the SNP genotype frequencies were quite different between Whites and Asians, we also conducted a subgroup analysis by race for each SNP. Analyses were performed using STATA® 14 (StataCorp, College Station, Texas). Haplotype analysis was performed using the THESIAS (Testing Haplotype Effects In Association Studies) software version 3.1 (22), which is based on the Stochastic expectation maximization algorithm (23). Hazard ratios (HR) and 95% confidence intervals (CI) adjusting for age, sex, and tumor stage were calculated using the most common haplotype as the reference. Haplotype analysis was performed on White patients only since the two SNPs were correlated among Whites and sample sizes for Asian and other races were limited.

**Results**

Selected characteristics of patients included in the survival analyses are presented in Table 2. The majority of patients were male, non-Hispanic White, and either current or former smokers. There was a slightly higher proportion of patients with limited stage than extensive stage SCLC. Approximately 90% of patients had died by the time of last follow-up and 87.5% of deceased patients died within 36 months after diagnosis of SCLC. The median follow-up time for patients who were alive at last follow-up was 73 months (ranged from 3 to 234 months). The allele frequencies of the two SNPs differed between persons of European and East Asian ancestry. The MAFs of rs942190 (G allele) for White and Asian patients in this study were 0.49 and 0.23, respectively, and for rs2401863 (C allele) were 0.38 and 0.52, respectively. Mean age at diagnosis was similar for patients in each genotype group. There was a slightly lower proportion of female and tumors of limited stage among patients with the rs942190 AA genotype compared to patients with the other two rs942190 genotypes. The proportions of tumor stage were comparable by rs2401863 genotype.

Kaplan-Meier (KM) analyses for all patients with known vital status and genotype demonstrated poorer survival for patients homozygous for the minor allele (GG) of rs942190 compared to those carrying the other two genotypes (Figure 1a). For rs2401863, better survival was associated with carrying both minor alleles (CC); however, the association was not statistically significant (Figure 1b).

The results from multivariable Cox regression analyses (Table 3) were consistent with the results from Kaplan-Meier analyses. Patients carrying GG of rs942190 had poorer survival compared to those with the AA genotype, with a HR of 1.36 (95% CI: 1.08-1.72, p-value=0.01), but no association with survival was observed for patients with the heterozygous (AG) genotype (HR=1.04, 95% CI: 0.84-1.29, p-value=0.72). The HRs associated with the presence of the GG genotype were in the same direction for Whites and Asians. For rs2401863, patients carrying two minor alleles (CC genotype) tended to have better survival than patients carrying the AA genotype (HR=0.79, 95% CI: 0.61-1.02, p-value=0.07); however, this inverse association with survival was observed only in White patients (HR=0.71, 95% CI: 0.54-0.94, p-value=0.02). The association was, if anything, in the opposite direction among Asian patients (HR=2.11, 95% CI: 0.90-4.95, p-value=0.09). The most common haplotype among White patients was the haplotype containing the G allele of rs942190 and the A allele of rs2401863. The haplotype containing the rs942190 A allele and the rs2401863 C allele was associated with better survival, compared to the most common haplotype (HR=0.84; 95% CI: 0.73-0.95, p-value=0.008).

We also examined potential functional consequences of the two SNPs using a single tissue expression quantitative trait loci (eQTL) analysis from the Genotype-Tissue Expression (GTEx) Project (www.gtexportal.org). The GTEx Project, funded by The National Institutes of Health Common Fund, has collected and analyzed genomic variation from blood and gene expression in multiple tissues of the non-diseased donor in order to determine how genetic variation affects gene expression in human tissues (24). Based on the analysis available from the GTex website, *TDP1* gene expression was higher in lung tissues of people with the GG genotype of rs942190 than of people with AG or AA genotypes (p-value = 0.0008) (Figure 2). In contrast, there was minimal difference of *TDP1* gene expression in lung tissue across rs2401863 genotypes (p-value=0.12).

**Discussion**

To our knowledge, the current study is the first to investigate germline variation of *TDP1* in relation to survival among SCLC patients. Leveraging data from ten ILCCO studies, we analyzed a fairly large cohort of SCLC patients with near complete follow-up at 36 months. Of the two SNPs examined, we found the rs942190 GG genotype to be associated with poorer overall survival compared to AA genotype.

Several lines of evidence support the potential function of rs942190 in influencing treatment response and survival of SCLC patients. Based on SNP functional prediction (<http://snpinfo.niehs.nih.gov/snpinfo/snpfunc.php>) (25), rs942190 is located in a transcription factor binding site that may affect TDP1 gene expression. It has been shown that overexpression of TDP1 in cell lines could counteract the effect of DNA topoisomerase inhibitors (26); therefore, one would expect that patients with higher TDP1 in lung tissue may have more resistance to treatment with DNA topoisomerase inhibitors. The observed higher *TDP1* expression in lung tissue of individuals with the rs942190 GG genotype from the GTEx analysis is in line with our finding that patients with the GG genotype had poorer survival than patients with the other two genotypes. Conversely, we did not find a clear association between rs2401863 genotype and survival, which is consistent with the lack of association between the rs2401863 genotype and *TDP1* expression in lung tissue. Our observed association of the rs2401863 genotype with survival among White patients only may be due to the linkage with rs942190 SNP.

Unlike a clinical trial, the data on treatment was not extensively collected since the majority of the SCLC patients in our pooled analysis came from case-control studies. Approximately 30% of all patients had unknown treatment modality, mostly unknown whether chemotherapy was given. However, because most of these patients had limited stage SCLC, we would expect that these patients likely received chemotherapy (cisplatin/carboplatin plus DNA topoisomerase inhibitor) with or without radiation as the standard of care.

Even for those known to receive chemotherapy, there was limited detail on chemo drugs, courses of treatment, chemotherapy completion, and response to chemotherapy. These data are important to determine the effect of individual SNPs on treatment response. When we examined the association of rs942190 with survival restricted to the subgroup of 526 patients known to receive chemotherapy, the magnitude of the association was attenuated (GG vs. AA genotype: HR= 1.12; 95% CI: 0.86-1.47). Ideally, the comparison would be restricted to patients who received the same treatment regimen since DNA topoisomerase inhibitors are used in combination with other drugs for treating SCLC patients. Moreover, there are two types of DNA topoisomerase inhibitors: TOP1 and TOP2 inhibitors. The effect of rs942190 may not be the same for different types of DNA topoisomerase inhibitors.

In conclusion, our study suggests an association between rs942190 genotype and overall survival at 36 months after SCLC diagnosis. However, without comprehensive data on treatment and treatment response, we are not able to determine whether the difference in survival is due to the influence of the SNP on response to DNA topoisomerase inhibitors. Further assessment of the genotype-survival association in a larger study with more detailed and complete treatment data is needed to confirm our findings.

**Disclaimers**

Views and opinions of, and endorsements by the author(s) do not reflect those of the US Army or the Department of Defense.

**Acknowledgements**

We would like to thank the study participants for their involvement.

**Grant Support**

ILCCO Data Repository was supported by Cancer Care Ontario Research Chair of Population Health, National Institute of Health (U19 CA148127), and Lunenfeld-Tanenbuaum Research Institute, Sinai Health System.

The Cancer de Pulmon en Asturias (CAPUA) study was supported by Fondo de investigación sanitaria. Instituto de Salud Carlos III. Consorcio de Investigación Biomédica en Red (CIBER) del Área de Epidemiología y Salud Pública and University of Oviedo.

The Environment And Genetics in Lung cancer Etiology (EAGLE) study was supported by the Intramural Research Program of NIH, NCI, Division of Cancer Epidemiology and Genetics.

The Epidemiology & Genetics of Lung cancer (EGLC) study was supported by the National Cancer Institute and National Institute of Health: R03-CA77118 (PI P Yang) and R01s-CA80127 (PI P Yang), CA84354 (PI P Yang), and HL107612 (MPI P Yang & C Wendt).

The Carotene and Retinol Efficacy Trial (CARET) was supported by the National Cancer Institute and National Institute of Health: 5-UM1-CA-167462, (PI Gary E. Goodman), U01-CA63673 (PIs G. Omenn, G. Goodman), R01-CA111703 (PI Chu Chen)

The Harvard Lung Cancer Study (LCS) was supported by National Institute of Health grants R01CA092824, R01CA074386, and P30 ES000002

The Japan lung cancer study was supported by the National Cancer Center Research and Development Fund (NCC Biobank).

Kentucky Lung Cancer Research Initiative (LCRI) was supported by the Department of Defense [Congressionally Directed Medical Research Program, U.S. Army Medical Research and Materiel Command Program] under award number: 10153006 (W81XWH-11-1-0781). This research was also supported by unrestricted infrastructure funds from the UK Center for Clinical and Translational Science, NIH grant UL1TR000117 and Markey Cancer Center NCI Cancer Center Support Grant (P30 CA177558) Shared Resource Facilities: Cancer Research Informatics, Biospecimen and Tissue Procurement, and Biostatistics and Bioinformatics.

The Liverpool Lung Project (LLP) is supported by the Roy Castle Lung Cancer Foundation, UK.

The work performed for the Toronto lung cancer study was supported by Ontario Institute for Cancer Research, the Canadian Cancer Society Research Institute (020214), Ontario Institute of Cancer and Cancer Care Ontario Chair Award to R.J.H. and G.L. and the Alan Brown Chair and Lusi Wong Programs at the Princess Margaret Hospital Foundation.

The Total Lung Cancer (TLC) study was supported by the following funding sources: James & Esther King Biomedical Research Program Grant (09KN-15), National Institutes of Health Specialized Programs of Research Excellence (SPORE) Grant (P50 CA119997), and an American Cancer Society Institutional Research Grant (93-032-13).

**References**

1. American Cancer S. Cancer Facts & Figures 2016. Atlanta: American Cancer Society; 2016.

2. Byers LA, Rudin CM. Small cell lung cancer: where do we go from here? Cancer **2015**;121(5):664-72 doi 10.1002/cncr.29098.

3. Kalemkerian GP. Small Cell Lung Cancer. Semin Respir Crit Care Med **2016**;37(5):783-96 doi 10.1055/s-0036-1592116.

4. Thakur MK, Gadgeel SM. Predictive and Prognostic Biomarkers in Non-Small Cell Lung Cancer. Semin Respir Crit Care Med **2016**;37(5):760-70 doi 10.1055/s-0036-1592337.

5. van Meerbeeck JP, Fennell DA, De Ruysscher DK. Small-cell lung cancer. Lancet **2011**;378(9804):1741-55 doi 10.1016/s0140-6736(11)60165-7.

6. Rudin CM, Ismaila N, Hann CL, Malhotra N, Movsas B, Norris K*, et al.* Treatment of Small-Cell Lung Cancer: American Society of Clinical Oncology Endorsement of the American College of Chest Physicians Guideline. J Clin Oncol **2015**;33(34):4106-11 doi 10.1200/jco.2015.63.7918.

7. Wang JC. Cellular roles of DNA topoisomerases: a molecular perspective. Nat Rev Mol Cell Biol **2002**;3(6):430-40 doi 10.1038/nrm831.

8. Hande KR. Etoposide: four decades of development of a topoisomerase II inhibitor. Eur J Cancer **1998**;34(10):1514-21.

9. Pommier Y. Drugging topoisomerases: lessons and challenges. ACS Chem Biol **2013**;8(1):82-95 doi 10.1021/cb300648v.

10. Nitiss KC, Malik M, He X, White SW, Nitiss JL. Tyrosyl-DNA phosphodiesterase (Tdp1) participates in the repair of Top2-mediated DNA damage. Proc Natl Acad Sci U S A **2006**;103(24):8953-8 doi 10.1073/pnas.0603455103.

11. Murai J, Huang SY, Das BB, Dexheimer TS, Takeda S, Pommier Y. Tyrosyl-DNA phosphodiesterase 1 (TDP1) repairs DNA damage induced by topoisomerases I and II and base alkylation in vertebrate cells. J Biol Chem **2012**;287(16):12848-57 doi 10.1074/jbc.M111.333963.

12. Dexheimer TS, Antony S, Marchand C, Pommier Y. Tyrosyl-DNA phosphodiesterase as a target for anticancer therapy. Anticancer Agents Med Chem **2008**;8(4):381-9.

13. Beretta GL, Cossa G, Gatti L, Zunino F, Perego P. Tyrosyl-DNA phosphodiesterase 1 targeting for modulation of camptothecin-based treatment. Curr Med Chem **2010**;17(15):1500-8.

14. Meisenberg C, Ward SE, Schmid P, El-Khamisy SF. TDP1/TOP1 Ratio as a Promising Indicator for the Response of Small Cell Lung Cancer to Topotecan. J Cancer Sci Ther **2014**;6(7):258-67 doi 10.4172/1948-5956.1000280.

15. Wu BT, Lin WY, Chou IC, Liu HP, Lee CC, Tsai Y*, et al.* Association of tyrosyl-DNA phosphodiesterase 1 polymorphism with Tourette syndrome in Taiwanese patients. J Clin Lab Anal **2013**;27(4):323-7 doi 10.1002/jcla.21606.

16. Hoskins JM, Marcuello E, Altes A, Marsh S, Maxwell T, Van Booven DJ*, et al.* Irinotecan pharmacogenetics: influence of pharmacodynamic genes. Clin Cancer Res **2008**;14(6):1788-96 doi 10.1158/1078-0432.ccr-07-1472.

17. Hoskins JM, Rosner GL, Ratain MJ, McLeod HL, Innocenti F. Pharmacodynamic genes do not influence risk of neutropenia in cancer patients treated with moderately high-dose irinotecan. Pharmacogenomics **2009**;10(7):1139-46 doi 10.2217/pgs.09.35.

18. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet **2004**;74(1):106-20.

19. Sakoda LC, Loomis MM, Doherty JA, Julianto L, Barnett MJ, Neuhouser ML*, et al.* Germ line variation in nucleotide excision repair genes and lung cancer risk in smokers. Int J Mol Epidemiol Genet **2012**;3(1):1-17.

20. Doherty JA, Sakoda LC, Loomis MM, Barnett MJ, Julianto L, Thornquist MD*, et al.* DNA repair genotype and lung cancer risk in the beta-carotene and retinol efficacy trial. Int J Mol Epidemiol Genet **2013**;4(1):11-34.

21. Amos CI, Dennis J, Wang Z, Byun J, Schumacher FR, Gayther SA*, et al.* The OncoArray Consortium: a Network for Understanding the Genetic Architecture of Common Cancers. Cancer Epidemiol Biomarkers Prev **2016** doi 10.1158/1055-9965.epi-16-0106.

22. Tregouet DA, Garelle V. A new JAVA interface implementation of THESIAS: testing haplotype effects in association studies. Bioinformatics **2007**;23(8):1038-9 doi 10.1093/bioinformatics/btm058.

23. Li SS, Khalid N, Carlson C, Zhao LP. Estimating haplotype frequencies and standard errors for multiple single nucleotide polymorphisms. Biostatistics (Oxford, England) **2003**;4(4):513-22.

24. The Genotype-Tissue Expression (GTEx) project. Nat Genet **2013**;45(6):580-5 doi 10.1038/ng.2653.

25. Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. Nucleic Acids Res **2009**;37(Web Server issue):W600-5 doi 10.1093/nar/gkp290.

26. Barthelmes HU, Habermeyer M, Christensen MO, Mielke C, Interthal H, Pouliot JJ*, et al.* TDP1 overexpression in human cells counteracts DNA damage mediated by topoisomerases I and II. J Biol Chem **2004**;279(53):55618-25 doi 10.1074/jbc.M405042200.

**Table 1.** Studies included in the pooled analysis

|  |  |  |  |
| --- | --- | --- | --- |
| **Study Name** | **Principal Investigator** | **Country** | **n** |
| CAncer de PUlmon en Asturias (CAPUA) | Adonina Tardón  | Spain | 137 |
| Environment and Genetics in Lung Cancer Etiology (EAGLE) | Maria Teresa Landi  | Italy | 189 |
| Epidemiology & Genetics of Lung cancer study (EGLC), Mayo Clinic | Ping Yang  | USA | 74 |
| FHCRC Molecular Epidemiology of Lung Cancer (Ancillary study to CARET) | Chu Chen | USA | 137 |
| Harvard Lung Cancer Study (LCS)  | David C. Christiani  | USA | 176 |
| Japan lung cancer study  | Kouya Shiraishi  | Japan | 87 |
| Kentucky Lung Cancer Research Initiative (LCRI) | Susanne M. Arnold | USA | 8 |
| Liverpool Lung Project (LLP) | John K. Field  | UK | 55 |
| Toronto lung cancer study\* | Rayjean J. Hung, Geoffrey Liu  | Canada | 25 |
| Total Lung Cancer: Molecular Epidemiology of Lung Cancer Survival (TLC) | Matthew B. Schabath | USA | 10 |

\* from Mount Sinai Hospital and Princess Margaret Cancer Centre (MSH-PMH) study and Great Toronto Area Study

**Table 2.** Selected characteristics of SCLC patients by genotype

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Total**  | **rs942190 (n=884)** | **rs2401863 (n=888)** |
|  | **(n=890)** | **AA (n=259)** | **AG (n=423)** | **GG (n=202)** | **AA (n=336)** | **AC (n=408)** | **CC (n=144)** |
| **Age at diagnosis, years**  |  |  |  |  |  |  |  |
|  Range | 24 - 87 | 24 - 87 | 39 - 85 | 39 - 86 | 34 - 86 | 39 - 85 | 24 - 87 |
|  Mean (SD) | 65.7 (8.8) | 65.4 (9.0) | 65.8 (8.3) | 65.9 (9.5) | 66.2 (9.0) | 65.2 (8.4) | 65.7 (9.4) |
| **Sex** |  |  |  |  |  |  |  |
|  Female | 294 (33.0%) | 77 (29.7%) | 149 (35.2%) | 67 (33.2%) | 110 (32.7%) | 140 (34.3%) | 44 (30.6%) |
|  Male | 596 (67.0%) | 182 (70.3%) | 274 (64.8%) | 135 (66.8%) | 226 (67.3%) | 268 (65.7%) | 100 (69.4%) |
| **Race** |  |  |  |  |  |  |  |
|  White\* | 798 (89.7%) | 206 (79.5%) | 395 (93.4%) | 195 (96.5%) | 312 (92.9%) | 367 (90.0%) | 118 (81.9%) |
|  Asian | 87 (9.8%) | 51 (19.7%) | 25 (5.9%) | 7 (3.5%) | 23 (6.9%) | 37 (9.1%) | 26 (18.1%) |
|  Others | 5 (0.6%) | 2 (0.8%) | 3 (0.7%) | 0 | 1 (0.3%) | 4 (1.0%) | 0 |
| **Ethnicity** |  |  |  |  |  |  |  |
|  Hispanic | 2 (0.6%) | 0 | 1 (0.7%) | 1 (1.3%) | 2 (1.5%) | 0 | 0 |
|  Not Hispanic | 354 (99.4%) | 128 (100%) | 146(99.3%) | 76 (98.7%) | 128 (98.5%) | 159 (100%) | 66 (100%) |
|  Unknown | 534 | 131 | 276 | 125 | 206 | 249 | 78 |
| **Smoking Status** |  |  |  |  |  |  |  |
|  Never | 43 (4.9%) | 15 (5.8%) | 13 (3.1%) | 14 (7.0%) | 24 (7.1%) | 10 (2.5%) | 9 (6.3%) |
|  Former | 306 (34.7%) | 84 (32.4%) | 158 (37.9%) | 62 (30.8%) | 110 (32.7%) | 145 (36.1%) | 51 (35.4%) |
|  Current | 534 (60.5%) | 160 (61.8%) | 246 (59.0%) | 125 (62.2%) | 201 (59.8%) | 247 (61.4%) | 84 (58.3%) |
|  Unknown | 7 | 0 | 6 | 1 | 1 | 6 | 0 |
| **Tumor Stage** |  |  |  |  |  |  |  |
|  Limited Stage | 418 (57.3%) | 117 (54.9%) | 203 (59.2%) | 97 (57.7%) | 162 (57.9%) | 186 (56.7%) | 69 (58.0%) |
|  Extensive Stage | 311 (42.7%) | 96 (45.1%) | 140 (40.8%) | 71 (42.3%) | 118 (42.1%) | 142 (43.3%) | 50 (42.0%) |
|  Unknown | 161 | 46 | 80 | 34 | 56 | 80 | 25 |
| **Vital Status at 3 years** |  |  |  |  |  |  |  |
|  Alive | 191 (21.5%) | 64 (24.7%) | 98 (23.2%) | 27 (13.4%) | 61 (18.2%) | 89 (21.8%) | 41 (28.5%) |
|  Death | 699 (78.5%) | 195 (75.3%) | 325 (76.8%) | 175 (86.6%) | 275 (81.8%) | 319 (78.2%) | 103 (71.5%) |
| \* included White and unknown race (imputed as White) |

**Table 3.** Results of multivariable Cox proportional regression analyses

|  |  |  |  |
| --- | --- | --- | --- |
| **SNP** | **Genotype** | **Adjusted HR\* [95% CI]** | **p-value** |
| rs942190 (all) | AA | 1.00 |  |
|  | AG | 1.04 [0.84, 1.29] | 0.719 |
|  | GG | 1.36 [1.08, 1.72] | 0.010 |
|  White\*\* only | AA | 1.00 |  |
|  | AG | 1.09 [0.87, 1.36] | 0.458 |
|  | GG | 1.39 [1.09, 1.77] | 0.008 |
|  Asian only | AA | 1.00 |  |
|  | AG | 0.50 [0.21, 1.19] | 0.116 |
|   | GG | 1.38 [0.63, 2.98] | 0.420 |
| rs2401863 (all) | AA | 1.00 |  |
|  | AC | 0.91 [0.76, 1.10] | 0.332 |
|  | CC | 0.79 [0.61, 1.02] | 0.071 |
|  White\*\* only | AA | 1.00 |  |
|  | AC | 0.91 [0.76, 1.11] | 0.354 |
|  | CC | 0.71 [0.54, 0.94] | 0.016 |
|  Asian only | AA | 1.00 |  |
|  | AC | 0.94 [0.40, 2.20] | 0.885 |
|   | CC | 2.11 [0.90, 4.95] | 0.085 |
|  | **Haplotype** |  |  |
| rs942190/rs2401863  | GA | 1.00 |  |
| (White\*\* only) | AC | 0.84 [0.73-0.95] | 0.008 |
|  | AA | 0.88 [0.73-1.06] | 0.165 |
|  | GC | 0.85 [0.51-1.42] | 0.541 |
| \* adjusted for age, sex, race, and tumor stage for all patients and adjusted for age, sex, and tumor stage for subgroup analyses |
| \*\* including White and unknown race (imputed as White) |

**Figure 1.**

**A B**

** **

Kaplan-Meier survival curves among 890 patients with SCLC. **A**, Stratified by rs942190 genotype. **B**, Stratified by rs2401863 genotype.

**Figure 2.**



Box plot from the Genotype-Tissue Expression (GTEx) Project demonstrated higher *TDP1* gene expression in lung tissues of individuals with rs942190 GG genotype compared to other genotypes. HomoRef, Het, and Homo Alt refer to individuals with AA, AG, and GG genotype, respectively.