The role of the CDCA gene family in ovarian cancer

Chongxiang Chen1,2#, Siliang Chen3#, Ma Luo4#, Honghong Yan2, Lanlan Pang5, Chaoyang Zhu5, Weiyan Tan5, Qingyu Zhao2, Jielan Lai6, Huan Li2

1Guangzhou Institute of Respiratory Diseases, State Key Laboratory of Respiratory Disease, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, China; 2Department of Intensive Care Unit, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China; 3Department of Hematology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China; 4Department of Interventional Radiology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China; 5Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510080, China; 6Department of Anesthesiology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong 510060, China

Contributions: (I) Conception and design: H Li, J Lai; (II) Administrative support: H Li; (III) Provision of study materials or patients: C Chen, S Chen; (IV) Collection and assembly of data: C Chen, S Chen, M Luo, H Yan, L Pang; (V) Data analysis and interpretation: H Li, J Lai, C Chen, S Chen, M Luo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

These authors contributed equally to this work.

Correspondence to: Huan Li. State Key Laboratory of Oncology in South China and Department of Intensive Care Unit, Sun Yat-sen University Cancer Center, 651 Dongfeng Dong Road, Guangzhou 510060, China. Email: lihuan@sysucc.org.cn; Jielan Lai. State Key Laboratory of Oncology in South China and Department of Anesthesiology, Sun Yat-sen University Cancer Center, 651 Dongfeng Dong Road, Guangzhou, Guangdong 510060, China. Email: laijl@sysucc.org.cn.

Background: Ovarian cancer is a frequently-occurring reproductive system malignancy in females, which leads to an annual of over 100 thousand deaths worldwide.

Methods: The electronic databases, including GEPIA, ONCOMINE, Metascape, and Kaplan-Meier Plotter, were used to examine both survival and transcriptional data regarding the cell division cycle associated (CDCA) gene family among ovarian cancer patients.

Results: All CDCA genes expression levels were up-regulated in ovarian cancer tissues relative to those in non-carcinoma ovarian counterparts. Besides, CDCA5/7 expression levels were related to the late tumor stage. In addition, the Kaplan-Meier Plotter database was employed to carry out survival analysis, which suggested that ovarian cancer patients with increased CDCA2/3/5/7 expression levels had poor overall survival (OS) (P<0.05). Moreover, ovarian cancer patients that had up-regulated mRNA expression levels of CDCA2/5/8 had markedly reduced progression-free survival (PFS) (P<0.05); and up-regulated CDCA4 expression showed remarkable association with reduced post-progression survival (PPS) (P<0.05). Additionally, the following processes were affected by CDCA genes alterations, including R-HAS-2500257: resolution of sister chromatid cohesion; GO:0051301: cell division; CORUM: 1118: Chromosomal passenger complex (CPC, including CDCA48, INCENP, AURKB and BIRC5); CORUM: 127: NDC80 kinetochore complex; M129: PID PLK1 pathway; and GO: 0007080: mitotic metaphase plate congression, all of which were subjected to marked regulation since the alterations affected CDCA genes.

Conclusions: Up-regulated CDCA gene expression in ovarian cancer tissues probably played a crucial part in the occurrence of ovarian cancer. The up-regulated CDCA2/3/5/7 expression levels were used as the potential prognostic markers to improve the poor ovarian cancer survival and prognostic accuracy. Moreover, CDCA genes probably exerted their functions in tumorigenesis through the PLK1 pathway.

Keywords: Ovarian cancer; cell division cycle associated (CDCA); prognosis

Submitted Oct 11, 2019. Accepted for publication Jan 03, 2020.
doi: 10.21037/atm.2020.01.99

View this article at: http://dx.doi.org/10.21037/atm.2020.01.99
Introduction

Ovarian cancer is a frequently-occurring reproductive system malignancy in females, which is estimated to cause an annual of over 100 thousand deaths worldwide (1). Disappointingly, more than 70% patients with ovarian cancer are diagnosed at the advanced stage due to the lack of efficient diagnostic method and early typical clinical symptom (2). Although treatments (such as surgery and targeted therapeutics) have been improved, they cannot achieve satisfactory progression-free survival (PFS) among patients with ovarian cancer; besides, the subsequent treatment for relapsed ovarian cancer is still encountered with great challenges (3). Additionally, the 5-year survival of ovarian cancer patients is only about 46.5% (4). Consequently, it is necessary to examine the underlying mechanisms regarding ovarian cancer tumorigenesis as well as progression and to identify the related tumor biomarkers that have high specificity and sensitivity. Nowadays, several molecular tests have been utilized prior to treatment for the risk screening of ovarian cancer, including BRCA1/2 mutations, microsatellite instability, and homologous recombination pathway genes. In addition, bevacizumab and Olaparib have been recommended by the National Comprehensive Cancer Network (NCCN) guideline for the treatment of ovarian cancer (5). All in all, BRCA1/2 mutations are not only used as the targets of agents like Olaparib (6), Rucaparib (7), and Niraparib (8), but also act as the high risk factors, which contribute to early screening (9-13). In patients with high risk factors (like BRCA mutation, family history), and cancer antigen 125 (CA-125), ultrasound is adopted to identify patients with ovarian cancer (14). More efforts should be made to search for more beneficial genes to predict cancer occurrence or targeted therapy. There are 8 members in the CDCA protein family, namely, CDCA1-8. Cell division plays an important role in the life process. It has been suggested in numerous studies that any dysregulation in the process of cell division may lead to malignancy (15-17). CDCA1 is known to be a member of a highly conserved Ndc80 complex that plays a crucial role in spindle checkpoint signaling (18). CDC42 plays a role in modulating response of DNA injury within cell cycle, which is achieved through binding onto protein phosphatase 1 γ (PP1γ) (19,20). CDCA3 functions to modulate the progression of cell cycle, and the expression level is regulated via protein degradation and transcription at G1 phase in cell cycle (21). Moreover, CDCA4 can regulate the cell cycle, which is related to transition of G1/S phase (22) and regulates the expression of p53 (23). CDCA5 serves as a primary regulatory factor for the sister-chromatid separation and cohesion (24). CDCA6, also named as CBX2, is a gene whose protein product forms part of the PRC1, a multi-protein complex that modifies histones (25). In the undifferentiated hematopoietic populations, CDCA7 can be triggered in the precursors of hematopoietic stem cells within murine embryo, which can be maintained afterwards. Besides, CDCA8 plays an essential role in regulating mitosis (26).

Nowadays, several studies on using some CDCA family genes as the prognostic factors have raised our attentions (27-29). However, there is little systematical analysis on the role of CDCA gene family in patients with ovarian cancer. The current research aimed to systematically evaluate the association of CDCA genes expression with ovarian cancer survival. Typically, the mRNA expression of CDCA genes was detected in both normal and ovarian cancer tissues. Then, the significance of all CDCA family members in predicting the prognosis for ovarian cancer was analyzed based on the Kaplan–Meier plotter database, and later the gene–gene interaction network was constructed for CDCA genes to examine the underlying mechanisms of action. This study explored the CDCA genes clinical value, so as to provide a certain theoretical foundation for making early diagnosis, prognosis evaluation, and specific treatment for ovarian cancer.

Methods

Each dataset used in this study was searched based on the published literature. The clinical tumor samples were collected during the first surgery, the normal specimens belonged to the same patients, and the threshold used to define low and high expression was 50% median. Additionally, the included literature datasets (TCGA datasets and GEO datasets) used for calculating Kaplan-Meier survival in Kaplan-Meier Plotter (www.Kmplot.com) are shown in Table S1.

ONCOMINE analyses

The transcription levels of CDCA genes among various cancer types were examined based on the online cancer microarray database, namely, the ONCOMINE gene expression array dataset (www.oncomine.org/). Moreover, CDCA5s mRNA expression was compared between the clinical tumor samples and normal specimens. The P value
was generated by Students’ t-test, and the threshold fold-change and P value were set at 2 and 0.01, respectively.

**The Gene Expression Profiling Interactive Analysis (GEPIA) dataset**

GEPIA, the newly designed interactive web server, was used to analyze RNA sequencing materials based on the GTEx and TCGA projects by the use of the normalized processing pipeline. Typically, GEPIA allows to offer the differential expression analyses on normal and tumor tissues, as well as the access to profiling of cancer type and pathological stage, analysis of patient survival, detection of similar gene, as well as dimensionality reduction and correlation analyses.

**Database Kaplan-Meier plotter**

Kaplan-Meier Plotter (www. Kmplot.com), the online database, was used to assess the prognostic significance of the mRNA expression levels of CDCA genes. For analyzing the ovarian cancer patient overall survival (OS), PFS, as well as the post progression survival (PPS), all specimens were divided as 2 groups according to the 50% median expression level (namely, low and high expression).

The baseline clinical data of included literature datasets in Kaplan-Meier Plotter for analyzing were collected, and then these data were included to perform bioinformatic Cox regression analysis for OS by R software (version 3.6.1). Later, the Kaplan-Meier survival plot was used for evaluation on the basis of hazard ratio (HR) and the corresponding 95% confidence intervals (CI), as well as the log-rank P value. The Kaplan-Meier plots were obtained through the CDCA’s JetSet best probe set alone, where the number at risk was suggested under the major plot.

**Bioinformatic analyses and functional enrichment**

The online database metascape (http://metascape.org) has integrated more than 40 bioinformatic knowledge bases, which enables to extract rich annotations, identify the enriched pathways, and construct the protein-protein interaction (PPI) (30) network based on the lists of protein and gene identifiers. The CDCA genes were analyzed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) approaches of Metascape, so as to search for those linked genes with the highest alteration frequency.

**Results**

Six CDCA factors are identified within the mammalian cells. In this study, the ONCOMINE databases were used to compare the CDCA gene transcriptional data between tumor tissues and normal specimens (Figure 1). According to our findings, CDCA2/3/4/5/7/8 expression remarkably increased among ovarian cancer patients derived from many databases. For CDCA2, the result from Yoshihara’s dataset showed that the fold change was 26.877 (31). TCGA dataset showed a fold change of 4.847 in CDCA3. In addition, TCGA dataset demonstrated that the fold change was 2.213 in CDCA4. For CDCA5, Yoshihara’s dataset revealed a fold change of 12.508 (31). For CDCA7, the results by Lu’s database suggested that the fold change was 2.646 (32), which were 8.261 as suggested by Yoshihara’s dataset (31). For CDCA8, Yoshihara's dataset showed a fold change of 20.335, which was 3.705 in TCGA dataset and 2.036 as suggested by the Bonome’s dataset (33) (Table 1).

**Associations of CDCAs mRNA expression with clinicopathological variables in ovarian cancer patients**

The GEPIA dataset (http://gepia.cancer-pku.cn/) was performed to compare CDCA’s mRNA expression in ovarian cancer tissues with that in normal ovarian tissues. According to our findings, the CDCA2/3/4/5/7/8 expression levels (tumor sample: n=426 vs. normal sample: n=88) were up-regulated within ovarian cancer tissues relative to those within normal tissues. (Figures 2, 3) Also, the association between CDCA genes expression and the ovarian cancer stage was analyzed. There were significant differences in the expression levels of all CDCA genes (Figure 4).

**Relationship between elevated CDCA 2/3/4/5/7/8 mRNA expression and dismal prognosis for ovarian cancer cases**

The Kaplan-Meier Plotter (www.Kmplot.com) was utilized to examine the relationship between CDCA’s mRNA expression and ovarian cancer patients’ survival by the use of the public datasets. The clinical baseline data of datasets included in Kaplan-Meier Plotter for analysis are displayed in Table S1. The multivariante analysis (using GSE 18920, GSE 26193, GSE 30161, GSE 63888) indicated that stage and CDCA7 were the independent risk factors (Table S2, Figure S1). Based on the long-rank test and Kaplan-Meier curve analyses, it was found that the elevated CDCA2/3/5/7...
Figure 1 CDCAs expression at the transcription level among various cancer types (the ONCOMINE). (Color red means high expression level in cancer sample, and color blue means low expression level in cancer sample).

Table 1 Remarkable CDCA expression changes at transcription level between ovarian cancer tissues and the non-carcinoma counterparts (Oncomine Database)

<table>
<thead>
<tr>
<th>CDCA genes</th>
<th>Type of ovarian cancer versus normal lung tissue</th>
<th>Fold change</th>
<th>P value</th>
<th>t-test</th>
<th>Source and/or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDCA2</td>
<td>Ovarian serous adenocarcinoma</td>
<td>26.877</td>
<td>1.07E-5</td>
<td>9.218</td>
<td>Yoshihara (31)</td>
</tr>
<tr>
<td>CDCA3</td>
<td>Ovarian serous cystadenocarcinoma</td>
<td>4.847</td>
<td>1.30E-9</td>
<td>21.227</td>
<td>TCGA</td>
</tr>
<tr>
<td>CDCA4</td>
<td>Ovarian serous cystadenocarcinoma</td>
<td>2.213</td>
<td>6.40E-6</td>
<td>9.679</td>
<td>TCGA</td>
</tr>
<tr>
<td>CDCA5</td>
<td>Ovarian serous adenocarcinoma</td>
<td>12.508</td>
<td>5.84E-13</td>
<td>15.754</td>
<td>Yoshihara (31)</td>
</tr>
<tr>
<td>CDCA7</td>
<td>Ovarian endometrioid adenocarcinoma</td>
<td>3.276</td>
<td>9.25E-5</td>
<td>6.009</td>
<td>Lu (32)</td>
</tr>
<tr>
<td></td>
<td>Ovarian serous adenocarcinoma</td>
<td>8.261</td>
<td>3.76E-6</td>
<td>6.923</td>
<td>Yoshihara (31)</td>
</tr>
<tr>
<td>CDCA8</td>
<td>Ovarian serous adenocarcinoma</td>
<td>20.335</td>
<td>3.63E-12</td>
<td>15.126</td>
<td>Yoshihara (31)</td>
</tr>
<tr>
<td></td>
<td>Ovarian serous cystadenocarcinoma</td>
<td>3.705</td>
<td>5.44E-7</td>
<td>13.126</td>
<td>TCGA</td>
</tr>
<tr>
<td></td>
<td>Ovarian carcinoma</td>
<td>2.036</td>
<td>6.46E-9</td>
<td>12.078</td>
<td>Bonome (33)</td>
</tr>
</tbody>
</table>
Figure 2 CDCAs expression within ovarian cancer (GEPIA).

Figure 3 CDCAs expression within ovarian cancer showed as boxplot (GEPIA). *P<0.05.
expression levels showed significant relationships with poorer patient OS. Ovarian cancer patients that had up-regulated CDCA2/5/8 mRNA expression levels had lower PFS, and the increased CDCA4 was only significantly associated with the lower PPS (Figure 5).

**Predicted functions and Pathway enrichment analyses and predicted functions of CDCA genes among ovarian cancer cases**

Genes showing co-expression with the CDCA gene family were examined using the String and the Functional protein association networks. According to the results, CDCA gene expression showed positive correlation with the up-regulated levels of the genes shown below: CDCA2, NU2, CDCA4, CDCA3, CDCA, CDCA5, CDCA8, CDCA7, CDCA20, AURKB, CBX2, CDK1, ZWINT, BUB1, NDC80, SPC24, SPC25, BIRC5, and INCENP (Figure 6). Then, the lists of all the CDCA genes expressed, together with linked genes displaying the highest alteration frequency, were compiled before they were analyzed by the KEGG and GO approaches in Metascape (Figure 7). Our results suggested that the processes below were subjected to the influence of CDCA gene alterations: R-HAS-2500257: resolution of sister chromatid cohesion; GO:0051301: cell division; CORUM: 1118: Chromosomal passenger complex (CPC, including CDCA8, INCENP, AURKB and BIRC5); CORUM: 127: NDC80 kinetochore complex; M129: PID PLK1 pathway; and GO: 0007080: mitotic metaphase plate congression.

**Discussion**

The guideline recommended treatments, namely, intravenous carboplatin/paclitaxel, or combined intravenous/intraperitoneal paclitaxel/cisplatin, may lead to complications like leukopenia, infection, renal toxicity, and neurotoxicity (34,35). Nowadays, BRCA1/2 mutations
CDCA2

PFS

HR = 1.35 (1.09−1.66)
logrank P = 0.0057

OS

HR = 1.57 (1.25−1.99)
logrank P = 0.00012

PPS

HR = 1.24 (0.98−1.57)
logrank P = 0.079

CDCA3

PFS

HR = 1.1 (0.95−1.27)
logrank P = 0.21

OS

HR = 1.15 (0.91−1.43)
logrank P = 0.38

PPS

HR = 1.22 (0.95−1.56)
logrank P = 0.11

CDCA4

PFS

HR = 1.06 (0.93−1.22)
logrank P = 0.37

OS

HR = 1.13 (0.98−1.3)
logrank P = 0.095

PPS

HR = 1.24 (0.98−1.57)
logrank P = 0.079

CDCA5

PFS

HR = 1.55 (1.22−1.97)
logrank P = 0.00033

OS

HR = 1.38 (1.09−1.74)
logrank P = 0.00069

PPS

HR = 0.81 (0.62−1.06)
logrank P = 0.13
have been tested for potential therapeutics, which result in less adverse effects. Therefore, in our opinion, CDCA genes probably could act as the effective therapeutics, which exert similar functions as BRCA1/BRCA2. Our findings suggested that CDCA genes expression levels in cancer patients were different from those in normal people. It is difficult to carry out early screening among patients with ovarian cancer, as suggested by our results discussing the different CDCA genes expression levels at different stages, but stage 1 ovarian cancer data are lacking.

CDCA2 can regulate H3 (the primary mitotic histone) phosphorylation depending on PP1 (36). Our results suggested that CDCA2 expression levels in ovarian cancer tissues were up-regulated compared with those in non-carcinoma counterparts. Additionally, the CDCA2 expression level was not correlated with cancer classification among ovarian cancer patients. The up-regulated CDCA2 expression level showed a significant correlation with dismal OS and PFS among ovarian cancer patients.

CDCA3 over-expression may be potentially related to the G1 arrest in the cell cycle, which serves as the crucial cell division checkpoint, leading to the cascade reactions that potentially result in cancer development. CDCA3 is suggested to be the possible breast cancer target. CDCA3 plays a role in triggering several types of cancer [such as liver cancer (37), as well as oral squamous cell cancer (38)]. It was found in this study that, CDCA3 expression level was up-regulated in human ovarian cancer tissues compared with that in the non-carcinoma counterparts, whereas such expression showed no association with cancer classification among ovarian cancer patients. Commonly, the high CDCA3 expression level showed a significant correlation with dismal OS for all ovarian cancer cases.

CDCA4 is one of the TRIP-Br transcription co-factor family members, which is demonstrated to play an important part in adjusting transcription factor (TF) activities, including p53 and E2F1 (22,39). It is important to determine specific target gene expression profiles in certain breast cancer subtypes, so as to develop the novel treatments. According to our results, CDCA4 expression...
Figure 6 CDCA genes co-expression among ovarian cancer patients (String).

level was higher in human ovarian cancer tissues compared with that in non-carcinoma counterparts, and its expression level showed no correlation with cancer classification among ovarian cancer patients. Commonly, the up-regulated CDCA4 expression level showed significant correlation with the dismal PPS among all ovarian cancer patients.

As one of the oncogenes, the over-expression of CDCA4 is detected in a variety of cancer types (40-42). CDCA4 exerts an important part for DNA repairs, as well as sister-chromatid cohesion and separation. The over-expression of CDCA5 has been shown to be related to the dismal prognosis of lung carcinoma (24). In addition, the over-expression of CDCA5 is associated with the malfunction of G1-S transition in bladder urothelial cancer (40). As found in this study, CDCA5 expression level increased in human ovarian cancer tissues compared with that in the non-carcinoma tissues, and such expression showed an association with cancer classification for ovarian cancer patients. Obviously, the high CDCA5 expression level displayed distinct correlation with dismal PFS and OS among ovarian cancer cases.

In this study, CDCA7 expression level was found to be up-regulated in human ovarian cancer tissues relative to that in the non-carcinoma counterparts, and such expression showed a correlation with cancer classification among ovarian cancer patients. Obviously, the high CDCA7 expression level showed significant correlation with dismal OS for ovarian cancer patients.

CDCA8, which is also referred to as the Borealin/Dasra B, has been identified to be one of the chromosomal passenger complex (CPC) members, and it plays an indispensable role in genomic transmission in the process of cell division (43). CDCA8, which can act as a cell mitosis regulator, is demonstrated to show association with lung carcinoma after being subjected to phosphorylation at 4 sites (44). According to the previous meta-analysis, the
Figure 7 Functions of CDCA genes as well as those showing significant association with CDCA genes alterations. (A) Heatmap of the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enriched terms colored by P values. (B) Network of GO and KEGG enriched terms colored by P values. (C) Network of GO and KEGG enriched terms colored by clusters.
over-expression of CDCA8 is related to the low survival of breast cancer patients (45). This study revealed that CDCA8 expression level was up-regulated in human ovarian cancer tissues compared with that in non-carcinoma counterparts, and such expression showed no correlation with cancer classification among ovarian cancer patients. Obviously, the increased CDCA8 expression level displayed significant correlation with the poor PFS among ovarian cancer patients.

Besides, KEGG and GO analyses were also carried out in this study to find the correlations between CDCA genes expression levels as well as linked genes of the highest alteration frequency and the prognosis for ovarian cancer. According to our results, attention should be paid to some pathways, which include: R-HAS-2500257: resolution of sister chromatid cohesion; GO: 0051301: cell division; CORUM: 1118: Chromosomal passenger complex (CPC, including CDCA8, INCENP, AURKB, and BIRC5); CORUM: 127: NDC80 kinetochore complex; M129: PID PLK1 pathway; GO: 0007080: mitotic metaphase plate congression. As a primary regulator of mitotic cell division and the DNA damage response, Polo-like kinase 1 (Plk1) is recognized as a new feasible biomarker in this area (46). PLK1 pathway has been revealed to exert certain function in patients with advanced solid malignancies (47), human hepatocellular carcinoma (HCC) (48,49), and glioma (50).

Certain limitations should be noted in this study. Ovarian cancer is associated with different histological subtypes, such as serous and endometrioid (51-54), which may generate different effects in gene targeted therapy. Consequently, there is still a long way to go for this finding.

The current research systemically examines CDCA genes expression levels and its prognostic significance within ovarian cancer, which sheds more light on the complexity and heterogeneity of ovarian cancer biological properties at the molecular level. This study suggests that up-regulation of CDCA genes expression levels in ovarian cancer tissues probably takes a crucial role in ovarian cancer oncogenesis. Besides, the high CDCA2/3/5/7 expression levels may serve as the potential prognostic markers of the poor survival and prognostic accuracy of ovarian cancer. Moreover, CDCA genes probably exert their functions in tumorigenesis through the PLK1 pathway.

Acknowledgments

Funding: This work was supported by the funds from the Medical Research Fund of Guangdong Province (No. A2018237).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

10. Marchetti C, De Felice F, Palaia I, et al. Risk-reducing salpingooophorectomy: a me-ta-analysis on impact on...
ovarian cancer risk and all cause mortality in BRCA 1 and BRCA 2 mutation carriers. BMC Womens Health 2014;14:150.


36. Qian J, Lesage B, Beullens M, et al. PP1/Repo-


<table>
<thead>
<tr>
<th>Dataset</th>
<th>Age</th>
<th>Stage</th>
<th>Grade</th>
<th>For drawing Kaplan-Meier plot (OS)</th>
<th>For drawing Kaplan-Meier plot (PFS)</th>
<th>For drawing Kaplan-Meier plot (PPS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSE 9891</td>
<td>–</td>
<td>Stage I: 24 (8.4%); stage II: 18 (6.3%); stage III: 217 (76.2%); stage IV: 22 (7.7%); unknown: 4 (1.4%)</td>
<td>Grade I: 20 (7.0%); grade II: 99 (34.7%); grade III: 161 (56.5%); unknown: 5 (1.8%)</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
</tr>
<tr>
<td>GSE 3149</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
</tr>
<tr>
<td>GSE 14764</td>
<td>Mean: 57; range: 36–81</td>
<td>Stage I: 8 (1.0%); stage II: 1 (1.3%); stage III: 69 (86.2%); stage IV: 2 (2.5%)</td>
<td>Grade I: 3 (3.8%); grade II: 23 (28.7%); grade III: 54 (67.5%)</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
</tr>
<tr>
<td>GSE 15622</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
</tr>
<tr>
<td>GSE 18520</td>
<td>61.9±12.7</td>
<td>Stage III–IV: 53 (100%)</td>
<td>Grade III: 53 (100%)</td>
<td>CDCA2-5, CDCA7-8</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSE 23554</td>
<td>–</td>
<td>Stage III–IV: 28 (100%)</td>
<td>Grade I: 2 (7.1%); grade II: 8 (28.6%); grade III: 18 (64.3%)</td>
<td>CDCA4, CDCA8</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSE 19829</td>
<td>63.61±10.51</td>
<td>Stage II: 2 (4.3%); stage III: 22 (81.4%); Stage IV 4 (14.3%)</td>
<td>Grade III: 28 (100%)</td>
<td>CDCA2-5, CDCA7-8</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSE 26712</td>
<td>–</td>
<td>Stage III–IV: 153 (100%)</td>
<td>Grade III: 153 (100%)</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
<td>–</td>
</tr>
<tr>
<td>GSE 27651</td>
<td>–</td>
<td>–</td>
<td>Grade I: 21 (48.8%); grade III: 22 (51.2%)</td>
<td>CDCA2-5, CDCA7-8</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSE 26193</td>
<td>–</td>
<td>Stage I: 20 (19.8%); stage II: 11 (10.9%); stage III: 59 (58.4%); stage IV: 11 (10.9%)</td>
<td>Grade I: 7 (6.9%); grade II: 33 (32.7%); grade III: 67 (66.3%)</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
</tr>
<tr>
<td>GSE 30161</td>
<td>62.57±10.51</td>
<td>Stage III: 53 (91.4%); stage IV: 5 (8.6%)</td>
<td>Grade I: 2 (3.4%); grade II: 19 (32.8%); grade III: 33 (56.9%); unknown: 4 (6.9%)</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
</tr>
<tr>
<td>GSE 63885</td>
<td>–</td>
<td>Stage II: 1 (1.4%); stage III: 59 (84.3%); stage IV: 10 (14.3%)</td>
<td>Grade I: 8 (11.4%); grade II: 44 (62.9%); grade IV: 18 (25.7%)</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
</tr>
<tr>
<td>GSE 51373</td>
<td>–</td>
<td>Stage II: 5 (18.5%); stage III: 19 (70.4%); stage IV: 3 (11.1%)</td>
<td>–</td>
<td>CDCA2-5, CDCA7-8</td>
<td>CDCA2-5, CDCA7-8</td>
<td>–</td>
</tr>
<tr>
<td>TCGA</td>
<td>59.74±11.52</td>
<td>–</td>
<td>Grade I: 6 (1.0%); grade II: 69 (11.8%); grade III: 495 (84.3%); grade IV: 1 (0.2%); unknown: 16 (2.7%)</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
<td>CDCA4, CDCA8</td>
</tr>
</tbody>
</table>

OS, overall survival; PFS, progress-free survival; PPS, post-progress survival.
<table>
<thead>
<tr>
<th>Factors</th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td>Grade</td>
<td>1.46 (1.15–1.85)</td>
<td>0.002</td>
</tr>
<tr>
<td>Stage</td>
<td>1.963 (1.548–2.490)</td>
<td>2.70E–08</td>
</tr>
<tr>
<td>CDCA2</td>
<td>1.089 (1.013–1.172)</td>
<td>0.021</td>
</tr>
<tr>
<td>CDCA3</td>
<td>1.078 (1.011–1.148)</td>
<td>0.021</td>
</tr>
<tr>
<td>CDCA4</td>
<td>1.094 (1.014–1.181)</td>
<td>0.021</td>
</tr>
<tr>
<td>CDCA5</td>
<td>1.076 (1.010–1.147)</td>
<td>0.024</td>
</tr>
<tr>
<td>CDCA7</td>
<td>1.075 (1.002–1.153)</td>
<td>0.044</td>
</tr>
<tr>
<td>CDCA8</td>
<td>1.062 (1.001–1.127)</td>
<td>0.045</td>
</tr>
</tbody>
</table>

GSE 18920, GSE 26193, GSE 30161, GSE 63888 were used for analysis.

**Figure S1** The multivariate analysis for risk factors and CDCA family genes.