PLCE1 mRNA and protein expression and survival of patients with esophageal squamous cell carcinoma and gastric adenocarcinoma

Wen-Qing Li *1, Nan Hu *1, Victoria H Burton *1, Howard H Yang 3, Hua Su1, Catherine M Conway 2, Lemin Wang 1, Chaoyu Wang 1, Ti Ding 4, Yi Xu 4, Carol Giffen 5, Christian C Abnet1, Alisa M Goldstein1, Stephen M Hewitt 2, and Philip R Taylor1

1 Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institute of Health, Bethesda, MD 20852
2 Tissue Array Research Program, Laboratory of Pathology, National Cancer Institute, Bethesda, MD
3 Office of the Director, Center for Cancer Research, NCI, Bethesda, MD
4 Shanxi Cancer Hospital, Taiyuan, Shanxi, PR China
5 Information Management Services, Inc., Silver Spring, MD

* These authors contributed equally to this work.

Corresponding author: Philip R Taylor, MD, ScD, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, 9609 Medical Center Drive, Room 6E444, Rockville, MD 20890. E-mail: ptaylor@mail.nih.gov. Telephone: 240-276-7235. Fax: 240-276-7837.

Grant support: The current analysis was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, the Division of Cancer Epidemiology and Genetics (WQ Li, N Hu, VH Burton, H Su, L Wang, C Wang, CC Abnet, AM Goldstein, and PR Taylor), and the Center for Cancer Research (HH Yang, CM Conway, and SM Hewitt).
Running head: PLCE1 expression and UGI cancer survival
ABSTRACT

Background Germline genetic variants in PLCE1 (10q23) have demonstrated consistent associations with risk of esophageal squamous cell carcinoma (ESCC) and gastric cancer among Chinese. We evaluated PLCE1 mRNA and protein expression in paired tumor-normal tissues, and their relationship with survival.

Methods PLCE1 mRNA was profiled using three probes in the Affymetrix GeneChip U133 for paired tumor-normal tissues of ESCC (n=132), gastric cardia adenocarcinoma (GCA, n=62) and gastric noncardia adenocarcinoma (GNCA, n=72). We used immunohistochemistry to detect PLCE1 protein on slides from tissue microarrays in paired tumor-normal tissues of ESCC (n=303), and tumors of GCA (n=298) and GNCA (n=124).

Results Compared with normal tissues, PLCE1 mRNA expression was significantly reduced in ESCC tumors (P=0.03, probe_205112_at), as well as in GCA and GNCA tumors (P<0.0001, each probe). Protein expression was non-significantly reduced in ESCC tumors (P=0.51). Increased tumor-normal mRNA fold change (probe_205112_at) was associated with longer survival in ESCC (9.6 months for highest vs lowest quartile; P-trend=0.02). Increased mRNA tumor-normal fold change (probe_205111_at) was associated with longer survival for GCA (10.7 months for highest quartile; P-trend=0.04), but not for GNCA cases (P=0.72). Similar to mRNA, elevated tumor-normal fold change for protein in ESCC was also associated with improved survival (8.1 months for highest quartile; P-trend=0.04).

Conclusions Dysregulated PLCE1 mRNA expression was observed for both ESCC (one probe only) and GCA tumors, and the altered PLCE1 expression appears to be associated with cancer prognosis.
Impact A potential role for PLCE1 in the early detection and/or therapy of ESCC and GCA warrants further investigation.
INTRODUCTION

Esophageal cancer and gastric cancer (GC) represent the sixth and second leading cause of cancer-related deaths worldwide, respectively (1-2). Esophageal cancer has two primary histologic types, esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (3). GC is usually subdivided into gastric cardia cancer (GCA) and gastric non-cardia cancer (GNCA) by anatomic location. Worldwide, over half of all esophageal cancer and more than 40% of gastric cancer deaths occur in China where ESCC is the predominant form of esophageal cancer in high-risk areas (4-5). Molecular mechanisms underlying the carcinogenesis of ESCC and GC remain poorly defined. A handful of genetic loci associated with risk of ESCC or GC have been identified from genome-wide association studies (6-14). Among them, genetic variants at 10q23 in the phospholipase C epsilon (PLCE1) gene have demonstrated the most consistent and strongest associations with risk of ESCC for Chinese (6-8,12), and also conveyed significant associations with risk of GCA, surpassing genome-wide significance6, offering promise in the further exploration of molecular events of PLCE1 in the development and progression of these cancers.

*PLCE1* encodes a phospholipase C enzyme that mediates the hydrolysis of phosphatidylinositol-4,5-bisphosphate to generate the Ca2+-mobilizing second messenger inositol 1,4,5-triphosphate and the protein kinase C-activating second messenger diacylglycerol, which subsequently play fundamental roles in cell growth, differentiation, and gene expression (15-16). PLCE1 is activated by the Ras and Rho family GTPases and heterotrimeric G proteins, uniquely in the phospholipase C family (17). Rs932764 and rs3765524 in *PLCE1* have been associated with blood pressure and dengue shock syndrome (18-19), suggesting possible pleiotropic effects of *PLCE1* variants.
Despite the current knowledge of PLCE1 germline variants present in ESCC and GC, very little is understood about PLCE1 somatic changes and their association with prognosis. We examined mRNA and protein expression in tumor and matched normal tissues, as well as their association with major demographic and clinical features and survival of patients with ESCC, GCA, and GNCA from high risk Chinese populations (4-5).

Materials and Methods

Study subjects

All study subjects were cases of ESCC, GCA, and GNCA who presented to the Surgery Department of the Shanxi Cancer Hospital in Taiyuan, Shanxi Province, between 1996 and 2007, and resided in the Taihang Mountain area of north central China (5). All cases were born in Shanxi with Shanxi as their ancestral home, had no prior therapy for their cancer, and underwent surgical resection at hospitalization as their primary therapy. Information on demographic and lifestyle characteristics, family history of upper gastrointestinal (UGI) cancers, and clinical data were collected. Cases were histologically confirmed by pathologists at both the Shanxi Cancer Hospital and the National Cancer Institute (NCI). Cases were followed-up periodically for vital status; for those who died, date and cause of death were determined. This study was approved by the Institutional Review Boards of the Shanxi Cancer Hospital and the NCI, and all subjects provided written informed consent prior to participation.

Tissue collection and total RNA preparation

Tumor tissues and matched normal tissues distant to the tumor were obtained during surgery, and were snap frozen in liquid nitrogen and stored at -130°C. Selection of patients for RNA studies
was based on the availability of appropriate tissues for RNA testing (i.e., testing of cases with available frozen tissues, tumor samples that were predominantly (>50%) tumor, and tissue RNA quality/quantity adequate for testing). Total RNA was extracted using the Trizol method following the protocol of the manufacturer (http://tools.invitrogen.com/content/sfs/manuals/trizol_reagent.pdf), or using Allprep RNA/DNA/Protein mini kit from Qiagen, following the manufacturer’s instructions (http://www.qiagen.com/literature/render.aspx?id=2067). For both methods, the quality and quantity of total RNA were determined on the RNA 6000 Labchip/Agilent 2100 Bioanalyzer (Agilent Technology, Inc.). The mRNA extracted from the tumor and matched normal tissues of 132 ESCC cases, 62 GCA cases, and 72 GNCA cases were included in the present study.

### Probe preparation and hybridization for mRNA microarrays

The mRNA detection for ESCC samples was evaluated using the Affymetrix Human U133A, U133A_2, or U133Plus_2 chips. The mRNA expression for all GCA and GNCA samples was detected using U133Plus_2 chips. The U133A, U133A_2 and U133Plus_2 chips all contained the same three probes for PLCE1 (ie, 205111_s_at, 205112_at, and 214159_at). Probes were prepared according to the protocol provided by the manufacturer (Affymetrix GeneChip expression analysis technical manual, available from: http://www.affymetrix.com/support/index.affx). Procedures included first strand synthesis, second strand synthesis, double-strand cDNA cleanup, in vitro transcription, cRNA purification, and fragmentation. Twenty micrograms of biotinylated cRNA were finally applied to the hybridization array of Affymetrix GeneChip. After hybridization at 45°C overnight, arrays were developed with phycoerythrin-conjugated streptavidin by using a fluidics station (GeneChip
Fluidics Station 450) and scanned (GeneChip Scanner 3000) to obtain quantitative gene expression levels. Paired tumor and normal tissue specimens from each patient were processed simultaneously during the RNA extractions and hybridizations.

**Tissue collection and tissue microarray construction**

For histopathologic and protein expression analyses, we constructed tissue microarrays (TMAs), for which the detailed methodology has been described previously (20). In brief, tumor only or matched tumor-normal tissues from the same case were obtained during surgery, fixed in 70% ethanol (for GC cases, all tumor tissues were fixed in ethanol; for ESCC, 41 of 275 ESCC cases with paired tumor-normal tissues analyzed in relation to survival, and 249 of 489 cases with tumor tissue only analyzed in relation to survival were fixed in ethanol) or 10% neutral buffered formalin (all other ESCC tumor or normal tissues), and embedded in paraffin. In an initial pilot study, a minor difference in staining intensity and color of the counter-stain was noticed between ethanol- and formalin-fixed tissues. A scoring algorithm developed using formalin-fixed tissues only was subsequently adjusted and re-validated on ethanol-fixed specimens with the same pathologist as reference. Following this adjustment, both formalin- and ethanol-fixed tissues produced similar results for mean intensity for both positive and counter stains. The paraffin-embedded tissue samples were cored using 0.6 or 1.5mm diameter needles and arrayed into two sets of separate TMA recipient blocks along with positive and negative control tissues. 5µm sections were cut and transferred to adhesive coated slides (PSA-CS4x, Instrumedics Inc., Hackensack, NJ) to support the attachment of 0.6-mm array cores, while 1.5-mm array cores were transferred to glass slides via standard water bath (42°C) adhesion. TMAs stained with H&E were reviewed for confirmation of histopathologic diagnosis, and cores with insufficient
tissue determined by TMA H&E review were eliminated from analyses. Tumor tissues from a total of 525 ESCCs (303 with matched normal tissue, 222 with tumor only), 298 GCAs, and 124 GNCAAs were included for protein expression analysis. Matched normal tissues were not obtained for GCA and GNCA cases.

**Immunohistochemical analysis**

TMA slides were deparaffinized in xylene, rehydrated in graded alcohols and pretreated with pH 9 retrieval buffer (Dako Corp, Carpinteria, CA) in a pressure cooker for 20 minutes. The tissues were blocked with 0.3% hydrogen peroxide in 95% ethanol at 37°C for 10 min, and incubated with a rabbit polyclonal antibody against human PLCE1 in a 1:100 dilution (Sigma-Aldrich, St. Louis, MO), either at room temperature using the DAKO Autostainer staining system (Model LV-1) for 60 min, or overnight manually in a humidity chamber. The antigen-antibody complex was visualized with the DAKO Envision+ detection system and 3,3’-diaminobenzidine solution (Dako Corp, Carpinteria, CA), counterstained with hematoxylin, dehydrated in ethanol, and cleared in xylene.

**Tissue image analysis**

Stained TMA slides were digitized with the NanoZoomer 2.0-HT scanner (Hamamatsu, Bridgewater, NJ) and uploaded into SlidePath Digital Image Hub (Leica, Dublin, Ireland) in Nanozoomer Digital Pathology Image (NDPI) format. Digital Image Hub is a web-based image analysis system designed to employ high throughput automated analysis of virtual histology slides for image analysis. Representative regions of tumor or normal epithelia in each tissue core
were annotated to define the area within the image to be analyzed (Supplementary Figure S1). The PLCE1 image analysis algorithm was deployed within SlidePath’s Tissue Image Analysis system to measure staining intensity, as well as percent of pixels above a defined cut-off for each annotated core. We used concentration as a quantitative measurement of PLCE1 protein expression, which was derived formulaically based on the quantitative algorithm output of intensity and percent positive pixels.

**PLCE1 germline SNPs**

Genome-wide scanning was performed using the Illumina 660W array. The genotyping and quality control filtering metrics have been detailed in our published GWAS on UGI cancer(6). There were a total of 63 SNPs located within PLCE1 and its flanking area (20kb upstream and 10kb downstream). We extracted genotype data for five SNPs which showed significant associations with ESCC and GCA (rs2274223, rs3765524, rs3781264, rs11187842, and rs753724), exceeding genome-wide significance level ($P \leq 2.74 \times 10^{-8}$)(6). For this analysis, 98 ESCCs, 35 GCAs, and 52 GNCAs had information on the mRNA expression in normal tissues and germline SNP data. A total of 286 ESCCs had information on protein expression in normal tissues and SNP data.

**Statistical analysis**

Information on PLCE1 mRNA expression was available using each of three probes for paired tumor-normal tissues of ESCC, GCA, and GNCA cases, and analyses were performed using individual probes as well as overall mRNA expression (calculated as the average mRNA expression across all three probes) (GEO accession number of RNA expression array data is
GSE2340021 for ESCC (21) and GSE2927222 for GCA and GNCA (22)). Differential expression for a case tumor sample compared with its normal epithelial counterpart was denoted as a fold change. For one subject, we calculated the fold change using each probe as well as the overall fold change, defined as the average fold change across all three probes.

Information on PLCE1 protein expression in tumor tissue was available for ESCC, GCA, and GNCA cases; protein expression in matched tumor-normal tissue was available only for ESCC cases. The tumor-normal protein differential expression was also denoted as a fold change (ie, expression in tumor tissue divided by expression in matched normal tissue).

The Wilcoxon signed rank-sum test was used to test differences in PLCE1 mRNA and protein expression between matched tumor and normal tissues. PLCE1 mRNA and protein expression in tumor and tumor-normal expression fold changes were compared by categories of characteristics (age, sex, smoking (ever/never), alcohol intake (weekly/less than weekly), and family history of UGI cancer) and clinical features (Lauren classification (for GC), tumor grade, stage, and metastasis), using the Wilcoxon rank-sum test.

To evaluate the relation of PLCE1 mRNA and protein expression to survival, expression values (tumor tissue expression and/or tumor-normal fold changes, as available) were categorized into quartiles for each of the three cancers evaluated. Kaplan-Meir survival curves were plotted by quartiles of expression, and overall differences in survival across different expression categories were evaluated using log rank tests. Subsequently, Cox proportional hazards regression analyses were conducted to calculate the hazard ratios (HRs) and 95% confidence intervals (CIs) for expression categories. Analyses were conducted using the lowest quartile as the reference, with adjustment for age, sex, Lauren classification (for GCA and GNCA only), tumor grade, stage, and metastasis. These covariates were controlled because of
their potential effect on the association between PLCE1 expression and mortality. Since smoking and alcohol intake are not major contributing factors for UGI cancers in this and other high-risk populations(23), and family history of UGI cancer was not correlated with PLCE1 expression, these variables were not adjusted for in the primary analyses. Analyses further adjusting for smoking, alcohol intake, and family history of UGI cancer were very similar to our primary analyses, thus only the primary analyses are shown.

We conducted a stratified analysis by fixation methods for the association between PLCE1 protein expression and survival, and evaluated the $P$ for heterogeneity between subgroups using Q statistics.

Spearman correlation analyses were conducted to examine the association between $PLCE1$ germline genetic variants (6-8) and $PLCE1$ mRNA expression in normal tissues for ESCC and GCA separately and combined. We also conducted Spearman correlation analyses to evaluate the association between PLCE1 mRNA and protein expression in normal tissues.

Statistical analyses were performed using the SAS 9.2 (SAS Institute Inc, Cary, NC) and R language, and statistical significance was set at $P<0.05$.

**Results**

Altogether we included paired tumor-normal tissues from 132 ESCCs, 62 GCAs, and 72 GNCAs for detection of $PLCE1$ mRNA expression. For protein expression, we included tumor tissues of 525 ESCCs (303 with paired tumor-normal tissues, 222 with only tumor tissues), 298 GCAs, and 124 GNCAs (Table 1).

Compared with paired normal tissues, GCA and GNCA tumor tissues had significantly decreased $PLCE1$ mRNA expression for each of the three probes as well as for the average of the
three probes ($P<0.0001$ for all) (Figure 1). In GCA, the median tumor-normal expression fold change is 0.75 for probe 205111_s_at, 0.66 for 205112_at, 0.91 for 214159_at, and 0.76 for average of three probes. In GNCA, the median fold change is 0.85 for probe 205111_s_at, 0.80 for 205112_at, 0.90 for 214159_at, and 0.85 for average of three probes. The mRNA expression in ESCC tumor tissues also tended to be lower than matched normal tissues, but only expression measured using probe 205112_at was significantly different (median fold change=0.96, $P=0.03$). For PLCE1 protein expression in ESCC, we observed slightly reduced expression in tumor tissues compared to normal tissues, but the difference was not significant (median fold change=0.98, $P=0.51$) (Figure 1).

For ESCC, PLCE1 mRNA and protein expression, neither tumor-normal fold change nor tumor expression varied significantly by selected risk factors or tumor characteristics (Supplementary Tables S1 and S2). However, PLCE1 protein expression was significantly lower in diffuse compared to intestinal type for both GCA ($P<0.001$) and GNCA ($P=0.02$). $PLCE1$ mRNA tumor-normal fold change in GNCA was lower in cases who ever smoked and who had higher tumor stage. We also found a lower mRNA expression in tumors of GNCA cases that were younger.

In survival analyses, ESCC cases with higher average mRNA tumor expression or tumor-normal fold change had non-significantly decreased mortality (Table 2). Analyses by individual probes showed that higher mRNA tumor expression or tumor-normal fold change was associated with longer survival for all three probes (Supplementary Table S3), although only the tumor-normal fold change for probe 205112_at was significant in both the log rank test ($P=0.01$, Figure 2) and the multivariate-adjusted Cox regression model ($P$-trend=$0.02$, Table 2). The median survival times of ESCC cases within each quartile of fold change (probe 205112_at)
were 16.7, 34.8, 24.4, and 26.3 months, respectively (from the lowest to the highest quartile of mRNA expression). Similarly, in GCA, probe-specific Cox regression analyses showed that higher mRNA expression using probe 205111_at was significantly associated with longer survival ($P$-trend=0.04 for both tumor expression and tumor-normal fold change, Table 3). Median survival times of GCA cases by quartile of tumor expression (probe 205111_at) were 13.2, 16.2, 16.0, and 28.9 months, respectively. Median survival times of GCA cases by quartile of fold change (probe 205111_at) were 10.2, 23.9, 36.4, and 20.9 months, respectively. For GNCA, no associations between mRNA expression and survival were observed (Table 3).

We also evaluated PLCE1 protein expression and survival (Table 4). Increased tumor-normal protein expression fold change was associated with reduced mortality in ESCC cases ($P$-trend=0.04; median survival times for quartiles 1 to 4 were 26.6, 23.5, 42.0, and 34.7 months, respectively). No evidence for heterogeneity was seen for this association between ethanol-fixed and formalin-fixed tissues ($P$ for heterogeneity=0.65). Only protein expression in tumor tissues was available for GCA and GNCA, and no significant associations between expression and survival were seen.

We had information on both PLCE1 genotypes and mRNA expression in normal tissues for a subset of cases, including 98 ESCCs, 35 GCAs and 52 GNCAs (Table 1); information on both mRNA and protein expression in normal tissue was available only for four ESCCs. A joint analysis of all UGI cancers found significant associations between two SNPs (rs2274223 and rs3765524) and PLCE1 mRNA expression (probe 205111_s_at and 205112_at) ($P<0.05$). However, no significant correlations between mRNA expression and selected germline SNPs were seen in normal tissues for any of the UGI cancers studied here (Supplementary Table S4). We did, however, observe a correlation between PLCE1 mRNA using probe 205111_s_at and
protein expression in the normal tissues of the four ESCC cases (Spearman correlation $r=0.95$, $P=0.05$). Because we did not have protein expression in normal tissues for GCA and GNCA, we cannot evaluate the association between mRNA and protein expression in GC.

Discussion

In the present study, we analyzed PLCE1 mRNA and protein expression in ESCC, GCA, and GNCA cases, and related expression to risk factors, clinical factors, and survival. We found that PLCE1 mRNA expression was down-regulated in all three tumors compared to normal tissues (for all three probes in GCA and GNCA; for only one probe in ESCC), and that reduced tumor expression was associated with poorer survival in ESCC and GCA. To our knowledge, this is the first study to systematically investigate the association between PLCE1 expression and survival in a large number of UGI cancer patients.

Previous studies conducted among Chinese have reported differential PLCE1 mRNA and protein expression between ESCC tumors and normal tissues, but with conflicting results (24-26). A study in Shanghai observed decreased PLCE1 mRNA in ESCC tumors compared with matched normal esophagus (n=36), but no change in protein expression (n=39, including the 36 cases tested for mRNA) (24). In contrast, another study observed significantly higher PLCE1 mRNA and protein expression in ESCC tumors than normal tissues (n=50) in Kazakh Chinese(25). In a third study, Wang et al showed increased protein expression in both ESCC and GCA tissues compared to normal tissues (n=94 for ESCC, and n=13 for GCA)(7). This same group also recently reported that PLCE1 mRNA was significantly higher in ESCC tumors than in matched normal tissues (n=26)(26). In all of the above-mentioned studies, mRNA was assayed by real-time RT-PCR, while protein was assayed by immunohistochemistry. The only published
study focused on GC overall showed up-regulated PLCE1 mRNA (assayed by RT-PCR) and protein expression (assayed by Western blotting) in tumor tissues compared to matched normal tissues (n=74 for both mRNA and protein) (27).

We observed lower PLCE1 mRNA expression in ESCC tumors compared to normal tissues when using probe 205112_at, however, although PLCE1 protein expression in ESCC tumors also decreased, this difference was not significant. Our results are similar to the report from Shanghai (24), but contrast with studies from the other two groups (25-26). Our observations of reduced PLCE1 expression in GCA and GNCA are also discrepant from other reported studies (7,27). Differences in results among studies for PLCE1 expression might be due to differences in sample size, population heterogeneity, environmental exposures, or laboratory assays. For ESCC, our sample size is much larger than previous studies (n=133 pairs for mRNA, and 303 pairs for protein). Our cases all came from Shanxi Province in the high-risk northwestern part of the Taihang Mountain area, and both intrinsic and extrinsic differences may exist between our subjects and those in previous studies (25-26). For GC, there are only two similar publications, but both had much smaller sample size than ours and used different methods. One is a preliminary report on GCA by Wang et al (7), and the other made no distinction regarding GC anatomic subsites, which may have affected findings (27).

PLCE1 has been implicated in tumorigenesis in functional studies. Silencing PLCE1 has been shown to inhibit bladder tumor growth in vitro and in vivo (28). PLCE1 stimulation has been shown to induce pro-inflammatory cytokine expression and augment inflammation in response to chemical carcinogens or UV irradiation (17,29-30), which could promote skin cancer development. In another report, PLCE1 increased inflammation and angiogenesis and was involved in intestinal tumorigenesis (31). In contrast, other studies have shown downregulation
of PLCE1 in colorectal cancers (32-33). Given these varying results from prior reports, further studies are merited to clarify whether PLCE1 might display pleiotropic effects in multiple cancers, either in the same or opposite directions.

We observed decreased PLCE1 mRNA expression in GC tumors for all three probes; however, decreased mRNA expression in ESCC tumors was only significant for one probe. We did not find significantly altered PLCE1 protein expression in ESCC tumors either. Reduced PLCE1 expression appeared to be more consistent in GC tumors than ESCC tumors, which is partly consistent with the stronger genetic effect of PLCE1 SNPs in GC, particularly in GCA, as compared with ESCC (6,34). However, the association of PLCE1 SNPs with GNCA was not significant (6,34). The differentially altered PLCE1 expression in UGI cancers warrants further investigation.

We found a significantly inverse association between PLCE1 tumor-normal mRNA fold change in GNCA and tumor stage, as well as an inverse association between PLCE1 protein expression in GCA tumor tissues and tumor grade. Most importantly, we found longer survival (10 months or more) among ESCC and GCA cases who had higher PLCE1 mRNA expression, but no association with mortality for GNCA. Pending further biological investigations, these findings likely suggest a possible role of PLCE1 in suppressing progression and prognosis of UGI cancers, particularly for ESCC and GCA.

In contrast to the results in mRNA, we did not find significantly decreased PLCE1 protein expression in ESCC tumors, nor did we find a significant association between PLCE1 protein expression in GC tumors and survival. Inconsistencies between mRNA and protein results are not unexpected, given the complicated regulatory processes that occur between transcription and translation (eg, post-transcriptional, translational, as well as protein degradation.
regulation (35)). In addition, the mRNAs detected by the three probes on the Affymetrix chip might correspond to different physical locations in the gene than the targeted location of the protein assay.

Prior studies have evaluated PLCE1 germline SNPs in conjunction with mRNA and protein expression, with mixed results (24,26). In one study, the rs2274223 G allele (the risk allele from GWAS (6-8,12)) was associated with increased PLCE1 mRNA and protein expression in ESCC tumor tissues and cancer cell lines(26). However, in a second study, the same rs2274223 G allele (the risk allele) showed reduced gene expression (n=39) in ESCC (24). Our study did not find significant associations between rs2274223 or any of the four other PLCE1 germline SNPs and mRNA expression. Our study might be limited in the statistical power for this association. It is also likely that PLCE1 expression is regulated by multiple mechanisms while most are unknown yet. For example, it would be useful to determine if epigenetic changes are involved in PLCE1 expression. It also remains to be seen whether SNPs in other genes regulate PLCE1 expression through cis- or trans-effects. As such, although we observed significant associations between PLCE1 expression and survival for ESCC and GCA, which reflected the combined effects of all potential mechanisms, expression did not vary across PLCE1 germline genotypes.

Our study was strengthened by its comprehensive examination of PLCE1 expression in paired tumor-normal tissues, detailed information on demographic and clinical characteristics, and evaluation of a large number of cases, especially for protein expression. Evaluation of the relation of expression to survival was enabled by prospective follow-up for vital status on nearly all cases. We acknowledge limitations of our study. The sample size exceeded most prior studies, but was still only modest, particularly for GCA and GNCA. Only 303 of 525 ESCC and no GC
tumors had paired normal tissues. The number of subjects with complete information on PLCE1 germline SNPs, mRNA, and protein was limited, partly due to the different tissue preservation and processing methods required for mRNA and protein assay. We also did not have normal tissues from GCA and GNCA for protein analysis. We lacked information on Helicobacter pylori (H. pylori) infection, however, since H. pylori positivity is nearly universal in the Taihang Mountain area (36), it is unlikely that our results on GC were distorted by unadjusted confounding of H. pylori infection.

In conclusion, our study provides evidence that PLCE1 mRNA expression is dysregulated in GC and ESCC (one probe only), and that ESCC and GCA cases with higher PLCE1 expression experience more favorable prognosis.
References


<table>
<thead>
<tr>
<th></th>
<th>Cases with tumor tissues</th>
<th>Cases with normal tissues</th>
<th>Cases with both tumor and normal tissues</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total No.</td>
<td>No. with follow-up for survival</td>
<td>Total No.</td>
</tr>
<tr>
<td>mRNA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESCC</td>
<td>132</td>
<td>125</td>
<td>132</td>
</tr>
<tr>
<td>GCA</td>
<td>62</td>
<td>55</td>
<td>62</td>
</tr>
<tr>
<td>GNCA</td>
<td>72</td>
<td>71</td>
<td>72</td>
</tr>
<tr>
<td>Protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESCC</td>
<td>525</td>
<td>489</td>
<td>337</td>
</tr>
<tr>
<td>GCA</td>
<td>298</td>
<td>297</td>
<td>-</td>
</tr>
<tr>
<td>GNCA</td>
<td>124</td>
<td>124</td>
<td>-</td>
</tr>
</tbody>
</table>

ESCC: esophageal squamous cell carcinoma; GCA: gastric cardia cancer; GNCA: gastric non-cardia cancer

\(^a\) 33 ESCCs, 1 GCA, and 16 GNCA had information on both mRNA and protein expression in tumor tissue(s). 4 ESCCs and no GCA or GNCA had information on both mRNA and protein expression in normal tissues.
Table 2. Hazard ratios (95% CIs) for the association between PLCE1 mRNA expression and mortality in ESCCs (n=125)\(^a\)

<table>
<thead>
<tr>
<th>Quartile</th>
<th>Median</th>
<th>(\frac{n_{\text{Alive}}}{n_{\text{deceased}}})</th>
<th>HR(^b) (95% CI)</th>
<th>Median</th>
<th>(\frac{n_{\text{Alive}}}{n_{\text{deceased}}})</th>
<th>HR(^b) (95% CI)</th>
<th>Median</th>
<th>(\frac{n_{\text{Alive}}}{n_{\text{deceased}}})</th>
<th>HR(^b) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quartile 1</td>
<td>3.53</td>
<td>8/23</td>
<td>1.00</td>
<td>0.76</td>
<td>9/22</td>
<td>1.00</td>
<td>0.61</td>
<td>6/25</td>
<td>1.00</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>3.88</td>
<td>10/21</td>
<td>0.66 (0.36-1.24)</td>
<td>0.96</td>
<td>13/18</td>
<td>0.64 (0.34-1.22)</td>
<td>0.88</td>
<td>16/15</td>
<td>0.39 (0.20-0.75)</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>5.18</td>
<td>15/17</td>
<td>0.54 (0.28-1.02)</td>
<td>1.03</td>
<td>13/19</td>
<td>0.67 (0.36-1.26)</td>
<td>1.03</td>
<td>9/23</td>
<td>0.68 (0.38-1.21)</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>5.32</td>
<td>12/19</td>
<td>0.76 (0.41-1.41)</td>
<td>1.43</td>
<td>10/21</td>
<td>0.58 (0.30-1.10)</td>
<td>1.36</td>
<td>14/17</td>
<td>0.38 (0.20-0.73)</td>
</tr>
<tr>
<td>(P)-trend</td>
<td></td>
<td>0.30</td>
<td>0.13</td>
<td></td>
<td>0.02</td>
<td>(P)-trend</td>
<td></td>
<td>0.02</td>
<td>(P)-trend</td>
</tr>
</tbody>
</table>

CI: confidence interval; ESCC: esophageal squamous cell carcinoma; HR: hazard ratio

\(^a\) Quartile 1 has the lowest expression level. The overall mRNA expression across three probes or mRNA using each probe respectively was evaluated. For the mRNA expression using individual probes, we show only the significant association between expression and mortality, due to space limit. Overall tumor expression and overall tumor-normal expression fold change are the average of the expression using three probes. The differences between numbers here and numbers in Table 1 are due to incomplete survival data for all subjects. Medians are for the expression signals or fold-changes that do not have units.

\(^b\) HRs and 95% CIs were calculated using Cox regression, with adjustment for age, sex, tumor grade, stage, and metastasis.
Table 3. Hazard ratios (95% CIs) for the association between PLCE1 mRNA expression and mortality of GCAs (n=55) and GNCA (n=71)\textsuperscript{a}

<table>
<thead>
<tr>
<th>Overall expression in tumor (average of three probes)</th>
<th>Overall tumor-normal fold change (average of three probes)</th>
<th>Expression in tumor (probe 205111_s_at)</th>
<th>Tumor-normal fold change (probe 205111_s_at)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median (\frac{n_{\text{alive}}}{n_{\text{deceased}}})</td>
<td>Median (\frac{n_{\text{alive}}}{n_{\text{deceased}}})</td>
<td>Median (\frac{n_{\text{alive}}}{n_{\text{deceased}}})</td>
<td>Median (\frac{n_{\text{alive}}}{n_{\text{deceased}}})</td>
</tr>
<tr>
<td>GCA Quartile 1 5.24 5/8 1.00</td>
<td>0.62 2/11 1.00</td>
<td>6.10 2/1 1.00</td>
<td>0.46 0/13 1.00</td>
</tr>
<tr>
<td>GCA Quartile 2 5.42 2/12 1.32 (0.46-3.84)</td>
<td>0.72 4/10 0.46 (0.17-1.26)</td>
<td>6.31 3/11 0.64 (0.26-1.57)</td>
<td>0.65 3/11 0.48 (0.17-1.31)</td>
</tr>
<tr>
<td>GCA Quartile 3 5.53 2/12 1.90 (0.66-5.45)</td>
<td>0.84 2/12 0.56 (0.23-1.41)</td>
<td>6.48 3/11 0.42 (0.16-1.15)</td>
<td>0.81 5/9 0.15 (0.06-0.42)</td>
</tr>
<tr>
<td>GCA Quartile 4 5.83 2/12 1.14 (0.38-3.40)</td>
<td>1.09 3/11 0.80 (0.33-1.95)</td>
<td>6.78 3/11 0.38 (0.15-0.99)</td>
<td>1.19 3/11 0.60 (0.25-1.45)</td>
</tr>
<tr>
<td>P-trend 0.72</td>
<td>0.74</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>GNCA Quartile 1 5.67 6/11 1.00</td>
<td>0.64 5/12 1.00</td>
<td>6.16 6/11 1.00</td>
<td>0.53 4/13 1.00</td>
</tr>
<tr>
<td>GNCA Quartile 2 5.92 5/13 1.09 (0.44-2.68)</td>
<td>0.79 4/14 0.87 (0.35-2.12)</td>
<td>6.50 6/12 1.12 (0.47-2.67)</td>
<td>0.75 9/9 0.38 (0.15-0.95)</td>
</tr>
<tr>
<td>GNCA Quartile 3 6.07 5/13 1.26 (0.54-2.95)</td>
<td>0.88 7/11 0.91 (0.39-2.14)</td>
<td>6.75 3/15 1.68 (0.73-3.83)</td>
<td>0.90 2/16 1.17 (0.54-2.55)</td>
</tr>
<tr>
<td>GNCA Quartile 4 6.36 4/14 1.45 (0.59-3.54)</td>
<td>1.09 4/14 1.36 (0.56-3.32)</td>
<td>7.07 5/13 1.62 (0.66-3.96)</td>
<td>1.14 5/13 0.85 (0.37-1.95)</td>
</tr>
<tr>
<td>P-trend 0.35</td>
<td>0.46</td>
<td>0.16</td>
<td>0.72</td>
</tr>
</tbody>
</table>

CI: confidence interval; GCA: gastric cardia cancer; GNCA: gastric non-cardia cancer; HR: hazard ratio

\textsuperscript{a} Quartile 1 has the lowest expression level. The overall (average) mRNA expression of all three probes and mRNA for each probe respectively was evaluated. For the mRNA expression using individual probes, we show only the significant association between expression and mortality, due to space limit. Overall tumor expression is the average of the expression using three probes. The difference between numbers here and numbers in Table 1 was due to incomplete survival data for all subjects. Medians are for the expression signals or fold-changes that do not have units.

\textsuperscript{b} HRs and 95% CIs were calculated using Cox regression, with adjustment for age, sex, tumor grade, stage, metastasis, and Lauren classification.
Table 4. Hazard ratios (95% CI) for the association between PLCE1 protein expression and mortality of ESCCs (n=489 for tumor and n=275 for tumor/normal expression ratio), GCAs (n=297), and GNCA (n=124)\(^a\)

<table>
<thead>
<tr>
<th>Expression in tumor</th>
<th>Tumor/normal expression ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
</tr>
<tr>
<td>ESCC</td>
<td></td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.20</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>0.73</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>1.31</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>2.90</td>
</tr>
<tr>
<td>P-trend</td>
<td></td>
</tr>
<tr>
<td>GCA</td>
<td></td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.22</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>0.77</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>1.32</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>2.34</td>
</tr>
<tr>
<td>P-trend</td>
<td></td>
</tr>
<tr>
<td>GNCA</td>
<td></td>
</tr>
<tr>
<td>Quartile 1</td>
<td>0.14</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>0.77</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>1.50</td>
</tr>
<tr>
<td>Quartile 4</td>
<td>3.17</td>
</tr>
<tr>
<td>P-trend</td>
<td></td>
</tr>
</tbody>
</table>

CI: confidence interval; ESCC: esophageal squamous cell carcinoma; GCA: gastric cardia cancer; GNCA: gastric non-cardia cancer; HR: hazard ratio

\(^a\) Quartile 1 has the lowest expression level. HRs and 95% CIs were calculated using Cox regression, with adjustment for age, sex, tumor grade, stage, metastasis, and Lauren classification (for GCA and GNCA only). Medians are for the expression signals or fold-changes that do not have units.

\(^b\) HRs and 95% CIs were calculated using Cox regression, with adjustment for age, sex, tumor grade, stage, metastasis, and Lauren classification (for GCA and GNCA).
Figure legends:

Figure 1. PLCE1 mRNA and protein expression in paired tumor-normal tissues of esophageal squamous cell carcinoma (ESCC), gastric cardia cancer (GCA), and gastric non-cardia cancer (GNCA). The mRNA expression signal using the three probes (205111_s_at, 205112_at, and 214159_at) and the average signal of the three probes are shown for ESCC (1st row), GCA (2nd row), and GNCA (3rd row) respectively. We used the same scale for the y axis for each cancer. We had protein expression signal for both tumor and matched normal tissues for ESCC only (4th row). The difference in mRNA or protein expression signal between tumor and normal tissues was calculated using Wilcoxon signed rank-sum test.

Figure 2. Kaplan-Meier survival curve of esophageal squamous cell carcinoma patients by quartiles of PLCE1 mRNA tumor-normal fold changes evaluated by probe 205112_at. Quartile 1 has the lowest expression level. The HR (95% CI) for each quartile in Cox regression was shown in Table 2-a. P-value=0.01 in the overall log rank test across quartiles.
Figure 2

The graph shows the survival rates over follow-up years for different quartiles. The x-axis represents the follow-up years, while the y-axis represents the survival rate. Four quartiles are compared, with Quartile 1 showing the highest survival rate and Quartile 4 showing the lowest. The survival rates decrease over time for all quartiles.