12-2017

Common *TDP1* Polymorphisms in Relation to Survival Among Small Cell Lung Cancer Patients: A Multicenter Study from the International Lung Cancer Consortium

Pawadee Lohavanichbutr
Fred Hutchinson Cancer Research Center

Lori C. Sakoda
Fred Hutchinson Cancer Research Center

Christopher I. Amos
Dartmouth College

Susanne M. Arnold
University of Kentucky, susanne.arnold@uky.edu

David C. Christiani
Harvard University

See next page for additional authors

Authors
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, and Chu Chen

Common *TDP1* Polymorphisms in Relation to Survival Among Small Cell Lung Cancer Patients: A Multicenter Study from the International Lung Cancer Consortium

Notes/Citation Information
Published in *Clinical Cancer Research*, v. 23, issue 24, p. 7550-7557.

Digital Object Identifier (DOI)
https://doi.org/10.1158/1078-0432.CCR-17-1401

Published in final edited form as:

Clin Cancer Res. 2017 December 15; 23(24): 7550–7557. doi:10.1158/1078-0432.CCR-17-1401.

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

Pawadee Lohavanichbutr[1], **Lori C. Sakoda**[1,2], **Christopher I. Amos**[3], **Susanne M. Arnold**[4], **David C. Christiani**[5], **Michael P.A. Davies**[6], **John K. Field**[6], **Eric B. Haura**[7], **Rayjean J. Hung**[8], **Takashi Kohno**[9], **Maria Teresa Landi**[10], **Geoffrey Liu**[11], **Yi Liu**[12], **Michael W. Marcus**[6], **Grainne M. O'Kane**[11], **Matthew B. Schabath**[13], **Kouya Shiraishi**[9], **Stacey A. Slone**[14], **Adonina Tardón**[15,16], **Ping Yang**[12], **Kazushi Yoshida**[9], **Ruyang Zhang**[5], **Xuchen Zong**[8], **Gary E. Goodman**[1], **Noel S. Weiss**[1,17], and **Chu Chen**[1,17,18]

[1]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington [2]Division of Research, Kaiser Permanente Northern California, Oakland, California [3]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire [4]Division of Medical Oncology, University of Kentucky Markey Cancer Center, Lexington, Kentucky [5]Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts [6]Roy Castle Lung Cancer Research Programme, Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool, UK [7]Department of Thoracic Oncology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida [8]Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario, Canada [9]Division of Genome Biology, National Cancer Research Institute, Tokyo, Japan [10]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland [11]Department of Medical Oncology, Princess Margaret Cancer Center and University Health Network, University of Toronto, Toronto, Ontario, Canada [12]Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota [13]Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida [14]Biostatistics and Bioinformatics Shared Resource Facility, University of Kentucky Markey Cancer Center, Lexington, Kentucky [15]CIBER Epidemiology and Public Health (CIBER-ESP), Health Research Institute Carlos III, Madrid, Spain [16]University Institute of Oncology, University of Oviedo, Oviedo, Asturias, Spain [17]Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington [18]Department of Otolaryngology: Head and Neck Surgery, School of Medicine, University of Washington, Seattle, Washington

Abstract

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; cchen@fredhutch.org.

Disclaimers

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

Disclosure of Potential Conflicts of Interest: The authors declare no potential conflicts of interest.

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). Results from the Genotype Tissue Expression (GTEx) Project, the Encyclopedia of DNA Elements (ENCODE), and the ePOSSUM web application support the potential function of rs942190.

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.

Keywords

TDP1; polymorphism; survival; SCLC

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 the 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 $r^2 \geq 0.8$ 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 r^2 of 0.63 ($r^2_{(\text{East Asian})}$ = 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). Chemotherapy drug use was classified as “TOP1 inhibitor” (received any TOP1 inhibitor along the courses of chemotherapy), “TOP2 inhibitor” (received any TOP2 inhibitor along the courses of chemotherapy), and “Other/Unknown” (i.e., not known to have receives any TOP1 or TOP2 inhibitor). The “TOP1 inhibitor” group and the “TOP2 inhibitor” group also contained patients who received both TOP1 and TOP2 inhibitors, either at the same time (n=3) or switching from one to the other during the course of chemotherapy (n=40). 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. Treatment was unknown for approximately 25% of the patients. Almost 90% of patients with known treatment received chemotherapy, among whom most received a TOP2 inhibitor. 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 risk haplotype). 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).

Based on SNP functional prediction (http://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html) (25), rs942190 may affect *TDP1* expression by residing in a transcription factor binding site (TFBS). We further investigated which transcription factors (TF) bind to this region using data from the Encyclopedia of DNA Elements (ENCODE). The ENCODE project has performed a large number of ChIP-seq experiments on multiple cell lines to identify TFBSs across the human genome (26–28). Table 4 shows the list of 21 TFs identified by ENCODE that could bind to the TFBS in the location that rs942190 resides. In addition, we used a freely accessible web-based application called ePOSSUM (http://mutationtaster.charite.de/ePOSSUM/) by Hombach *et al.* (29) to assess the impact of the T(A) and C(G) alleles of rs942190 on TF binding. ePOSSUM allows user to enter either the genomic position of the SNP (based on human genome assembly GRCh37) or the wild-type and variant sequences. The output shows predicted TF binding scores of 81 TFs from three sources (JASPAR, HT-SELEX, and hPDI) as well as a summary of prediction whether the genetic alteration leads to the gain or loss of TF binding. Of the 21 TFs identified by ENCODE, nine have data available on the ePOSSUM website. Using rs942190 location (Chr14:90422414T>C) as an input on the website, six of the nine TFs were predicted to have a different binding affinity to the T(A) and C(G) allele. These included CTCF, MAX, MYBL2, RBBP5, TFAP2A, and TFAP2C (Table 4).

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 including the results from GTEx, ENCODE, and ePOSSUM. (25)It has been shown that overexpression of TDP1 in cell lines could counteract the effect of DNA topoisomerase inhibitors (30); 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. However, this observation was based on healthy tissue; the effect of rs942190 GG genotype may be different in tumor tissue. No known studies have compared *TDP1* expression in tumor versus adjacent non-tumor tissue from SCLC patients; however, increased *TDP1* expression has been found in tumor tissue relative to adjacent non-tumor tissue from NSCLC patients (31,32). 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.

The difference in *TDP1* expression among different rs942190 genotypes may be the result of differences in TF binding affinity. This SNP is located in the TFBS where several TFs bind (confirmed by cell line experiments from ENCODE project). At least six of them were predicted (based on in silico analysis) to have different binding affinity between T(A) and C(G) alleles of rs942190. Although these six TFs are mostly well known, the effect of these TFs specifically on *TDP1* gene expression has not been reported. CTCF could function as an enhancer or repressor (33); thus the attenuation of binding may result in either increasing or decreasing transcription. MAX could also be either an enhancer (forming heterodimers with MYC, MYC-MAX) or a repressor (forming heterodimers with MAD, MAD-MAX or homodimers, MAX-MAX) (34–36). Overexpression of *MYBL2* has been found in several cancer types and associated with poor patient outcomes (review in (37)). Moreover, studies in cell lines suggest that overexpression of *MYBL2* is associated with resistance to chemotherapeutic agents (including etoposide) and radiation (38–40). It is plausible that one mechanism of resistance to topoisomerase inhibitor or radiation is through activation of *TDP1* expression by MYBL2. Patients with the rs942190 GG genotype could have higher MYBL2 binding affinity, thus having higher *TDP1* expression that causes their tumors to be relatively more resistant to the treatment. Further study is needed to investigate this possibility. The protein encoded by *TFAP2A* and *TFAP2C* (AP-2α and AP-2γ) could activate or repress transcription of their target genes (41,42). Although the effect of AP-2α and AP-2γ on *TDP1* expression has not been reported, there was an evidence showing that breast cancer cell lines expressing *TFAP2A* and *TFAP2C* were more resistant to a topoisomerase inhibitor and radiation than constructed breast cancer cells depleting in AP-2 function (43).

The majority of the data used in this analysis came from etiologic studies of lung cancer, and so the data on treatments received by patients often were limited. The treatment method and the name of chemotherapeutic agents used were not available for approximately 30% of patients in this study. The majority of patients with unavailable treatment data were from the Harvard cohort. However, the chemotherapy regimen most commonly used in initial treatment of SCLC at Harvard is etoposide (TOP2 inhibitor) plus cisplatin or carboplatin, a regimen similar to that of other institutions. Thus, we would expect that the majority of patients with unknown treatment would have received similar treatment to the rest of patients. Based on the available data, we explored whether the association of rs942190 with survival differed between patients who received TOP1 and TOP2 inhibitors. We found a stronger association among patients who received TOP1 inhibitor (HR comparing GG vs. AA adjusting for age, sex, race, and tumor stage = 1.58; 95% CI: 0.87–2.87) compared to those receiving TOP2 inhibitor (aHR=0.99; 95% CI: 0.73–1.34). When we excluded patients who received both TOP1 and TOP2 inhibitors, the magnitude of association was stronger among patients receiving a TOP1 inhibitor (n=47, aHR = 1.92; 95% CI: 0.73–5.06). The adjusted HR for those receiving a TOP2 inhibitor without a TOP1 inhibitor (n=354) was 0.96 (95% CI: 0.69–1.33). However, since the sample size for patients who received a TOP1 inhibitor is quite small, and important data such as chemotherapy completion and response

to treatment were unavailable, we are not able to conclude that the association of rs942190 with survival differs among patients receiving different type of topoisomerase inhibitors.

In addition to repairing DNA damage produced by TOP1 and TOP2 inhibitors, an effect of TDP1 on DNA repair caused by radiation has been reported (11,44). Thus, we further explored the association of rs942190 genotype with overall survival among patients known to have received radiation (n=290), and found that patients with the GG genotype tended to have poorer survival compared to patients with the AA genotype (aHR=1.36; 95% CI: 0.95–1.97). The association was stronger among patients who received both a TOP1 inhibitor and radiation therapy (n=36, aHR=4.31; 95% CI: 1.1–16.80) but not for those who received a TOP2 inhibitor and radiation (n=221, aHR=1.18; 95%CI: 0.78–1.81). We did not observe an association with survival among those who did not receive radiation therapy (n=287, aHR comparing rs942190 GG to AA genotype = 1.09; 95% CI: 0.76–1.55).

In conclusion, our study suggests an association between rs942190 genotype and overall survival at 36 months after SCLC diagnosis. The association may be different among patients who received different treatment regimens, with respect to both chemotherapy and radiation. Further assessment of the genotype-survival association in a larger study with more detailed and complete treatment data is needed to confirm our findings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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 Society. 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 [PubMed: 25336398]
3. Kalemkerian GP. Small Cell Lung Cancer. Semin Respir Crit Care Med. 2016; 37(5):783–96. DOI: 10.1055/s-0036-1592116 [PubMed: 27732999]
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 [PubMed: 27732997]
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 [PubMed: 21565397]
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 [PubMed: 26351333]
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 [PubMed: 12042765]
8. Hande KR. Etoposide: four decades of development of a topoisomerase II inhibitor. Eur J Cancer. 1998; 34(10):1514–21. [PubMed: 9893622]
9. Pommier Y. Drugging topoisomerases: lessons and challenges. ACS Chem Biol. 2013; 8(1):82–95. DOI: 10.1021/cb300648v [PubMed: 23259582]
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 [PubMed: 16751265]
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 [PubMed: 22375014]
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. [PubMed: 18473723]
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. [PubMed: 20166932]
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 [PubMed: 25232464]
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 [PubMed: 23852793]

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 [PubMed: 18347181]
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 [PubMed: 19604089]
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. [PubMed: 14681826]
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. [PubMed: 22493747]
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. [PubMed: 23565320]
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 [PubMed: 17308338]
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 [PubMed: 23715323]
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 [PubMed: 19417063]
26. The ENCODE (ENCyclopedia Of DNA Elements) Project. Science. 2004; 306(5696):636–40. DOI: 10.1126/science.1105136 [PubMed: 15499007]
27. A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. 2011; 9(4):e1001046.doi: 10.1371/journal.pbio.1001046 [PubMed: 21526222]
28. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489(7414):57–74. DOI: 10.1038/nature11247 [PubMed: 22955616]
29. Hombach D, Schwarz JM, Robinson PN, Schuelke M, Seelow D. A systematic, large-scale comparison of transcription factor binding site models. BMC Genomics. 2016; 17:388.doi: 10.1186/s12864-016-2729-8 [PubMed: 27209209]
30. 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 [PubMed: 15494395]
31. Liu C, Zhou S, Begum S, Sidransky D, Westra WH, Brock M, et al. Increased expression and activity of repair genes TDP1 and XPF in non-small cell lung cancer. Lung Cancer. 2007; 55(3):303–11. [PubMed: 17118488]
32. Jakobsen AK, Lauridsen KL, Samuel EB, Proszek J, Knudsen BR, Hager H, et al. Correlation between topoisomerase I and tyrosyl-DNA phosphodiesterase 1 activities in non-small cell lung cancer tissue. Exp Mol Pathol. 2015; 99(1):56–64. DOI: 10.1016/j.yexmp.2015.05.006 [PubMed: 25987486]
33. Lu Y, Shan G, Xue J, Chen C, Zhang C. Defining the multivalent functions of CTCF from chromatin state and three-dimensional chromatin interactions. Nucleic Acids Res. 2016; 44(13):6200–12. DOI: 10.1093/nar/gkw249 [PubMed: 27067545]
34. Kretzner L, Blackwood EM, Eisenman RN. Transcriptional activities of the Myc and Max proteins in mammalian cells. Curr Top Microbiol Immunol. 1992; 182:435–43. [PubMed: 1490382]

35. Hurlin PJ, Ayer DE, Grandori C, Eisenman RN. The Max transcription factor network: involvement of Mad in differentiation and an approach to identification of target genes. Cold Spring Harb Symp Quant Biol. 1994; 59:109–16. [PubMed: 7587059]

36. Amati B, Land H. Myc-Max-Mad: a transcription factor network controlling cell cycle progression, differentiation and death. Curr Opin Genet Dev. 1994; 4(1):102–8. [PubMed: 8193530]

37. Musa J, Aynaud MM, Mirabeau O, Delattre O, Grunewald TG. MYBL2 (B-Myb): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. Cell Death Dis. 2017; 8(6):e2895.doi: 10.1038/cddis.2017.244 [PubMed: 28640249]

38. Grassilli E, Salomoni P, Perrotti D, Franceschi C, Calabretta B. Resistance to apoptosis in CTLL-2 cells overexpressing B-Myb is associated with B-Myb-dependent bcl-2 induction. Cancer Res. 1999; 59(10):2451–6. [PubMed: 10344757]

39. Levenson VV, Davidovich IA, Roninson IB. Pleiotropic resistance to DNA-interactive drugs is associated with increased expression of genes involved in DNA replication, repair, and stress response. Cancer Res. 2000; 60(18):5027–30. [PubMed: 11016623]

40. Ahlbory D, Appl H, Lang D, Klempnauer KH. Disruption of B-myb in DT40 cells reveals novel function for B-Myb in the response to DNA-damage. Oncogene. 2005; 24(48):7127–34. DOI: 10.1038/sj.onc.1208869 [PubMed: 16170378]

41. Hilger-Eversheim K, Moser M, Schorle H, Buettner R. Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control. Gene. 2000; 260(1–2):1–12. [PubMed: 11137286]

42. Eckert D, Buhl S, Weber S, Jager R, Schorle H. The AP-2 family of transcription factors. Genome Biol. 2005; 6(13):246.doi: 10.1186/gb-2005-6-13-246 [PubMed: 16420676]

43. Thewes V, Orso F, Jager R, Eckert D, Schafer S, Kirfel G, et al. Interference with activator protein-2 transcription factors leads to induction of apoptosis and an increase in chemo- and radiation-sensitivity in breast cancer cells. BMC Cancer. 2010; 10:192.doi: 10.1186/1471-2407-10-192 [PubMed: 20459791]

44. El-Khamisy SF, Hartsuiker E, Caldecott KW. TDP1 facilitates repair of ionizing radiation-induced DNA single-strand breaks. DNA Repair (Amst). 2007; 6(10):1485–95. DOI: 10.1016/j.dnarep.2007.04.015 [PubMed: 17600775]

Translational Relevance

Small cell lung cancer (SCLC) is the most aggressive form of lung cancer. Currently, there are very few markers to predict survival or to guide treatment selection for SCLC patients. *TDP1* gene plays a role in repairing DNA topoisomerases-mediated DNA damage and is believed to be responsible for drug resistance to DNA topoisomerase inhibitors (one of the common chemotherapeutic agents used for treating SCLC). To our knowledge, this is the first study to investigate germ line variation of *TDP1* in relation to survival among SCLC patients. We found rs942190 GG genotype to be associated with poor survival among 890 SCLC patients. If confirmed in a large study, *TDP1* rs942190 genotype may be used as a prognostic marker for patients with SCLC or a predictive marker for treatment response to DNA topoisomerase inhibitors.

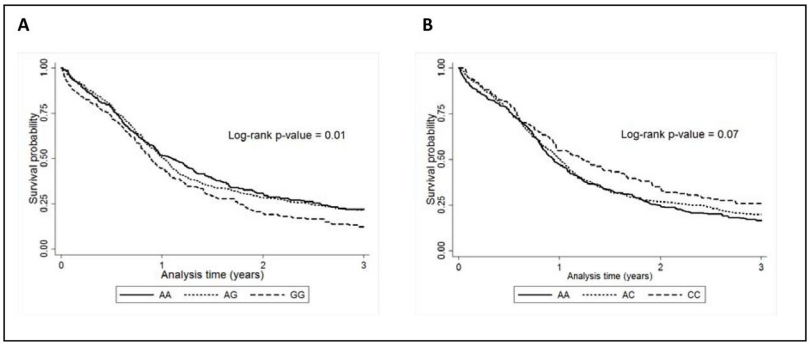

Figure 1.

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

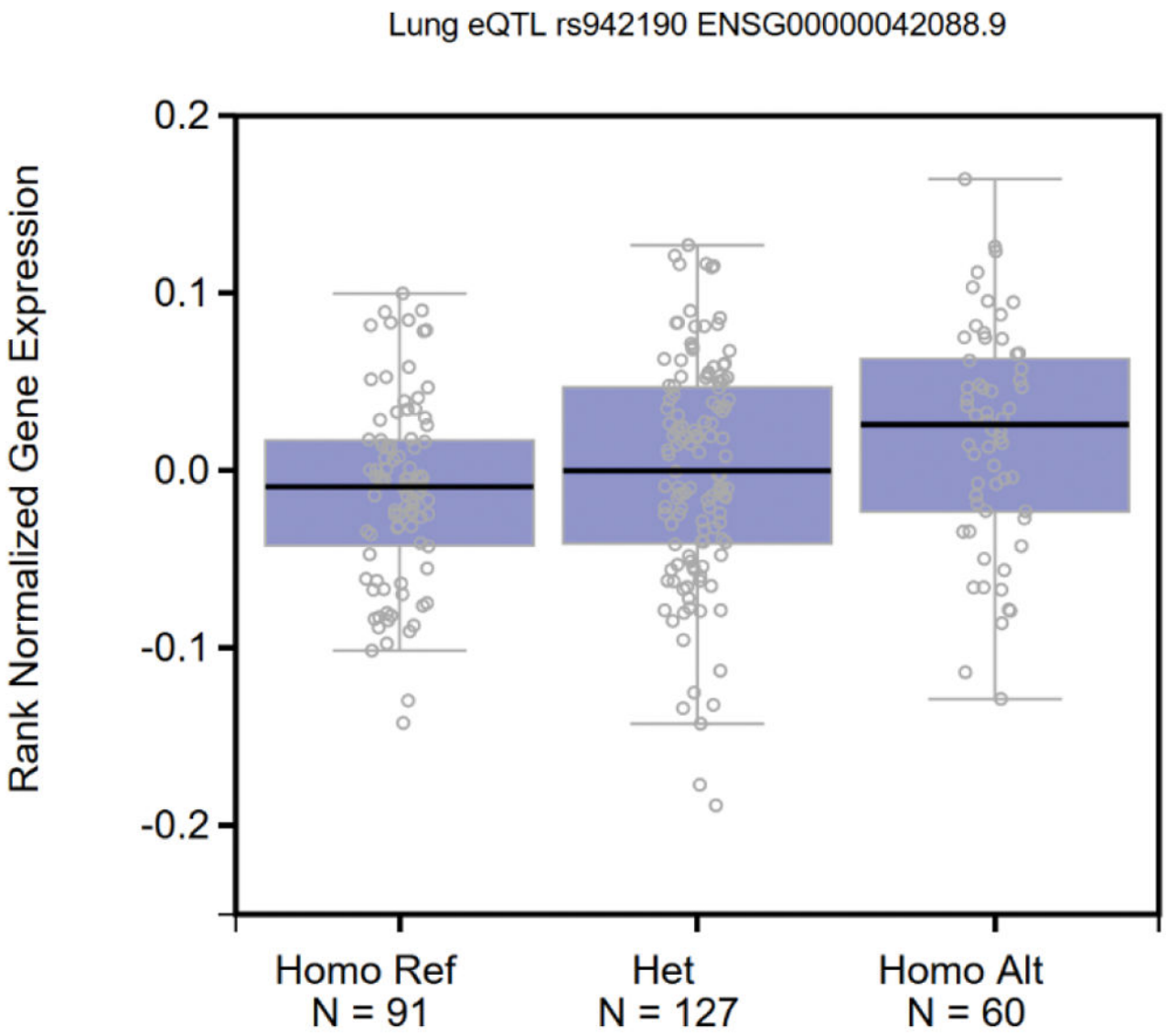

Figure 2.

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.

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

Clin Cancer Res. Author manuscript; available in PMC 2017 December 16.

Table 2

Selected characteristics of SCLC patients by genotype

		rs942190 (n=884)			rs2401863 (n=888)		
	Total (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
Chemotherapy							
Yes	598 (88.7%)	186 (88.2%)	265 (89.5%)	142 (87.7%)	228 (87.0%)	269 (90.6%)	99 (87.6%)
No	76 (11.3%)	25 (11.8%)	31 (10.5%)	20 (12.3%)	34 (13.0%)	28 (9.4%)	14 (12.4%)
Unknown	216	48	127	40	74	111	31

		rs942190 (n=884)			rs2401863 (n=888)		
	Total (n=890)	**AA (n=259)**	**AG (n=423)**	**GG (n=202)**	**AA (n=336)**	**AC (n=408)**	**CC (n=144)**
Chemotherapy Drug **							
TOP1 inhibitor	94 (15.7%)	35 (18.8%)	39 (14.7%)	17 (12.0%)	29 (12.7%)	44 (16.4%)	20 (20.2%)
TOP2 inhibitor	434 (72.6%)	130 (69.9%)	189 (71.3%)	113 (79.6%)	176 (77.2%)	190 (70.6%)	67 (67.7%)
Other/Unknown	113 (18.9%)	31 (16.7%)	58 (21.9%)	24 (16.9%)	40 (17.5%)	55 (20.4%)	18 (18.2%)
Radiation							
Yes	316 (47.7%)	90 (43.1%)	144 (50.3%)	80 (49.7%)	128 (49.0%)	141 (49.1%)	46 (41.1%)
No	346 (52.3%)	119 (56.9%)	142 (49.7%)	81 (52.1%)	133 (51.0%)	146 (50.9%)	66 (58.9%)
Unknown	228	50	137	41	75	121	32
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)

** The denominator for the percentage of this variable is the total number of patients who received chemotherapy. The counts presented are not mutually exclusive, since some patients received both TOP1 and TOP2 inhibitors.

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 (White** only)	GA	1.00	
	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)

Table 4

Transcription factors (TFs) identified by ENCODE that could bind to the transcription factor binding site at the location that rs942190 resides and the prediction of TF binding comparing between T and C allele of rs942190 based on ePOSSUM.

Transcription Factor	Summary prediction based on ePOSSUM
AP-2gamma	not available on ePOSSUM
ATF2	not available on ePOSSUM
CCNT2	not available on ePOSSUM
CHD1	not available on ePOSSUM
CTCF	attenuation of TF binding for C allele compared to T allele
E2F1	no definite result
E2F6	no definite result
ELF1_(SC-631)	not available on ePOSSUM
HA-E2F1	not available on ePOSSUM
HDAC1	not available on ePOSSUM
MAX	attenuation of TF binding for C allele compared to T allele
MYBL2	enhancement of TF binding for C allele compared to T allele
MYC	not available on ePOSSUM
PHF8	not available on ePOSSUM
Pol2	not available on ePOSSUM
Pol2-4H8	not available on ePOSSUM
POLR2A	not available on ePOSSUM
RBBP5	enhancement of TF binding for C allele compared to T allele
TCF3	no definite result
TFAP2A	enhancement of TF binding for C allele compared to T allele
TFAP2C	enhancement of TF binding for C allele compared to T allele