Published in final edited form as:
*Clin Cancer Res*. 2011 June 1; 17(11): 3830–3840. doi:10.1158/1078-0432.CCR-10-2877.

# Genetic Variations in Multiple Drug Action Pathways and Survival in Advanced-Stage Non-small Cell Lung Cancer Treated with Chemotherapy

**Yafei Li**[a,b], **Zhifu Sun**[c], **Julie M Cunningham**[d], **Marie C. Aubry**[e], **Jason A. Wampfler**[c], **Gary A. Croghan**[f], **Cassandra Johnson**[a], **Danli Wu**[a], **Jeremiah A. Aakre**[c], **Julian Molina**[f], **Liewei Wang**[g], **V. Shane Pankratz**[c], and **Ping Yang**[a,*]

[a]Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, MN, USA

[b]Department of Epidemiology, College of Preventive Medicine, Third Military Medical University, Chongqing, People's Republic of China

[c]Division of Biomedical Statistics and Informatics, Mayo Clinic, College of Medicine, Rochester, MN, USA

[d]Genomics Shared Resource, Mayo Clinic, College of Medicine, Rochester, MN, USA

[e]Department of Laboratory Medicine and Pathology, Mayo Clinic, College of Medicine, Rochester, MN, USA

[f]Division of Medical Oncology, Mayo Clinic, College of Medicine, Rochester, MN, USA

[g]Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, College of Medicine, Rochester, MN, USA

## Abstract

**Purpose—**Variations in genes related to anticancer drugs' biologic activity could influence treatment responses and lung cancer prognosis. Genetic variants in four biological pathways, i.e., glutathione metabolism, DNA repair, cell cycle, and EGFR, were systematically investigated to examine their association with survival in advanced-stage NSCLC treated with chemotherapy.

**Experimental Design—**A total of 894 tagging single-nucleotide polymorphisms (tagSNPs) in 70 genes from the four pathways were genotyped and analyzed in a 1076-patient cohort. Association with overall survival was analyzed at single-SNP and whole-gene levels within all patients and major chemotherapy agent combination groups.

**Results—**A poorer overall survival was observed in patients with genetic variations in *GSS* (glutathione pathway) and *MAP3K1* (EGFR pathway) (HR=1.45, 95% CI=1.20–1.70 and HR=1.25, 95% CI=1.05–1.50, respectively). In stratified analysis on patients receiving platinum

*Corresponding Author: Ping Yang, M.D., Ph.D., Mayo Clinic, Department of Health Sciences Research, 200 First Street SW, Rochester, MN 55905. Phone: (507) 266-5369, Fax: (507) 266-2478, yang.ping@mayo.edu.

**Contributions** Ping Yang led the study by designing and conducting it, interpreting results, writing the manuscript, and obtaining funding. Yafei Li coordinated the data analysis, results interpretation, and writing of the manuscript. Zhifu Sun searched the literature for the genes in the pathways, conducted tagSNP selection/custom panel design, and contributed to the final manuscript writing. V. Shane Pankratz led the statistical analyses and contributed to the final manuscript writing. Jason A. Wampfler and Jeremiah A. Aakre participated in the data analysis. Marie C. Aubry undertook pathology verification. Julie M. Cunningham led the genotyping of the Mayo Clinic samples. Julian Molina and Gary A. Croghan coordinated participant eligibility determination and enrollment. Danli Wu and Cassandra Johnson abstracted medical records data. Liewei Wang contributed to data collection and manuscript writing. All authors contributed to the final paper.

**Conflicts of interest** None.

plus taxane treatment, we observed a hazardous effect on overall survival by *MAP3K1* variant (HR=1.38, 95% CI =1.11–1.72) and a protective effect by *RAF1* (HR=0.64, 95% CI=0.5–0.82) in the EGFR pathway. In patients receiving platinum plus gemcitabine treatment, *RAF* and *GPX5* (glutathione pathway) genetic variations showed protective effects on survival (HR=0.54, 95% CI=0.38–0.77; HR=0.67, 95% CI=0.52–0.85, respectively); in contrast, *NRAS* (EGFR pathway) and *GPX7* (glutathione pathway) variations showed hazardous effects on overall survival (HR=1.91, 95% CI=1.30–2.80; HR=1.83, 95% CI=1.27–2.63, respectively). All genes that harbored these significant SNPs remained significant by whole-gene analysis.

**Conclusion—**Common genetic variations in genes of *EGFR* and glutathione pathways may be associated with overall survival among patients with advanced-stage NSCLC treated with platinum, taxane, and/or gemicitabine combinations.

### Keywords

non-small cell lung cancer; survival; single-nucleotide polymorphisms; pathway; chemotherapy

## Introduction

Non-small cell lung cancer (NSCLC) represents more than 80% of lung cancer diagnoses.(1) For advanced-stage (stage III and stage IV) NSCLC, 5-year survival varies widely (3–50%) depending on the number of lymph nodes involved, resectability, and tumor histology.(2) Established prognostic factors include stage, gender, age, and performance status.(3) Chemotherapy remains the major component of the standard care in conjunction with radiation therapy and supportive care for patients with advanced-stage lung cancer. For a given treatment, usually prescribed by a standard dosing protocol, survival time widely varies even after stratification by tumor stage, histology, and other clinical information, highlighting the need for improved predictive markers. In addition to clinical and biological behaviors of the tumor, multiple drug metabolism systems of the host also have an impact on the outcomes for lung cancer patients. Pharmacogenomic studies have defined gene variations responsible for varied drug efficacy. Patient-specific genetic profiles of genetic variants related to drugs' biological activity may help improve drug selection.

The glutathione metabolic pathway (GSH pathway) is directly involved in the detoxification or inactivation of multiple anticancer drugs, with evolving clinical relevance in lung cancer treatment.(4) The efficacy of anticancer drugs is highly influenced by cellular DNA repair capacity; increasing evidence shows reduced DNA repair capacity resulting from genetic polymorphisms of various DNA repair genes being associated with improved survival after platinum-based chemotherapy.(5–10) Taxanes (paclitaxel and docetaxel) are a class of anticancer agents, which exert their cytotoxic effects on spindle microtubules dynamics causing cell cycle arrest and apoptosis.(11, 12) These agents have been widely used as active chemotherapeutic agents in the treatment of NSCLC. Genes that regulate the cell cycle are thought to be involved in the biological activities of taxanes, although the impact of genetic variants in cell cycle pathway genes on clinic outcomes is unclear. The epidermal growth factor receptor (EGFR) is a tyrosine kinase transmembrane receptor in the ErB family of receptors expressed on the surface of epithelial cells, including NSCLC, regulating important processes in cell survival, cell cycle progression, tumor invasion, and angiogenesis.(13, 14) The EGFR pathway has been recognized as an important drug target. The EGFR tyrosine kinase inhibitors (TKIs), namely Gefitinib (IRESSA) and Erlotinib (Tarceva) are extensively studied in NSCLC and are currently incorporated into clinical practice. The EGFR pathway is associated with platinum-chemotherapy sensitivity in cancer patients in a recent study.(15)

A variety of chemotherapeutic drugs are used, and most advanced-stage NSCLC patients receive more than one drug; therefore, genetic variants of genes in multiple major drug action pathways may have combined effects on outcomes. However, most studies thus far have focused on limited candidate SNPs from a few selected genes. The influence of genetic variation in multiple pathways on NSCLC prognosis remains poorly understood. In this study, genetic variants of 70 key genes in four candidate pathways related to major chemotherapy drug actions, i.e., GSH, DNA repair, cell cycle, and EGFR pathways (Supplementary Figure 1), were systematically investigated in a prospectively followed cohort of advanced-stage NSCLC patients. Our goal was to examine whether SNP markers from the selected pathways are collectively associated with lung cancer overall survival among NSCLC patients receiving chemotherapy.

## Materials and Methods

### Patient Cohort

Lung cancer patients were identified and enrolled between 1997 and 2008. Detailed procedures of patient enrollment, diagnosis, data collection, and follow-up have been previously described.(16–18) Briefly, new cases diagnosed with lung cancer were identified by a daily electronic pathology reporting system. Once identified, patients were invited to participate and enrolled after their consent. The overall participation and blood sample donation rates were 87%. Demographic and tobacco history data were obtained from medical records and interview. When the patient received any therapy elsewhere, authorization for the release of medical information was requested, and copies of the relevant medical records were abstracted. Treatment information was obtained from oncology records, treatment data, outside medical records, and follow up questionnaires. Surgery was defined as a treatment when the patient had any pulmonary resection for the primary tumor, including pneumonectomy, bilobectomy, lobectomy, segmentectomy, and wedge resection. Chemotherapy agents, dosage, and date treatment started and ended were abstracted. Radiation was categorized as radiation to the chest only, radiation to another place with specific location the radiation occurred, gamma knife, and PCI; elapsed days for radiation, energy, radiation dose, times per day, and fractions were also abstracted. Clinical staging and recurrence or progression was determined by results from available chest radiography, computerized tomography, bone scans, position emission tomography scans, and magnetic resonance imaging. All patients were actively followed beginning within 6 months of diagnosis, with subsequent annual follow-up by mailed questionnaires. Annual verification of patients' vital status was accomplished through the use of Mayo Clinic's electronic medical notes and registration database, next-of-kin reports, death certificates, and obituary documents filed in the patients' medical records, as well as through the Mayo Clinic Tumor Registry and Social Security Death Index website. Well-trained abstractors conducted information abstraction and data entry. Research protocols were approved by the Mayo Clinic's Institutional Review Board.

### Pathway, gene and SNP selection

We focused on four common candidate pathways related to major chemotherapy drug actions, i.e., GSH, DNA repair for platinum based agents, cell cycle for taxanes, and EGFR pathways following a review of the literature in lung cancer therapy. Note that each of the four pathways may relate to multiple drug actions, particularly for individual genes; specifically, a gene can participate in multiple drug action pathways other than the one that was selected. Seventy genes were selected from four pathways following a review of the literature (Supplementary Table 1). TagSNPs (19, 20) were identified via SNPApp, a tagSNP selection program developed by the Bioinformatics Core at Mayo Clinic, Rochester, MN. SNPApp queries multiple public SNP data repositories (Hapmap, SeattleSNPs, and

National Institute for Environmental Health Science SNPs) that contain information on known SNPs, using ldSelect(21) to identify tagSNPs for the genes or regions of interest. SNPs within 5kb of each gene with a minor allele frequency (MAF) ≥ 0.05 for European populations were used as candidate SNPs, and tagSNPs were identified with a pair-wise linkage disequilibrium threshold of $r^2 \geq 0.8$. For a given gene, the SNP selection procedure used the SNP data repository with the greatest number of SNPs with a MAF ≥ 0.05 and the greatest number of linkage disequilibrium bins that met Illumina Golden Gate Assay quality score thresholds. Nonsynomyous SNPs were preferentially selected as tags when they were identified, provided that their metrics were equivalent to those of alternative tagSNPs.

## Genotyping and Quality control

A total of 1025 tagSNPs were genotyped in the Mayo Clinic Genomics Shared Resource using a custom-designed Illumina GoldenGate panel. Concordance among the three genomic control DNA samples present in duplicate was 100%. Subjects with a call rate was over 90% for all subjects; 111 SNPs failed genotyping. Of the SNPs with genotyping data, the call rate was >95%, and the minor allele frequency was >0.001 for all. SNPs with a Hardy-Weinberg equilibrium test p-value$<1\times10^{-7}$ (n=11) and/or monomorphic (n=9) were excluded, resulting in 894 SNPs in the analyses.

## Chemotherapy agent information

Chemotherapy regimens mostly fall into four groups (Supplementary Table 2): (1) platinum agents (P) including carboplatin and cisplatin, (2) taxane agents (T), paclitaxel and docetaxel, (3) gemcitabine (G), and (4) EGFR inhibitors (E), Gefitinib (IRESSA) and Erlotinib (Tarceva). Only a very small group of patients received drugs outside of these four groups; therefore, patients were divided into two subgroups based on the most commonly used combinations: platinum plus taxanes with or without other agents (PT group), platinum plus gemcitabine with or without other agents (PG group). We performed subsequent analyses on the significant SNPs in more restricted drug combinations: (1) platinum plus taxanes alone (PT-only group); (2) platinum plus any other agents except taxanes (P/no-T group); (3) platinum plus any other agents except gemcitabine (P/no-G group); (4) platinum plus taxanes without EGFR inhibitors (PT/no-E); (5) platinum plus gemcitabine without EGFR inhibitors (PG/no-E); and (6) EGFR inhibitors alone or plus any other agents (E group).

## Outcomes

Overall survival time was used as the primary endpoint, defined as the time from lung cancer diagnosis to either death or the last known date alive. Patients known to be alive were censored at the time of last contact.

## Statistical analysis

**Single-SNP association analysis—**A backward selection process was used to screen potential confounders: age at diagnosis, sex, race, smoking status, stage, histological types, comorbidity, and treatment modality. The significant variables were retained as covariates in all subsequent analyses. A Cox regression model was used to assess the associations between each SNP's genotypes and overall survival. The primary test of association was based on a genetic model free scheme (2 degrees of freedom). In subgroups, for SNPs where five or fewer minor allele homozygotes were observed, homozygote genotypes were combined with heterozygotes. If combined frequency was still five or fewer, then the SNP was removed. Hazard ratios (HRs) and 95% confidence intervals (CI) were estimated, comparing patients carrying one and two minor alleles individually to patients carrying two major alleles. To account for multiple comparisons in the SNP-based analysis, q-values set

at 0.20 were computed using the single-SNP p-values to quantify the probability that a p-value might be a false positive,(22) accepting a false discovery rate of 20%. The significant SNPs were further tested in refined treatment groups using the same methods outlined above. Kaplan-Meier curves and log-rank tests were used to assess the differences in survival time by individual SNPs.

**Principal Components Analysis (PCA) by whole-gene—**In order to assess whether different analytical approaches resulted in consistent findings, we performed whole-gene tests of association using a PCA approach. We utilized a genetic model free scheme that used two indicator variables for each SNP, one for the heterozygous and one for the homozygous minor allele carriers and extracted sufficient principal components to capture 90% of the SNP. The resulting collection of principal component variables was used for an omnibus test of significance for the association between each gene and survival in the multivariable Cox proportional hazards regression models. P-values for the global tests were obtained, along with summaries of the PCA.

All reported p-values were based on a two-sided test. All statistical analyses were performed using SAS, version 9 (SAS Institute, Inc.) or R-project software.

## Results

Table 1 provides descriptive characteristics of 1076 advanced-stage NSCLC patients. The median overall survival time was 1.7 years with 83.8% deceased within the 11-year follow-up period. Of the 1076 patients, 962 (89.4%) had chemotherapy, of which 657 (68.3%) were treated with PT, and 305 (31.7%) with PG (Supplementary Table 2). Eighty-eight percent of the patients had documented lung cancer progression or recurrence. In the PT group, gender and treatment modality were significant confounders; and in the PG group, gender, histological types, and treatment modality were significant confounders (Supplementary Table 3).

The analytic strategy is presented in Figure 1. Significant SNP associations in each treatment group are summarized in Table 2. A poorer overall survival was observed in all patients with minor alleles of rs17309872 in *GSS* from the GSH pathway and rs17661089 in *MAP3K1* from the EGFR pathway [HR=1.45, 95% CI=1.20–1.7 and HR=1.25, 95% CI=1.05–1.50, respectively] (Table 2 and Supplementary Figure 1). By gene-level tests, *GSS* and *MAP3K1* remained significant ($p=3.76\times10^{-2}$ and $1.80\times10^{-2}$, respectively).

In stratified analysis of the PT group, 4 SNPs, rs17661089, rs16886403, rs726501 (in the *MAP3K1* gene of the *EGFR* pathway), and rs11710163 (in *RAF1* of *EGFR* pathway) were significantly associated with overall survival. Patients with rare homozygote genotypes of these variants in *MAP3K1* had significantly poorer overall survival, while patients with heterozygotes at rs11710163 (in *RAF1*) had better overall survival compared to having common homozygotes (Table 2 and Supplementary Figure 3). By gene-level tests, *MAP3K1* and *RAF1* were the top two significant genes ($p=6.10\times10^{-3}$ and $3.14\times10^{-2}$, respectively) in the PT group, consistent with the single-SNP level analysis (Table 3). We further analyzed the four significant SNPs in the PT-only, P/no-T, P/no-E, and E groups, accordingly. All of the four SNPs were replicated in the P/no-E group; three SNPs in *MAP3K1* were replicated in the PT-only group; none showed significance in the P/no-T or E groups (Supplementary Table 4).

In the PG group, rs11710163 in *RAF1* was also found to be significant; patients with a heterozygote genotype showed a nearly doubled overall survival compared to patients with a common homozygote genotype (HR=0.54, 95% CI=0.38–0.77). The heterozygote genotype

at rs451774 in the *GPX5* gene also had a protective effect on overall survival, with the HR=0.61 (95% CI=0.47–0.79). In contrast, heterozygote genotypes at rs1065634 in *NRAS* and rs12118636 in *GPX7* showed a nearly doubled risk for death (HR=1.91; 95% CI=1.30–2.80 and HR=1.83; 95% CI=1.27–2.63, respectively) (Table 2 and Supplementary Figure 4). The genes, *RAF1*, *GPX5*, *GPX7*, and *NRAS*, which harbored the significant SNPs, were also confirmed to be significant (Table 3). We analyzed the four significant SNPs in the P/no-G, PG/no-E, and E groups, and all four SNPs were replicated in the PG/no-E group; however, none remained significant in the P/no-G or E groups (Supplementary Table 4).

## Discussion

In this study, we systematically investigated the associations between 894 SNPs in 70 key genes from four drug action pathways and overall survival in patient groups receiving different chemotherapy agent combinations. Genetic variations in *MAP3K1* in the EGFR pathway and *GSS* in the GSH pathway were associated with overall survival in the analysis of all patients. In the stratified analysis on patients receiving platinum plus taxane treatment, *MAP3K1* and *RAF1* variations in the EGFR pathway were associated with survival. In patients receiving platinum plus gemcitabine treatment, *RAF* and *GPX5* (GSH pathway), *NRAS* (EGFR pathway), and *GPX7* (GSH pathway) showed predictive effects on overall survival.

The glutathione synthetase gene (*GSS*) encodes the enzyme glutathione synthetase. Mutations in *GSS* prevent cells from making adequate levels of glutathione, leading to glutathione synthetase deficiency (23). Genetic variations in *GSS* were also found to be associated with overall survival of small cell lung cancer after treatment (24). It is interesting to observe *GSS* variants, instead of other commonly reported genetic variations, such as *GSTM1*, *GSTP1*, and *GSTT1*, as the most significant predictive markers in our systematic GSH pathway SNP analysis. It is proactive to observe that genetic variations in *MAP3K1* from the EGFR pathway are also significant predictive markers for survival in patients receiving the most commonly used doublet agents, cisplatin or carboplatin and paclitaxel or docetaxel, even after restricting down to patients who did not receive EGFR inhibitor agents.

The single-SNP and whole-gene analyses consistently indicated that genetic variations in *MAP3K1* and *RAF1* were significantly associated with overall survival in the PT group. Three SNPs, rs17661089, rs16886403, and rs726501 in *MAP3K1*, were significant. The associations of these SNPs in *MAP3K1* were further replicated in the PT-only group, suggesting that polymorphic variants in the *MAP3K1* were specifically associated with overall survival in patients receiving platinum and taxanes. All four SNPs were replicated in the P/non-E group, but not in the E group, suggesting the associations of these SNPs with overall survival may be independent of the effect of the EGFR inhibitors, although both *MAP3K1* and *RAF1* regulate the EGFR pathway.

*MAP3K1* has a pivotal role in a network of phosphorylating enzymes integrating cellular responses to a number of mitogenic and metabolic stimuli. *MAP3K1* phosphorylates and activates MAPK kinase (MAPK2), which in turn phosphorylates MAPK/ERK to produce downstream signaling effects on a variety of cancer genes (Supplementary Figure 1).(25, 26) Researchers have recently reported an association between the EGFR pathway and platinum-chemotherapy sensitivity in colorectal cancer.(15) Other genes in the EGFR pathway, such as *EGFR* itself, also have been demonstrated to be associated with lung cancer outcomes.(13, 14) For the three significant SNPs in *MAP3K1*, rs17661089 and rs16886403 are in relatively high linkage disequilibrium with rs726501 ($r^2$=0.64 and 0.72, respectively). rs17661089 is located in the flanking region of 5'-UTR, and the other two

SNPs are located in intron 1 of the gene. rs17661089 was predicted to be in the transcription factor binding sites (TFBS); this SNP may be a functional SNP or other variants in high linkage disequilibrium with it possibly influence the expression level of the gene product, requiring further functional analyses.

The RAF family is composed of three related serine/threonine protein kinases - RAF1, A-RAF and B-RAF— which act, in part, as downstream effectors of the RAS pathway.(27) Activated RAS interacts directly with the amino-terminal regulatory domain of the RAF kinase, which results in a cascade of reactions including direct activation of MEK.(28–30) Constitutively active mutated RAF can transform cells *in vitro*.(31) RAF may play a broader role in tumorigenesis and promotes the expression of the multi-drug resistance gene MDR1.(30) RAF1 holds an important role in cell growth, proliferation, and cell survival.(32, 33) The RAF1 protein has been found to be amplified in different lung cancer cell lines. One study showed that the RAF expression level is critical in tumor development.(34) Ravi *et al* reported that activated RAF1 causes growth arrest in human small cell lung cancer cells.(35) Immunohistochemical staining indicated that RAF1 was present in 49/53 ovarian adenocarcinomas and high c-RAF expression correlated significantly with poor survival in ovarian cancer.(36) The SNP, rs11706408, that has high linkage disequilibrium (r2=1) with rs11710163 in RAF1 was predicted to be in the TFBS; their functional roles in NSCLC chemotherapy agent actions and patient survival need to be further investigated.

Four SNPs in *RAF1*, *GPX5*, *GPX7*, and *NRAS* were found to be significantly associated overall survival in the PG group. All four SNPs were replicated in the PG/non-E group but were not replicated in the E group, indicating the association of being independent of the effect of the EGFR inhibitors. None of the four SNPs remained significant in the P/non-G group, suggesting the association of the variations in *RAF1*, *GPX5*, *GPX7*, and *NRAS* with overall survival in the PG group may reflect the effect of gemcitabine.

The GSH pathway has been related with detoxification or inactivation of platinum drugs. Glutathione peroxidases (GPXs) are a major antioxidative damage enzyme family that catalyzes the reduction of hydrogen peroxide, organic hydroperoxide, and lipid peroxides by reduced glutathione.(37–39) GPX has at least seven allozymes, distributed in different organs. In this study, *GPX5* rs451774 and *GPX7* rs12118636 were found to be associated overall survival in the PG group. The association of *GPX5* and *GPX7* were also confirmed by whole gene-based analysis. Function prediction indicated that rs451774 located at 3' UTR of *GPX5* was in both TFBS and miRNA binding sites, and is a potential functional SNP. *GPX5* has structural similarity to GPX1.(40) A recent study reported *GPX1* may be an inherited factor in predicting patients' quality of life.(41) In previous studies, variation in *GPX1* was linked tumor recurrence of bladder cancer.(42) *GPX7* shares a similar structural domain to *GPX4*. Reported findings suggest a potential role for *GPX7* in alleviating oxidative stress induced by dietary consumption of fatty acids in breast cancer cells.(43) Common variations in *GPX4* were associated with prognosis after a diagnosis of breast cancer.(44) However, no reports are currently available on the association of *GPX5* and *GPX7* variations and clinical outcomes. Our findings showed that genetic variations in *GPX5* tend to predict a better survival, while variations in *GPX7* were associated with worse survival. Current evidence also indicates the diversified roles of the individual *GPX*;(45) individual *GPX* might have different impact on the clinical outcome of lung cancer.

The SNP, rs1065634, is located at 5' near the gene region of *NRAS*. This SNP also was found to be significant in the PG group. *NRAS* is a member of the *RAS* proto-oncogene family. Point mutations of the *RAS* proto-oncogene family members are among the most frequent genetic alterations found in human cancer.(46, 47) Until now, *NRAS* gene mutations were mainly associated with hematopoietic malignancies(48), melanomas(49),

and bladder cancers.(50) rs1065634 was predicted to be in both TFBS and miRNA binding sites, and may be an important functional SNP; its potential functional significance is required to further study.

To our knowledge, our study presents the first and largest effort to comprehensively characterize the associations of gene alterations in four chemotherapy drug action pathways and survival in advanced-stage lung cancer patients. Platinum drugs, combined with other cytotoxic agents, continue to be the first-line chemotherapy for advanced-stage lung cancers and other metastatic cancers. Our study is strengthened by using different agent combinatorial groups for stratified analysis, which more closely mimics clinic practice than studies using single drug-based groups. We also used different analytic strategies and subgroups to confirm our results.

There are several limitations. First, despite efforts to account for potential confounding factors and careful adjustment for multiple tests in the analyses, false-positive associations cannot be completely ruled out. Second, although the significant SNPs were carefully analyzed in different combinatorial agent subgroups, chemotherapy heterogeneity may also modify the true association between genetic variation and clinical outcomes. Third, although our patient cohort represents one of the largest among all published combinatorial agent pharmacogenetics studies for late stage NSCLC, the power of this study to detect association was still modest, and it is possible that we missed some real associations with a small effect. For example, in the PG group, using a type I error level of $6\times10^{-5}$ (0.05/984), we had 86% power to detect effects from SNPs with a minor allele frequency of 0.3, which confers a per-allele HR of 1.6. If the per-allele HR is instead 1.4, the power is only 35%. Fourth, a thorough evaluation of the gemcitabine pathway SNPs was not included in this study. An additional limitation was the lack of patients treated with first-line EGFR inhibitor drugs, so the specific SNP effects could not be detected or evaluated relevant to these drugs. Our study focused on late-stage NSCLC patients who were treated with platinum drugs with or without other treatments, which will limit the ability to compare our results to other data sets. Nonetheless, our study is useful for testing a novel hypothesis, which may be confirmed in future specifically-designed clinical studies.

In summary, our results suggest that variations of *MAP3k1* and *RAF1* may be associated with overall survival among patients treated with platinum and taxane agents. Polymorphisms in the *RAF1*, *NRAS*, *GPX5*, and *GPX7* genes were found to be associated with overall survival among patients treated with platinum and gemcitabine agents. In the current clinical practice, oncologists usually select patients who they expect to benefit from chemotherapy based on available clinical information. Validated biomarkers, even with small effects, may have utility in optimizing patient therapies. Therefore, our findings could be considered in prospective biomarker-stratified clinical trials.

Statement of Translation Relevance

The prognosis for patients with advanced-stage non-small cell lung cancer (NSCLC) remains very poor. Chemotherapy remains the major component of the standard care for patients with advanced-stage lung cancer. This study comprehensively characterized the associations of gene alterations in four chemotherapy drug action pathways i.e., glutathione, DNA repair, cell cycle, and EGFR, and survival in advanced-stage NSCLC patients. Our results suggest that variations of *MAP3K1* and *RAF1* may be associated with overall survival among patients treated with platinum and taxane agents. Polymorphisms in the *RAF1*, *NRAS*, *GPX5*, and *GPX7* genes were found to be associated with overall survival among patients treated with platinum and gemcitabine agents. These

genetic variations may be applied to future prospective biomarker-stratified clinical trials, and help in designing patient-specific treatment and in predicting patients' survival.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Susan Ernst, M.A. for her technical assistance with the manuscript and David N. Rider (Division of Biomedical Statistics and Informatics, Mayo Clinic) for his assistance in SNP selection.

**Research Support:** This study was supported by U.S. National Institutes of Health (NIH) research grants, R01-CA80127 (to PY) and R01-CA84354 (to PY), and Mayo Foundation funds (to PY). Support was also provided by the NIH grant, P30 CA15083 (Mayo Clinic Comprehensive Cancer Center-Gene Analysis Shared Resource).

## References

1. Walker S. Updates in non-small cell lung cancer. Clin J Oncol Nurs. 2008; 12:587–96. [PubMed: 18676326]
2. Spira A, Ettinger DS. Multidisciplinary management of lung cancer. N Engl J Med. 2004; 350:379–92. [PubMed: 14736930]
3. Moore, DJ.; Lee, J. Lung Cancer: Principles and Practice. Pass, H.; Mitchell, J.; Johnson, D., et al., editors. Lipincott-Raven Publishers; Philadelphia: 1996. p. 495-509.
4. Yang P, Ebbert JO, Sun Z, Weinshilboum RM. Role of the glutathione metabolic pathway in lung cancer treatment and prognosis: a review. J Clin Oncol. 2006; 24:1761–9. [PubMed: 16603718]
5. Kalikaki A, Kanaki M, Vassalou H, Souglakos J, Voutsina A, Georgoulias V, et al. DNA repair gene polymorphisms predict favorable clinical outcome in advanced non-small-cell lung cancer. Clin Lung Cancer. 2009; 10:118–23. [PubMed: 19362955]
6. Gurubhagavatula S, Liu G, Park S, Zhou W, Su L, Wain JC, et al. XPD and XRCC1 genetic polymorphisms are prognostic factors in advanced non-small-cell lung cancer patients treated with platinum chemotherapy. J Clin Oncol. 2004; 22:2594–601. [PubMed: 15173214]
7. Ryu JS, Hong YC, Han HS, Lee JE, Kim S, Park YM, et al. Association between polymorphisms of ERCC1 and XPD and survival in non-small-cell lung cancer patients treated with cisplatin combination chemotherapy. Lung Cancer. 2004; 44:311–6. [PubMed: 15140544]
8. Isla D, Sarries C, Rosell R, Alonso G, Domine M, Taron M, et al. Single nucleotide polymorphisms and outcome in docetaxel-cisplatin-treated advanced non-small-cell lung cancer. Ann Oncol. 2004; 15:1194–203. [PubMed: 15277258]
9. Zhou W, Gurubhagavatula S, Liu G, Park S, Neuberg DS, Wain JC, et al. Excision repair cross-complementation group 1 polymorphism predicts overall survival in advanced non-small cell lung cancer patients treated with platinum-based chemotherapy. Clin Cancer Res. 2004; 10:4939–43. [PubMed: 15297394]
10. Wu X, Lu C, Ye Y, Chang J, Yang H, Lin J, et al. Germline genetic variations in drug action pathways predict clinical outcomes in advanced lung cancer treated with platinum-based chemotherapy. Pharmacogenet Genomics. 2008; 18:955–65. [PubMed: 18854777]
11. Gligorov J, Lotz JP. Preclinical pharmacology of the taxanes: implications of the differences. Oncologist. 2004; 9(Suppl 2):3–8. [PubMed: 15161985]
12. Zhao J, Kim JE, Reed E, Li QQ. Molecular mechanism of antitumor activity of taxanes in lung cancer (Review). Int J Oncol. 2005; 27:247–56. [PubMed: 15942666]
13. Nicholson RI, Gee JM, Harper ME. EGFR and cancer prognosis. Eur J Cancer. 2001; 37(Suppl 4):S9–15. [PubMed: 11597399]
14. Onn A, Correa AM, Gilcrease M, Isobe T, Massarelli E, Bucana CD, et al. Synchronous overexpression of epidermal growth factor receptor and HER2-neu protein is a predictor of poor

outcome in patients with stage I non-small cell lung cancer. Clin Cancer Res. 2004; 10:136–43. [PubMed: 14734462]

15. Zhang W, Stoehlmacher J, Park DJ, Yang D, Borchard E, Gil J, et al. Gene polymorphisms of epidermal growth factor receptor and its downstream effector, interleukin-8, predict oxaliplatin efficacy in patients with advanced colorectal cancer. Clin Colorectal Cancer. 2005; 5:124–31. [PubMed: 16098254]
16. Yang P, Bamlet WR, Sun Z, Ebbert JO, Aubry MC, Krowka MJ, et al. Alpha1-antitrypsin and neutrophil elastase imbalance and lung cancer risk. Chest. 2005; 128:445–52. [PubMed: 16002971]
17. Yang P, Sun Z, Krowka MJ, Aubry MC, Bamlet WR, Wampfler JA, et al. Alpha1-antitrypsin deficiency carriers, tobacco smoke, chronic obstructive pulmonary disease, and lung cancer risk. Arch Intern Med. 2008; 168:1097–103. [PubMed: 18504338]
18. Yang P, Wentzlaff KA, Katzmann JA, Marks RS, Allen MS, Lesnick TG, et al. Alpha1-antitrypsin deficiency allele carriers among lung cancer patients. Cancer Epidemiol Biomarkers Prev. 1999; 8:461–5. [PubMed: 10350443]
19. Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, et al. Haplotype tagging for the identification of common disease genes. Nat Genet. 2001; 29:233–7. [PubMed: 11586306]
20. 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; 74:106–20. [PubMed: 14681826]
21. Carlson CS, Eberle MA, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. Nature. 2004; 429:446–52. [PubMed: 15164069]
22. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A. 2003; 100:9440–5. [PubMed: 12883005]
23. Njalsson R, Ristoff E, Carlsson K, Winkler A, Larsson A, Norgren S. Genotype, enzyme activity, glutathione level, and clinical phenotype in patients with glutathione synthetase deficiency. Hum Genet. 2005; 116:384–9. [PubMed: 15717202]
24. Sun Z, Chen J, Aakre J, Marks RS, Garces YY, Jiang R, et al. Genetic variation in glutathione metabolism and DNA repair genes predicts survival of small-cell lung cancer patients. Ann Oncol. 2010; 21:2011–6. [PubMed: 20439344]
25. Witowsky JA, Johnson GL. Ubiquitylation of MEKK1 inhibits its phosphorylation of MKK1 and MKK4 and activation of the ERK1/2 and JNK pathways. J Biol Chem. 2003; 278:1403–6. [PubMed: 12456688]
26. Xu S, Cobb MH. MEKK1 binds directly to the c-Jun N-terminal kinases/stress-activated protein kinases. J Biol Chem. 1997; 272:32056–60. [PubMed: 9405400]
27. Kerkhoff E, Rapp UR. Cell cycle targets of Ras/Raf signalling. Oncogene. 1998; 17:1457–62. [PubMed: 9779991]
28. Beeram M, Patnaik A, Rowinsky EK. Regulation of c-Raf-1: therapeutic implications. Clin Adv Hematol Oncol. 2003; 1:476–81. [PubMed: 16258435]
29. Kolch W, Heidecker G, Kochs G, Hummel R, Vahidi H, Mischak H, et al. Protein kinase C alpha activates RAF-1 by direct phosphorylation. Nature. 1993; 364:249–52. [PubMed: 8321321]
30. Cornwell MM, Smith DE. A signal transduction pathway for activation of the mdr1 promoter involves the proto-oncogene c-raf kinase. J Biol Chem. 1993; 268:15347–50. [PubMed: 8101839]
31. Stanton VP Jr. Cooper GM. Activation of human raf transforming genes by deletion of normal amino-terminal coding sequences. Mol Cell Biol. 1987; 7:1171–9. [PubMed: 3561413]
32. Morrison DK, Cutler RE. The complexity of Raf-1 regulation. Curr Opin Cell Biol. 1997; 9:174–9. [PubMed: 9069260]
33. Kolch W. Meaningful relationships: the regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. Biochem J. 2000; 351(Pt 2):289–305. [PubMed: 11023813]
34. Kerkhoff E, Fedorov LM, Siefken R, Walter AO, Papadopoulos T, Rapp UR. Lung-targeted expression of the c-Raf-1 kinase in transgenic mice exposes a novel oncogenic character of the wild-type protein. Cell Growth Differ. 2000; 11:185–90. [PubMed: 10775035]

35. Ravi RK, Weber E, McMahon M, Williams JR, Baylin S, Mal A, et al. Activated Raf-1 causes growth arrest in human small cell lung cancer cells. J Clin Invest. 1998; 101:153–9. [PubMed: 9421477]
36. McPhillips F, Mullen P, Monia BP, Ritchie AA, Dorr FA, Smyth JF, et al. Association of c-Raf expression with survival and its targeting with antisense oligonucleotides in ovarian cancer. Br J Cancer. 2001; 85:1753–8. [PubMed: 11742498]
37. Halliwell B. Antioxidant defence mechanisms: from the beginning to the end (of the beginning). Free Radic Res. 1999; 31:261–72. [PubMed: 10517532]
38. Miyamoto Y, Koh YH, Park YS, Fujiwara N, Sakiyama H, Misonou Y, et al. Oxidative stress caused by inactivation of glutathione peroxidase and adaptive responses. Biol Chem. 2003; 384:567–74. [PubMed: 12751786]
39. Ray G, Husain SA. Oxidants, antioxidants and carcinogenesis. Indian J Exp Biol. 2002; 40:1213–32. [PubMed: 13677623]
40. Vernet P, Rock E, Mazur A, Rayssiguier Y, Dufaure JP, Drevet JR. Selenium-independent epididymis-restricted glutathione peroxidase 5 protein (GPX5) can back up failing Se-dependent GPXs in mice subjected to selenium deficiency. Mol Reprod Dev. 1999; 54:362–70. [PubMed: 10542376]
41. Yang P, Mandrekar SJ, Hillman SH, Allen Ziegler KL, Sun Z, Wampfler JA, et al. Evaluation of glutathione metabolic genes on outcomes in advanced non-small cell lung cancer patients after initial treatment with platinum-based chemotherapy: an NCCTG-97-24-51 based study. J Thorac Oncol. 2009; 4:479–85. [PubMed: 19347979]
42. Zhao H, Liang D, Grossman HB, Wu X. Glutathione peroxidase 1 gene polymorphism and risk of recurrence in patients with superficial bladder cancer. Urology. 2005; 66:769–74. [PubMed: 16230136]
43. Utomo A, Jiang X, Furuta S, Yun J, Levin DS, Wang YC, et al. Identification of a novel putative non-selenocysteine containing phospholipid hydroperoxide glutathione peroxidase (NPGPx) essential for alleviating oxidative stress generated from polyunsaturated fatty acids in breast cancer cells. J Biol Chem. 2004; 279:43522–9. [PubMed: 15294905]
44. Udler M, Maia AT, Cebrian A, Brown C, Greenberg D, Shah M, et al. Common germline genetic variation in antioxidant defense genes and survival after diagnosis of breast cancer. J Clin Oncol. 2007; 25:3015–23. [PubMed: 17634480]
45. Brigelius-Flohe R, Kipp A. Glutathione peroxidases in different stages of carcinogenesis. Biochim Biophys Acta. 2009; 1790:1555–68. [PubMed: 19289149]
46. Bos JL. ras oncogenes in human cancer: a review. Cancer Res. 1989; 49:4682–9. [PubMed: 2547513]
47. Barbacid M. ras genes. Annu Rev Biochem. 1987; 56:779–827. [PubMed: 3304147]
48. Bartram CR, Ludwig WD, Hiddemann W, Lyons J, Buschle M, Ritter J, et al. Acute myeloid leukemia: analysis of ras gene mutations and clonality defined by polymorphic X-linked loci. Leukemia. 1989; 3:247–56. [PubMed: 2564452]
49. Ball NJ, Yohn JJ, Morelli JG, Norris DA, Golitz LE, Hoeffler JP. Ras mutations in human melanoma: a marker of malignant progression. J Invest Dermatol. 1994; 102:285–90. [PubMed: 8120410]
50. Karimianpour N, Mousavi-Shafaei P, Ziaee AA, Akbari MT, Pourmand G, Abedi A, et al. Mutations of RAS gene family in specimens of bladder cancer. Urol J. 2008; 5:237–42. [PubMed: 19101897]

**Figure 1. Study design and key findings**

PT group, platinum plus taxanes with or without other drugs; PG group, platinum plus gemcitabine with or without other agents; PT-only group, platinum plus taxanes alone; P/no-T group, platinum plus any other agents except taxanes; PT/no-E group, both platinum and taxanes without EGFR inhibitors; E group, EGFR inhibitors alone or plus any other agents; P/no-G group, platinum plus any other agents except gemcitabine; PG/ no-E group, both platinum and gemcitabine without EGFR inhibitors.

## Table 1

Characteristics of the patients

| Patient characteristics | All patients n=1076 (%) | Patients by treatment groups | | |
|---|---|---|---|---|
| | | PT group[a] n=657 (%) | PG group[b] n=305 (%) | Other patients[c] n=133 (%) |
| **Alive/death** | 174/902 (16.2/83.8) | 108/549 (16.4/83.6) | 27/278 (8.9/91.2) | 23/110 (17.3/82.7) |
| **Survival time (years)** | | | | |
| Median (range) | 1.7 (0.1–11.0) | 1.8 (0.1–10.7) | 1.9 (0.2–10.1) | 1.4 (0.2–8.2) |
| Mean(SD) | 2.4(2.1) | 2.6(2.2) | 2.3(1.7) | 2.1(1.6) |
| **Age at diagnosis (years)** | | | | |
| Mean(SD) | 62.4(11.0) | 61.0(10.5) | 61.0(10.9) | 64.9(11.4) |
| ≤ 50 | 176(16.4) | 125(19.0) | 59(19.3) | 16(12.0) |
| 50–70 | 592(55.0) | 383(58.3) | 169(55.4) | 67(50.4) |
| ≥70 | 308(28.6) | 149(22.7) | 77(25.3) | 50(37.6) |
| **Men/women** (%) | 597/479 (55.5/44.5) | 360/297 (54.8/45.2) | 159/146 (52.1/47.9) | 74/59 (55.6/44.4) |
| **Race** | | | | |
| Non-Caucasian | 57(5.3) | 37(5.6) | 14(4.6) | 6(4.5) |
| Caucasian | 1019(94.7) | 620(94.4) | 291(95.4) | 127(95.5) |
| **Smoking history at diagnosis (%)** | | | | |
| Never | 197(18.3) | 133(20.2) | 79(25.9) | 24(18.1) |
| Former | 542(50.4) | 314(47.8) | 142(46.6) | 69(51.9) |
| Current | 337(31.3) | 210(32.0) | 84(27.5) | 40(30.1) |
| **Histological type** | | | | |
| Adenocarcinoma/BAC[d] | 637(59.2) | 387(58.9) | 193(63.3) | 81(60.9) |
| Squamous | 197(18.3) | 118(18.0) | 43(14.1) | 22(16.5) |
| Other and unclassified NSCLC[e] | 242(22.5) | 152(23.1) | 69(22.6) | 30(22.6) |
| **Stage** | | | | |
| IIIA | 304(28.3) | 182(27.7) | 60(19.7) | 29(21.8) |
| IIIB | 263(24.4) | 160(24.4) | 81(26.6) | 29(21.8) |
| IV | 509(47.3) | 315(48.0) | 164(53.8) | 75(56.4) |
| **Comorbidity diseases** | | | | |
| Pulmonary diseases (%) | | | | |
| Yes | 542(50.4) | 322(49.0) | 123(40.3) | 67(50.4) |
| No | 534(49.6) | 335(51.0) | 182(59.7) | 66(49.6) |
| Non-pulmonary diseases (%) | | | | |
| Yes | 692(64.3) | 434(66.1) | 187(61.3) | 72(54.1) |
| No | 384(35.7) | 223(33.9) | 118(38.7) | 61(45.9) |
| Other cancers[f] (%) | | | | |
| Yes | 292(27.1) | 172(26.2) | 72(23.6) | 35(26.3) |
| No | 784(72.9) | 485(73.8) | 233(76.4) | 98(73.7) |
| **Lung cancer recurrence/ progression/new primary** | | | | |

| Patient characteristics | All patients n=1076 (%) | Patients by treatment groups | | |
|---|---|---|---|---|
| | | PT group[a] n=657 (%) | PG group[b] n=305 (%) | Other patients[c] n=133 (%) |
| Yes (%) | 947(88.0) | 581(88.4) | 290(95.1) | 113(85.0) |
| No (%) | 129(12.0) | 76(11.6) | 15(4.9) | 20(15.0) |
| **Treatment modality** | | | | |
| Only chemotherapy (%) | 319(29.7) | 209(31.8) | 116(38.0) | 52(39.1) |
| Only radiation (%) | 36(3.4) | 0(0) | 0(0) | 0(0) |
| Both chemotherapy and radiation (%) | 358(33.3) | 250(38.1) | 118(38.7) | 49(36.8) |
| Both chemotherapy and surgery (%) | 86(8.0) | 62(9.4) | 21(6.9) | 11(8.3) |
| Both radiation and surgery (%) | 42(3.9) | 0(0) | 0(0) | 0(0) |
| Chemotherapy, radiation, and surgery (%) | 199(18.5) | 136(20.7) | 50(16.4) | 21(15.8) |
| Other treatment[g] | 36(3.4) | 0(0) | 0(0) | 0(0) |

[a] Platinum plus taxanes with or without other drugs.

[b] Platinum plus gemcitabine with or without other agents, not mutually exclusive from the PT group.

[c] Patients not receiving platinum, taxanes, and gemcitabine.

[d] Bronchioalveolar carcinoma.

[e] Non-small cell lung cancer.

[f] Excludes non-melanoma skin cancer.

[g] Includes all other treatment except for chemotherapy, radiation, and surgery to lung.

**Table 2**

Chemotherapy agent-relevant pathway gene polymorphisms and overall survival

| Pathway | Gene | SNP | Genotype | Alive/Death | MAF[a] | Median survival (years) | HR (95% CI)[b] | p-value[b] | q-value |
|---|---|---|---|---|---|---|---|---|---|
| **All patients (n=1076)** | | | | | | | | | |
| GSH | GSS | rs17309872 | | 174/902 | 0.06 | | | | |
| | | | TT | 161/783 | | 1.80 | Ref | | |
| | | | TA | 13/117 | | 1.46 | 1.44(1,19–1.76) | 2.39E-04 | |
| | | | TA+AA | | | | 1.45(1.20–1.77) | 1.47E-04 | |
| | | | Model free test[d] | | | | | 3.28E-04 | 1.50E-01 |
| EGFR | MAP3K1 | rs17661089 | | 174/902 | 0.08 | | | | |
| | | | AA | 151/750 | | 1.79 | Ref | | |
| | | | AG | 23/145 | | 1.55 | 1.22(1.02–1.46) | 2.92E-02 | |
| | | | GG | 0/7 | | 0.86 | 4.27(1.90–9.59) | 4.32E-04 | |
| | | | AG+GG | | | | 1.25(1.05–1.50) | 1.14E-02 | |
| | | | Model free test[d] | | | | | 2.59E-04 | 1.50E-01 |
| **PT group (n=657)** [c] | | | | | | | | | |
| EGFR | *MAP3K1* | rs17661089 | | 108/549 | 0.09 | | | | |
| | | | AA | 96/449 | | 1.92 | Ref | | |
| | | | AG | 12/94 | | 1.58 | 1.32(1.06, 1.65) | 1.46E-02 | |
| | | | GG | 0/6 | | 0.83 | 4.92(2.18, 11.12) | 1.24E-04 | |
| | | | AG+GG | | | | 1.38(1.11, 1.72) | 3.79E-03 | |
| | | | Model free test[d] | | | | | 5.24E-05 | 4.23E-02 |
| EGFR | *MAP3K1* | rs16886403 | | 108/549 | 0.11 | | | | |
| | | | AA | 94/422 | | 1.92 | Ref | | |
| | | | AG | 14/117 | | 1.68 | 1.31(1.06, 1.61) | 1.10E-02 | |
| | | | GG | 0/10 | | 0.99 | 2.62(1.39, 4.92) | 2.79E-03 | |
| | | | AG+GG | | | | 1.36(1.11, 1.66) | 2.61E-03 | |
| | | | Model free test[d] | | | | | 7.82E-04 | 1.58E-01 |

| Pathway | Gene | SNP | Genotype | Alive/Death | MAF[a] | Median survival (years) | HR (95% CI)[b] | p-value[b] | q-value |
|---|---|---|---|---|---|---|---|---|---|
| EGFR | *MAP3K1* | rs726501 | | 108/549 | 0.11 | | | | |
| | | | GG | 94/422 | | 1.92 | Ref | | |
| | | | GA | 14/117 | | 1.68 | 1.31(1.06, 1.61) | 1.10E-02 | |
| | | | AA | 0/10 | | 0.99 | 2.62(1.39, 4.92) | 2.79E-03 | |
| | | | GA+AA | | | | 1.36(1.11, 1.66) | 2.61E-03 | |
| | | | Model free test[d] | | | | | 7.82E-04 | 1.58E-01 |
| EGFR | *RAF1* | rs11710163 | | 108/549 | 0.08 | | | | |
| | | | AA | 86/472 | | 1.80 | Ref | | |
| | | | AG | 22/73 | | 2.18 | 0.62(0.49, 0.8) | 2.02E-04 | |
| | | | AG+GG | | | | 0.64(0.50, 0.82) | 3.78E-04 | |
| | | | Model free test[d] | | | | | 6.70E-04 | 1.58E-01 |
| **PG Group (n=305)** [e] | | | | | | | | | |
| EGFR | *RAF1* | rs11710163 | | 27/278 | 0.08 | | | | |
| | | | AA | 20/239 | | 1.83 | Ref | | |
| | | | AG | 6/38 | | 2.23 | 0.55(0.38, 0.79) | 1.09E-03 | |
| | | | AG+GG | | | | 0.54(0.38, 0.77) | 6.84E-04 | |
| | | | Model free test[d] | | | | | 6.84E-04 | 1.98-01 |
| EGFR | *NRAS* | rs1065634 | | 27/278 | 0.05 | | | | |
| | | | AA | 27/247 | | 1.92 | Ref | | |
| | | | AG | 0/30 | | 1.67 | 1.92(1.30–2.83) | 1.03E-03 | |
| | | | AG+GG | | | | 1.91(1.30, 2.80) | 1.00E-03 | |
| | | | Model free test[d] | | | | | 1.00E-03 | 1.98-01 |
| GSH | *GPX5* | rs451774 | | 27/278 | 0.29 | | | | |
| | | | AA | 9/148 | | 1.66 | Ref | | |
| | | | AG | 15/104 | | 2.27 | 0.61(0.47, 0.79) | 2.06E-04 | |
| | | | GG | 3/26 | | 2.15 | 0.99(0.65, 1.51) | 9.55E-01 | |
| | | | AG+GG | | | | 0.67(0.52, 0.85) | 1.06E-03 | |

| Pathway | Gene | SNP | Genotype | Alive/Death | MAF[a] | Median survival (years) | HR (95% CI)[b] | p-value[b] | q-value |
|---|---|---|---|---|---|---|---|---|---|
| | | | Model free test[d] | | | | | 7.25E-04 | 1.98-01 |
| GSH | *GPX7* | rs12118636 | | 27/278 | 0.06 | | | | |
| | | | GG | 27/241 | | 2.00 | Ref | | |
| | | | GA | 0/37 | | 1.28 | 1.83(1.27, 2.63) | 1.21E-03 | |
| | | | Model free test[d] | | | | | 1.21E-03 | 1.98-01 |

[a] Minor allele frequency.

[b] Hazard ratios (HRs), 95% confidence intervals (CIs), and p-values were calculated using COX regression analysis adjusted for gender, treatment modality, and stage for all patients; gender and treatment modality for the PT group; and gender, histological cell types, and treatment modality for the PG group.

[c] Platinum plus taxanes with or without other drugs.

[d] The association analysis was performed based on the genetic model free test (2 degrees of freedom).

[e] Platinum plus gemcitabine with or without other agents, not mutually exclusive from PT group.

## Table 3

Chemotherapy agent-relevant pathway gene and overall survival

| Pathway | Gene | Number of SNPs | Number of SNP captured 90% of the variability | p-value[a] |
|---|---|---|---|---|
| **All patients (n=1077)** | | | | |
| EGFR | NRAS | 14 | 10 | 6.09E-04 |
| EGFR | MAP3K1 | 26 | 10 | 1.80E-02 |
| GSH | GSTM1 | 2 | 2 | 2.75E-02 |
| GSH | GSS | 18 | 10 | 3.76E-02 |
| DNA | RAD52 | 30 | 10 | 4.63E-02 |
| **PT group[b] (n=657)** | | | | |
| EGFR | *MAP3K1* | 26 | 10 | 6.10E-03 |
| EGFR | *RAF1* | 22 | 9 | 3.14E-02 |
| GSH | *GPX5* | 12 | 8 | 3.18E-02 |
| DNA | *APEX1* | 10 | 8 | 3.20E-02 |
| GSH | *GSS* | 18 | 10 | 4.24E-02 |
| GSH | *GSTA5* | 14 | 9 | 4.34E-02 |
| **PG group[c] (n=305)** | | | | |
| GSH | *GPX6* | 18 | 8 | 8.33E-05 |
| GSH | *GPX5* | 12 | 8 | 2.22E-04 |
| EGFR | *NRAS* | 14 | 10 | 3.50E-03 |
| GSH | *GPX7* | 12 | 6 | 4.70E-03 |
| GSH | *GCLC* | 40 | 10 | 2.93E-02 |
| GSH | *GSTM3* | 10 | 5 | 4.09E-02 |
| DNA | *RAD52* | 30 | 10 | 4.38E-02 |
| EGFR | *RAF1* | 22 | 9 | 4.49E-02 |
| Taxa | *MAPK9* | 38 | 10 | 4.87E-02 |

[a] P-values for the global tests summarizing the outcome of the PCA were calculated using a multivariable Cox proportional hazards regression model based on the genetic model free test. We adjusted for gender, treatment modality, and stage for all patients; gender and treatment modality for the PT group; and gender, histological cell types, and treatment modality for the PG group.

[b] Platinum plus taxanes with or without other drugs.

[c] Platinum plus gemcitabine with or without other agents, not mutually exclusive from the PT group.