Identification of a Prognostic Signature Associated With DNA Repair Genes in Ovarian Cancer

Hengzi Sun, Dongyan Cao, Xiangwen Ma, Jiaxin Yang, Peng Peng, Mei Yu, Huimei Zhou, Ying Zhang, Lei Li, Xiao Huo* and Keng Shen*

Department of Obstetrics and Gynecology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Introduction: Ovarian cancer is a highly malignant cancer with a poor prognosis. At present, there is no accurate strategy for predicting the prognosis of ovarian cancer. A prognosis prediction signature associated with DNA repair genes in ovarian cancer was explored in this study.

Methods: Gene expression profiles of ovarian cancer were downloaded from the GEO, UCSC, and TCGA databases. Cluster analysis, univariate analysis, and stepwise regression were used to identify DNA repair genes as potential targets and a prognostic signature for ovarian cancer survival prediction. The top genes were evaluated by immunohistochemical staining of ovarian cancer tissues, and external data were used to assess the signature.

Results: A total of 28 DNA repair genes were identified as being significantly associated with overall survival (OS) among patients with ovarian cancer. The results showed that high expression of XPC and RECQL and low expression of DMC1 were associated with poor prognosis in ovarian cancer patients. The prognostic signature combining 14 DNA repair genes was able to separate ovarian cancer samples associated with different OS times and showed robust performance for predicting survival (Training set: $p < 0.0001$, AUC = 0.759; Testing set: $p < 0.0001$, AUC = 0.76).

Conclusion: Our study identified 28 DNA repair genes related to the prognosis of ovarian cancer. Using some of these potential biomarkers, we constructed a prognostic signature to effectively stratify ovarian cancer patients with different OS rates, which may also serve as a potential therapeutic target in ovarian cancer.

Keywords: ovarian cancer, DNA damage response, genome integrity, prognostic signature, therapeutic target

INTRODUCTION

Ovarian cancer is the fifth leading cause of cancer-related death in women (Siegel et al., 2018). The average lifetime risk of developing ovarian cancer is 1.3%, and the 5-year survival rate ranges from 29% to 93% depending on the spread of the cancer at diagnosis (Torre et al., 2018). In developed countries, the mortality rate of ovarian cancer has declined only slightly over the past 3 decades, despite advances in treatment strategies and techniques, largely because nearly 60% of cases are diagnosed at an advanced stage due to a lack of obvious symptoms

and adequate screening tests (Siegel et al., 2018). Standard treatment for advanced disease involves debulking surgery and chemotherapy. However, most patients experience relapse within 2 or 3 years after receiving the first-line chemotherapy regimen and die due to chemoresistance (Dinh et al., 2008). Considering the critical role that tumor molecular biology plays in the initiation and progression of tumors, researchers and clinicians have to date focused on effective targeted prognostic and treatment strategies for ovarian cancer (Klinck et al., 2008).

Genomic instability is a hallmark of cancer. Cells develop a complex DNA damage response (DDR) to repair DNA damage and promote maintenance of genome integrity. Defects in DDR are associated with failure to accurately repair damaged DNA in cells, leading to the transformation of normal cells into cancer cells with accumulated genetic changes (Minchom et al., 2018). Due to defects in DDR, cancer cells often show a reduced capacity to repair DNA compared to normal cells and are more reliant on other subsets of repair pathways for survival, which results in DNA replication stress and accumulation of DNA damage. Indeed, approximately 50% of epithelial ovarian cancers are characterized by inactivation of genes required for homologous recombination (Spriggs and Longo, 2018). BRCA 1/2 is well known for its role in DNA repair of double-stranded breaks and recombination (Spriggs and Longo, 2018). Moreover, BRCA mutations have been demonstrated to be associated with improved overall survival (OS) and progression-free survival (PFS) in patients with ovarian cancer. Although the reason remains unclear, it may be due to "synthetic lethality", which leads to cell death due to a synergistic effect in an already DDR-deficient background, such as PARP inhibition (Kaelin, 2005; Moeller et al., 2009). Synthetic lethality between repair pathways also has provided advanced clinical strategies for targeting DNA repair/DDR (Brown et al., 2017). Such issues render DDR targeting a powerful prognostic strategy.

Therefore, we integrated expression profiling based on the Gene Expression Omnibus (GEO), UCSC Xena, and The Cancer Genome Atlas (TCGA) databases to explore DNA repair genes related to the prognosis of ovarian cancer and potentially to explore DDR targeting biomarkers to improve the survival of ovarian cancer patients.

MATERIALS AND METHODS

Data Sources

After obtaining expression data from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), we integrated 2 sets of data, GSE14001 and GSE14407, containing raw data for 32 ovarian cancer samples and 15 para-carcinoma samples. The HT_HG-U133A dataset from UCSC Xena (https://genome-cancer.ucsc.edu/), which contains 565 samples with prognostic follow-up information, was used to analyse the relationship between the DNA repair gene expression profile and prognosis. Finally, we downloaded another set of RNA-Seq data from TCGA (the July version) containing 379 samples

and 364 samples with available follow-up information as an independent validation dataset for survival analysis.

Data Pre-Processing and Differential Expression Analysis

The data were preprocessed as follows. 1) The original chip data of the GEO dataset were downloaded. 2) The R package (R 3.4.0 version) affy was used to process the chip raw data. 3) RMA standardization was performed to convert the data into expression spectrum chip data. 4) The sva combat method of the R package was further employed to remove the batch effect, and probes were then mapped to the genes. The no-load probe was removed, and multiple probes were designed to correspond to a median of 1 gene. 5) Each dataset was further quantified by quantile normalization to extract expression profiles for the DNA repair genes. 6) We screened DNA repair genes in samples with different grade levels by the Wilcoxon rank test and analyzed genes with expression levels that were significantly different among samples with low, high and normal levels using ANOVA, with p-values <0.05 indicating significant differences. 7) Hierarchical clustering was performed for cluster analysis of the DNA repair gene expression profile for each sample, and the prognoses for samples with different expression patterns were examined. 8) Univariate analysis was performed using the R package survival, and a gene with a significance level of p < 0.05 was selected as a prognostic DNA repair gene. 9) Ovarian cancer molecular subtypes were constructed by unsupervised clustering based on prognostic DNA repair genes, and multivariate Cox proportional hazards models were applied to observe their impacts on prognosis. Next, stepwise regression was applied to establish a prognostic signature. 10) Validation was performed using external TCGA (July version) RNA-Seq data containing 379 samples and 364 samples with follow-up information.

Immunohistochemical Staining (IHC)

We collected a total of 200 human ovarian tissue samples, 160 of which had accompanying follow-up information, and 40 cancer-adjacent ovarian tissue samples from archives of paraffin-embedded tissues between February 2009 and February 2013 at the Department of Pathology of Peking Union Medical College Hospital. The follow-up was performed until March 30, 2018. The pathological diagnoses were reconfirmed by a pathologist. The project was approved by the Ethical Committee (Peking Union Medical College Hospital), and informed consent was acquired from patients or family members. IHC was performed as previously described (Li et al., 2010). Antibodies against the following were used: XPC 1:200 abcam ab203693; RECQL 1:50 abcam ab203693; DMC1 1:200 abcam ab203693. The scoring details have been described previously (Zhang et al., 2015). The intensity of immunostaining was graded as follows: 1+, weak; 2+, moderate; 3+, strong or 4+, very strong. The area of positive cancer cells in each microscopic field was categorized as follows: 1+, 0 to 25%; 2+, 25 to 50%; 3+,

50 to 75% or 4+, 75 to 100%. The sum between 5 and 80 was obtained by multiplying the 2 scores by 5. A sum from 0 to 42 was assigned as "low expression" and that from 43 to 80 as "high expression." All pathological diagnoses were confirmed in a blinded manner by 3 expert pathologists.
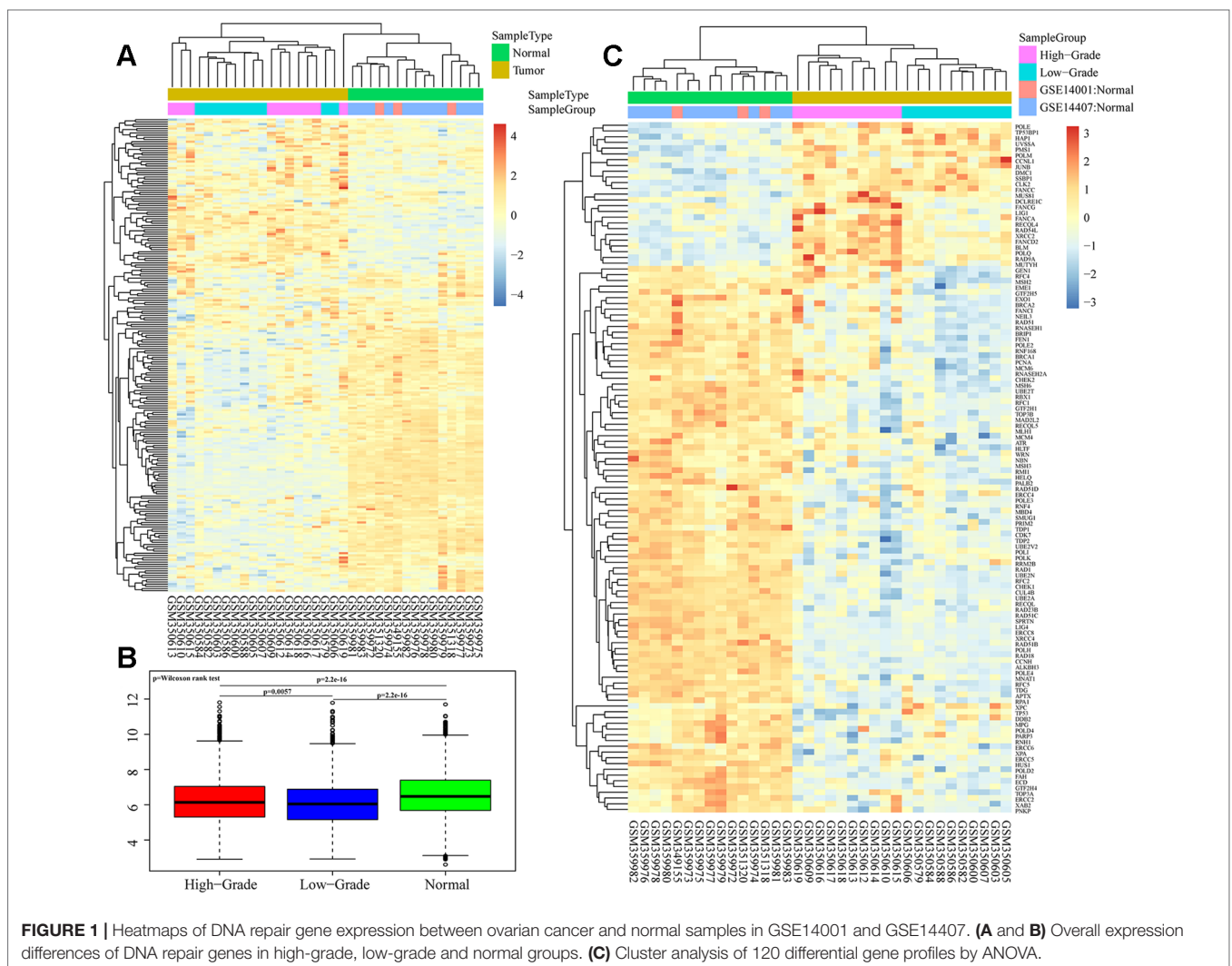
RESULTS

Data Standardization

Two sets of raw chip data (GSE14001 and GSE14407) were downloaded from GEO: GSE14001 contains 10 low-grade ovarian serous carcinoma samples, 10 high-grade ovarian serous carcinoma samples and 3 short-term primary cultures of human ovarian cells; GSE14407 contains 12 ovarian surface epithelial cell lines and 12 ovarian cancer epithelial cell lines. We standardized the raw chip data using the R software package RMA method, and the standardized results are shown in **Supplementary S1a**. The batch effect was further removed using the R package sva combat method, and the results are shown in **Supplementary S1b**. Finally, the results of quantile

normalization are shown in Supplementary **Figure S1C**. The testing set that included 727 DNA repair genes was obtained from the DNA repair-related pathways in KEGG and the literature (**Supplementary S2**).

Expression Patterns of DNA Repair Genes in High-Grade Ovarian Cancer, Low-Grade Ovarian Cancer, and Normal Samples

We analyzed overall expression differences in DNA repair genes in the different groups. As shown in **Figures 1A**, **B**, significant differences were found between the 3 groups. Expression of the DNA repair genes in the normal group was significantly higher than that in the other 2 groups ($p < 0.001$). Moreover, expression of the DNA repair genes in the high-grade group was higher than that in the low-grade group ($p = 0.0057$). We assessed each DNA repair gene using ANOVA and ultimately obtained 120 differentially expressed genes (FDR < 0.05) (**Supplementary S3**). Further cluster analysis of these 120 gene expression profiles showed different expression patterns for these 120 genes in the 3 groups (**Figure 1C**).

FIGURE 1 | Heatmaps of DNA repair gene expression between ovarian cancer and normal samples in GSE14001 and GSE14407. **(A** and **B)** Overall expression differences of DNA repair genes in high-grade, low-grade and normal groups. **(C)** Cluster analysis of 120 differential gene profiles by ANOVA.

The Relationship Between Different Expression Patterns of DNA Repair Genes and Prognosis in Cancer Samples

We quantified the UCSC HT_HG-U133A chip data, containing 565 ovarian cancer cases with follow-up information, and further extracted 102 gene expression profiles for the abovementioned 120 genes.

Hierarchical clustering was used to cluster the expression profiles of DNA repair genes for each sample. The samples were divided into 3 clusters, as shown in **Supplementary S4a**. Only 1 sample was included in Cluster 3. The overall expression level of these 102 DNA repair genes in Cluster 2 was significantly higher than that in Cluster 1, and a significant prognostic difference was identified between Cluster 1 and Cluster 2, indicating that DNA repair was more active in the sample with the better prognosis than in the sample with the poorer prognosis (p = 0.026) (**Supplementary S4b**). A total of 28 DNA repair genes related to prognosis were obtained by univariate survival analysis. The most significant top 20 genes are shown in **Supplementary S5**, with a hazard ratio <1, indicating that lower expression of these genes was related to poor prognosis.

Unsupervised clustering based on these 28 prognostic DNA repair genes is shown in **Figure 2A**. Twenty-eight genes divided the samples into 3 clusters, though Cluster 3 contained only 1 sample. Expression of DNA repair genes in the Cluster 2 and Cluster 1 samples was significantly different, as was the prognosis of the 2 groups (p = 0.0079) (**Figure 2B**).

Prognostic Signature Identification for Ovarian Cancer Patients

In this study, we aimed to obtain a prognostic signature for ovarian cancer prediction. Therefore, we selected the 28 DNA repair genes for constructing a proportional hazard model by multivariate regression and a prognostic signature of DNA repair genes by stepwise regression. A total of 14 DNA repair genes were included. The risk model was as follows: Risk Score = 0.38*XPC − 0.24*PALB2 + 0.29*RECQL − 0.18*XRCC2 + 0.32*GTF2H5 −

0.19*GTF2H4 − 0.22*SSBP1 + 0.24*RAD54L − 0.25*MUTYH − 0.3*SMUG1 − 0.16*TDP1 − 0.24*DDB2 + 0.26*RNH1 + 0.18* TP53BP1. The forest map of these genes is shown in **Figure 3**. Eight genes had a hazard ratio <1, and 6 genes had a hazard ratio >1. To observe the stability of the model, we performed ROC analysis; as shown in **Figure 4A**, the AUC was 0.759. Kaplan-Meier survival analysis of the 2 groups after classifying the samples using the optimal threshold of 0.078 showed a significant difference in prognosis (p < 0.0001) (**Figure 4B**).
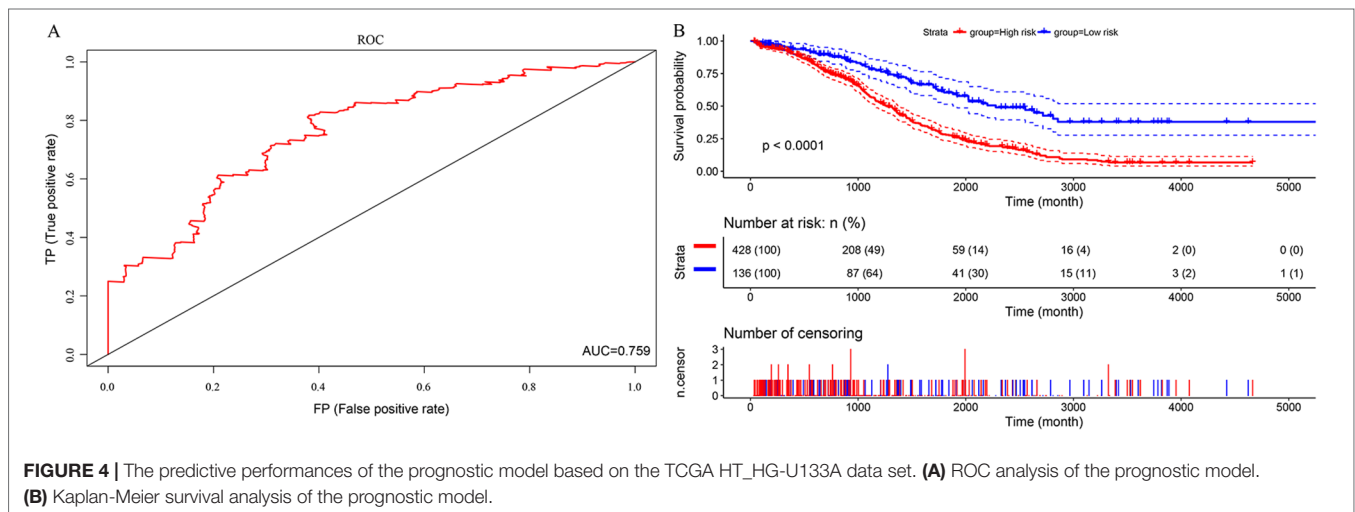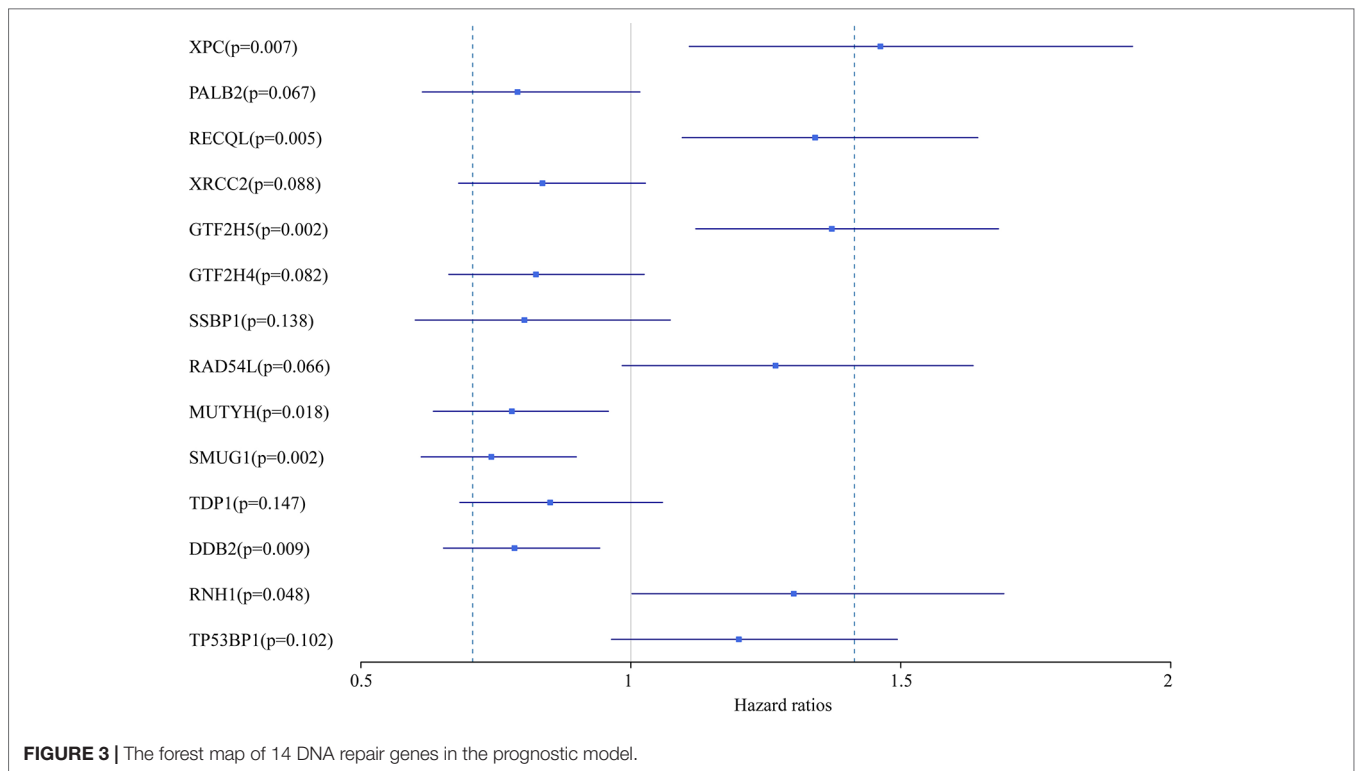
Prognostic Signature Evaluation

To verify that these 14 genes were reproducible and portable in relation to ovarian cancer prognosis, we obtained another RNA-Seq dataset from TCGA containing 364 samples of cancer patients with follow-up information as an independent validation dataset for survival analysis. Expression of these 14 DNA repair genes in the 364 samples was extracted for multivariate analysis with the Cox regression model, and the ROC curves, with AUC values of 0.76, are shown in **Supplementary S6a**. Further analysis of prognosis after classification of the samples is shown in **Supplementary S6b**. The 14 genes had a significant classification effect on ovarian cancer prognosis in the external data, as the survival of the low-risk group was significantly better than that of the high-risk group (p < 0.0001), further demonstrating that the 14 DNA repair genes that we screened are key genes related to the prognosis of ovarian cancer.

Evaluation of the Prognosis of Ovarian Cancer and DNA Repair Genes by IHC

From February 2009 to February 2013, 200 human ovarian tissue samples, 160 of which had accompanying follow-up information, and 40 cancer-adjacent ovarian tissue samples from archives of paraffin-embedded tissues were collected at the Department of Pathology of Peking Union Medical College Hospital. The follow-up was performed until March 30, 2018. Supplementary S7 summarizes the characteristics of all patients, including age, disease stage, and tumor grade. We selected the top 3 genes (*XPC*, *RECQL*, and *DMC1*, excluding *SUMG1* and *GTF2H5*, which have been evaluated in
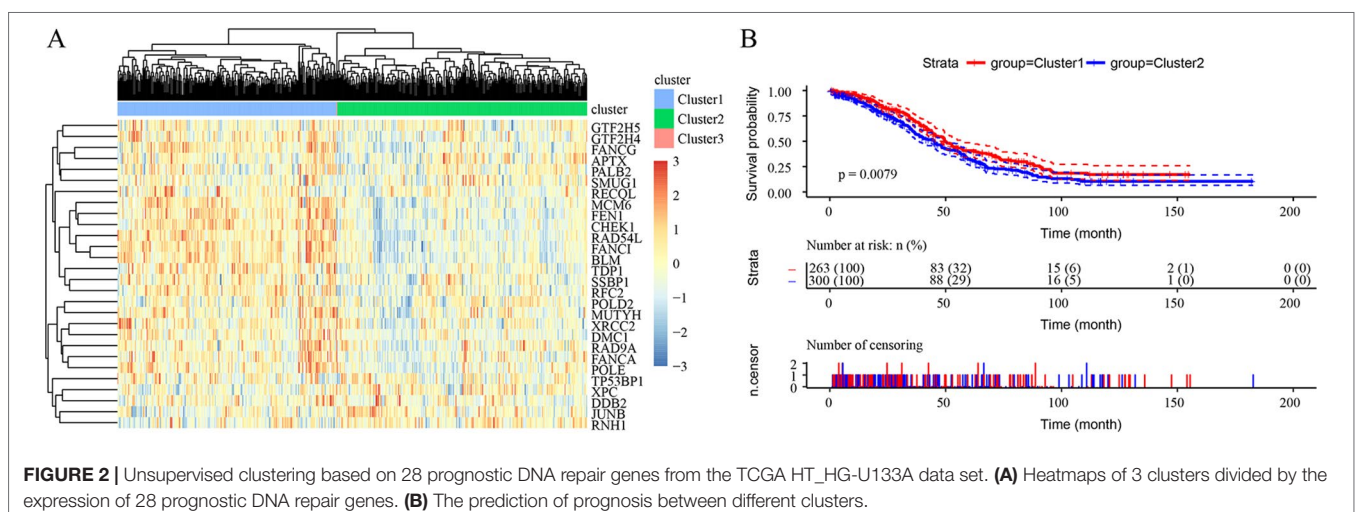
FIGURE 2 | Unsupervised clustering based on 28 prognostic DNA repair genes from the TCGA HT_HG-U133A data set. **(A)** Heatmaps of 3 clusters divided by the expression of 28 prognostic DNA repair genes. **(B)** The prediction of prognosis between different clusters.

FIGURE 3 | The forest map of 14 DNA repair genes in the prognostic model.

FIGURE 4 | The predictive performances of the prognostic model based on the TCGA HT_HG-U133A data set. **(A)** ROC analysis of the prognostic model. **(B)** Kaplan-Meier survival analysis of the prognostic model.

previous studies of ovarian cancer) that have rarely been studied in ovarian cancer to evaluate gene expression values by IHC. The differences in *XPC, RECQL,* and *DMC1* expression between ovarian cancer tissues and adjacent normal ovarian tissues are shown in **Figure 5**. Expression of *XPC* (50.18 ± 1.2 vs 23.13 ± 2.8, $p < 0.01$) and *RECQL* (46.20 ± 1.0 vs 25.25 ± 2.3, $p < 0.01$) was significantly higher in ovarian cancer than in adjacent cancer tissue. Conversely, *DMC1* (28.28 ± 1.5 vs 57.63 ± 2.7, $p < 0.05$) showed lower expression in ovarian cancer tissue. In addition, the correlation between expression of these genes and ovarian cancer prognosis is shown in **Figure 6**. These data reveal that high expression of *XPC* (OS, HR = 1.473, 95% CI 1.032–2.264, $p = 0.043$; PFS, HR = 1.403, 95% CI 1.005–2.114, $p = 0.053$) and *RECQL* (OS, HR = 1.658, 95% CI 1.085–3.032, $p = 0.027$; PFS, HR = 1.668, 95% CI 1.201–2.906, $p = 0.007$) and low expression of *DMC1* (OS, HR = 1.483, 95% CI 0.9710–2.225, $p = 0.071$; PFS, HR = 1.762, 95% CI 1.233–2.479, $p = 0.002$) are associated with poor prognosis in patients with ovarian cancer.
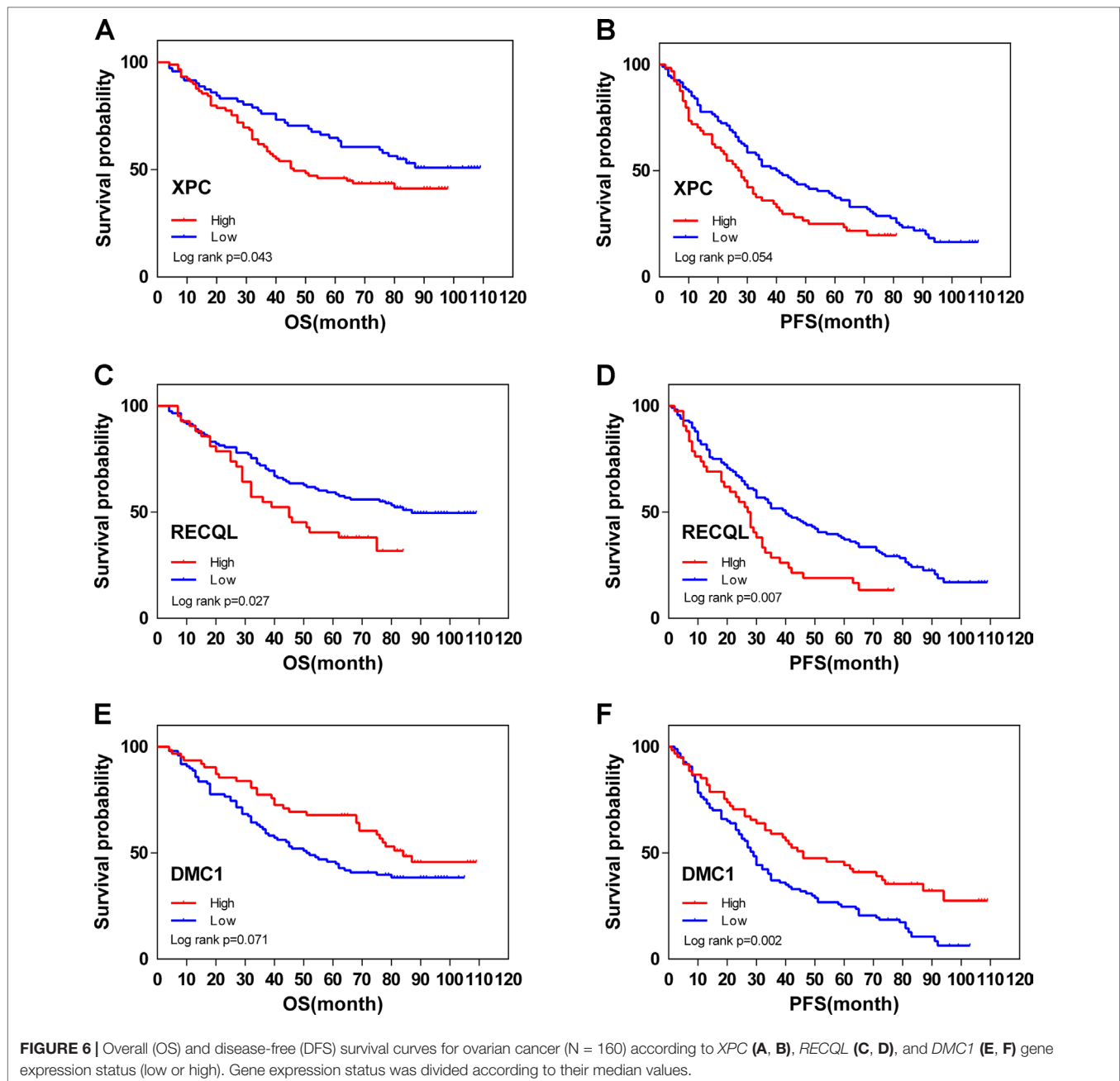
DISCUSSION

Approximately 151,900 women die from ovarian cancer every year worldwide due to the advanced stage of the disease at
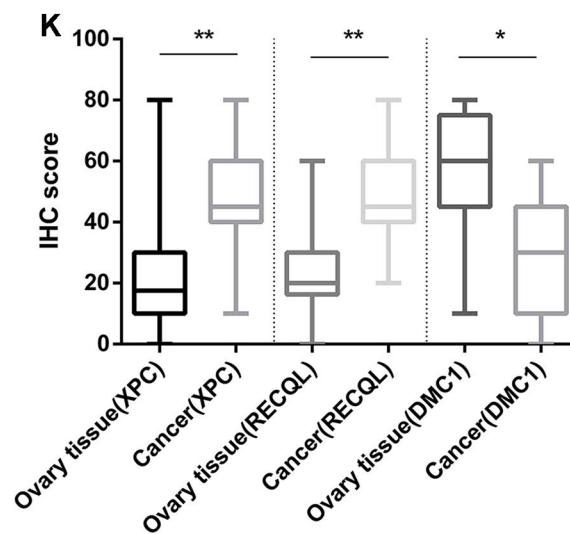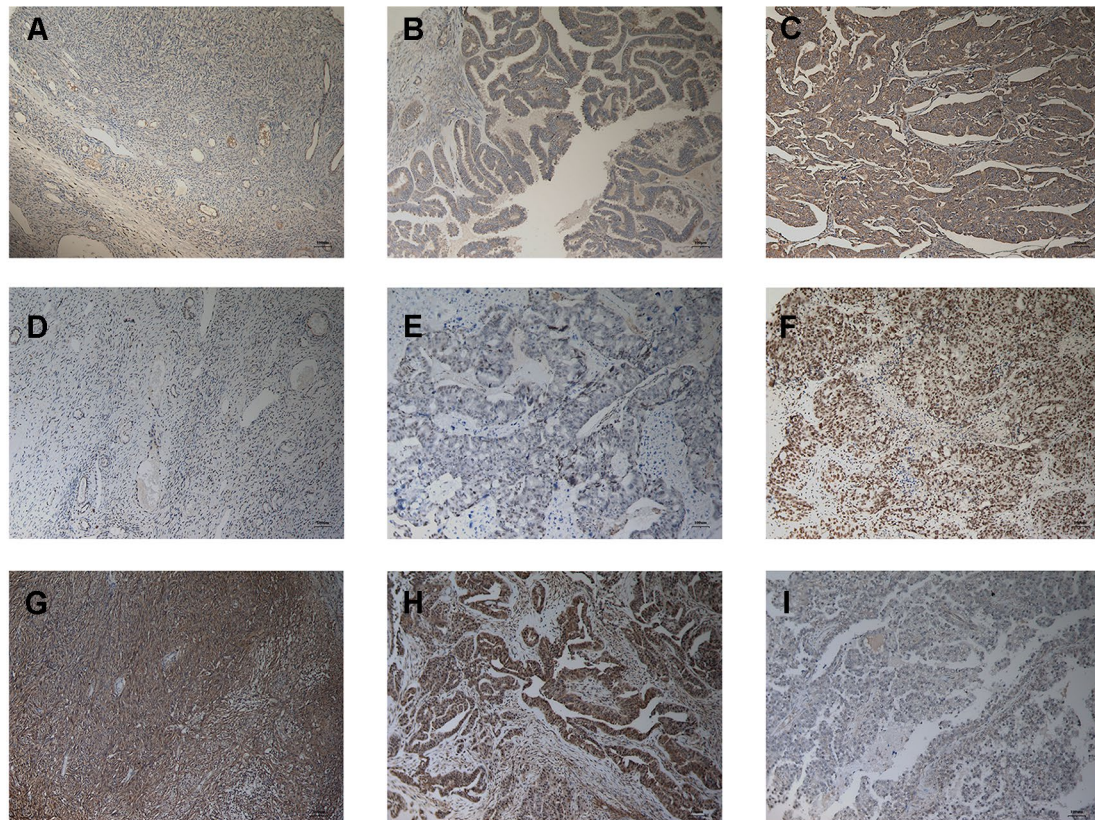
FIGURE 5 | Immunohistochemistry for *XPC*, *RECQL,* and *DMC1*. Samples of ovarian tissue (N = 40) and ovarian cancer (N = 200). Cancer-adjacent ovarian tissue samples with weak immunostaining scores for **(A)** *XPC* or **(D)** *RECQL* or strong immunostaining scores for **(G)** *DMC1*. Ovarian cancer sample of weak and strong immunostaining score for *XPC* **(B**, **C)** or *RECQL* **(E**, **F)** and *DMC1* **(I**, **H)**, respectively. Expression of the *XPC*, *RECQL,* and *DMC1* genes is depicted in **(K)** slides (X 100). *p < 0.05, **p < 0.01.

primary diagnosis and subsequent chemoresistance (Torre et al., 2015). Therefore, effective prognostic and therapeutic strategies to reduce the mortality rate of ovarian cancer are being actively explored. Together with an increased understanding of the role of DNA repair systems, targeting DNA repair/DDR defects through synthetic lethality has provided a paradigm for advanced clinical strategies (Kaelin, 2005; Oza et al., 2015). Approximately 50% of high-grade serous ovarian cancers are

FIGURE 6 | Overall (OS) and disease-free (DFS) survival curves for ovarian cancer (N = 160) according to *XPC* **(A, B)**, *RECQL* **(C, D)**, and *DMC1* **(E, F)** gene expression status (low or high). Gene expression status was divided according to their median values.

characterized by HR deficiency, which involves mutations in *BRCA1/2*, the *MRN* complex (*MRE11*, *RAD50, NBS1*), *ATM*, *RAD51C/D*, *PALB2*, *BRIP1*, *BARD1,* and other genes (O'Connor, 2015; Moschetta et al., 2016). It has also been reported that DNA repair/DDR defects are associated with the prognosis of ovarian cancer (Konstantinopoulos et al., 2015). In the present study, we used a high-throughput (based on TCGA and GEO data) method to search for genetic differences in terms of DNA repair genes associated with ovarian cancer prognosis and conducted a comprehensive analysis to obtain more reliable targets of DNA repair genes and a prognostic signature for ovarian cancer survival prediction. Ovarian

cancer is a highly heterogeneous tumor with a wide variety of types, of which ovarian epithelial cancer is the most common, accounting for 80–90% of ovarian cancer cases. Ovarian cancer is divided into various types according to histological and pathological morphological differences, as follows: serous carcinoma, mucinous carcinoma, endometrioid carcinoma, clear cell carcinoma, and other types of tumors. In 2004, Professor Kurman, proposed a binary model theory based on a series of morphological and molecular genetics data: type I and type II (Shih Ie and Kurman, 2004). Ovarian serous carcinoma is divided into "high-grade serous carcinoma" and "low-grade serous carcinoma." The 2 types of ovarian cancer

have obvious clinical pathological and molecular differences. Once it seemed obvious that ovarian cancers, including high-grade serous ovarian cancer (HGSC), originated in the ovarian surface epithelium (Bell, 2005). In some patients, ovarian cancer is confined to the ovary. Though observed in different patients, viewing these tumors as different phases of the same malignancy, ovarian cancer with advanced-stage disease was assumed to have originated in the ovary (Scully, 1995; Feeley and Wells, 2001). Most advanced-stage ovarian cancers are high-grade serous carcinoma (HGSC). Hence, HGSC is also thought to arise from the ovary (Bell, 2005; Scott and McCluggage, 2006). Moreover, Ovarian carcinomas histopathologically resembling human HGSC can also arise from the ovarian surface epithelium (Flesken-Nikitin et al., 2003; Szabova et al., 2012; Tanwar et al., 2014). Mutations in the p53 gene (TP53 in humans; Trp53 in mice is the most common genetic event observed in human HGSC (2011). At present, it seems that many HGSC arises from serous tubal intraepithelial carcinoma (STIC) formed in the distal fallopian tube epithelium, however, logical and intuitive as it may seem, many of precursor or premalignant lesions, despite consisting of microscopically and genetically cancerous cells, would not progress to lethal malignancies (Esserman et al., 2013; Nikiforov et al., 2016; Brawley, 2017). Mouse studies suggest that many STIC lesions may not progress to invasive, and more critically, metastatic malignancies (Powell et al., 2011; Perets et al., 2013; Zhai et al., 2017). Studies of human HGSCs also note that most STIC lesions likely do not advance to metastatic HGSC, and may thus be classified as low grade (Carlson et al., 2008; Jarboe et al., 2008). Hence, it remains to be elucidated whether STIC could be a bona fide precursor lesion for HGSC in women in the general population who are at average risk, and yet who account for most cases of HGSC. Also, STIC lesions may not be unique to HGSC. Though STICs are associated chiefly with HGSC, they are not exclusive to HGSCs of the ovary, fallopian tube, and peritoneum. These findings suggest that a significant number of HGSCs may derive from precursors independent of STICs (Lim and Oliva, 2013; Howitt et al., 2015; Kim et al., 2018). The overall fraction of ovarian cancers that originate in the fallopian tube is not known. Therefore, both may be the cell of origin in different fractions of cases. In present study, using this dataset in the analysis, we found a high proportion of abnormal expression of DNA repair genes in ovarian cancer (17%). Based on this finding, we speculated that DNA repair genes may have potential as prognostic markers for ovarian cancer, and thus analysis of subsequent large sample queues. In fact, most of the current research is compared the ovarian cancer and adjacent tissues so that to obtain differential expression genes. We will pay more attention to the study of the fallopian tubes in the future. In the present study, expression of DNA repair genes in the normal group was significantly higher than that in the other 2 groups, and DNA repair gene expression in the high-grade group was also higher than that in the low-grade group, demonstrating the differences in gene expression between "high-grade serous carcinoma" and "low-grade serous carcinoma." In the present study, 32 ovarian cancer samples and 15 matched para-carcinoma samples were

analyzed as a training set to identify potential survival-related biomarkers. Another dataset containing 565 ovarian cancer samples with follow-up information was used to re-screen biomarkers related to survival. After hierarchical clustering and univariate survival analysis, 28 survival-related DNA repair genes were retained. We also constructed a proportional hazard model and a prognostic signature of DNA repair genes by stepwise regression based on 14 of the 28 genes. Ultimately, 5 genes showed significance in univariate analysis.

XPC, SMUG1, and GTF2H5 were the top 3 most significant genes associated with ovarian cancer survival according to the prognostic signature. XPC plays a central role in the early steps of global genome nucleotide excision repair (NER), including damage sensing and DNA binding, and shows a preference for single-stranded DNA. Mutations in XPC can result in a rare autosomal recessive disorder termed Xeroderma pigmentosum, which is characterized by increased sensitivity to sunlight and the development of carcinomas at an early age (Sugasawa, 2016). Recently, XPC polymorphisms have been demonstrated to be associated with an increased risk for several types of human malignancies, such as lung, bladder, breast, and esophageal cancers (Zhu et al., 2008). In addition, the XPC rs2228001 A>C polymorphism has a significant association with an increased risk of ovarian cancer, whereas the variant rs2228000 C>T has the opposite association (Zhao et al., 2018). Zhao et al. also reported 3 intronic XPC SNPs (XPC-PAT, rs3731108 and rs1124303) to be associated with prolonged PFS in ovarian cancer, possibly suggesting improved platinum sensitivity (Fleming et al., 2012). However, these authors did not measure mRNA or protein expression of XPC to validate their findings. Therefore, we evaluated the correlation between XPC expression and the prognosis of ovarian cancer by IHC, showing that higher expression of XPC was associated with a poor prognosis. Therefore, we believe that XPC plays a crucial role in the DNA repair pathway of ovarian cancer. SMUG1 is a key enzyme involved in BER that functions by removing uracil from single- and double-stranded DNA and is always associated with rectosigmoid junction neoplasms and bone lymphoma. In addition, increased expression of SMUG1 is often found in various tumor types (including bladder, gastric, breast, esophageal, and cervical cancers) with adverse clinicopathological features, such as poorly differentiated and chemoradiotherapy-resistant tumors (Abdel-Fatah et al., 2013; An et al., 2013; Korourian et al., 2017). Further study has shown increased survival in colorectal cancer for carriers of the genotype SMUG1 rs2233921 TT compared with those with the GT/GG genotypes (Pardini et al., 2013). GTF2H5 participates in the general and transcription-coupled NER of damaged DNA by opening the DNA around a lesion and initiating RNA transcription. Gayarre J et al. evaluated the prognostic and predictive value of GTF2H5 based on IHC staining in 139 ovarian cancer samples and demonstrated that low GTF2H5 expression is associated with an improved prognosis in ovarian cancer patients, which may be due to cisplatin sensitization (Gayarre et al., 2016).

In addition to the abovementioned genes, many cancer susceptibility, progression, and chemotherapy resistance-related DNA repair genes, such as RECQL and DMC1, were

included among the 28 genes in our univariate survival analysis. *RECQL* is involved in various types of DNA repair, including mismatch repair, NER, and direct repair. *RECQL* has been found to be overexpressed in many other tumors and plays a critical role in malignant progression and PARPi resistance and may serve as a potential prognostic or predictive factor for ovarian cancer (Li et al., 2016; Vittori et al., 2017; Viziteu et al., 2017). *DMC1* has been reported to be an essential recombinase for meiotic homologous recombination and plays an important role in generating the diversity of genetic information. Loss of *DMC1* expression is found in multiple human cancers, and SNPs for *DMC1* are associated with cervical cancer. More importantly, *DMC1* interacts directly with the DNA repair gene *BRCA2*, which may provide possibilities for synthetic lethality targets for ovarian cancer (Harada et al., 2001; Martinez et al., 2016). In the present study, we evaluated expression of *RECQL* and *DMC1* in ovarian cancer samples and found that higher expression of *RECQL* and lower expression of *DMC1* were associated with poor prognosis, suggesting that *DMC1* deficits contribute to the progression of ovarian cancer and subsequently enhance other DNA repair genes, such as *XPC* and *RECQL*.

In recent years, many prognostic signatures based on oxidative stress- and immunogenomic-related genes, alternative splicing, genomic and epigenomic mechanisms, DNA methylation, miRNA, and long non-coding RNAs have been constructed, and the predictive performances of these models in chemotherapy-sensitive ovarian cancer have been validated. Nonetheless, few DNA repair gene-related models have been constructed to explore the prognosis of ovarian cancer. Notably, Kang et al. hypothesized a DNA repair pathway-focused score for predicting outcomes and sensitivity to platinum-based chemotherapy in ovarian cancer; however, the accuracy was not high, as the area under the ROC curve was only 0.65, and no evaluation in ovarian cancer tissue was performed (Kang and D'Andrea, 2012). Wang et al. recently stratified ovarian cancer into 7 subgroups within histotypes based on their diverse DNA repair deficiency-related signatures, which were characterized by mutation signatures associated with mismatch repair deficiency, the AID/APOBEC family of cytidine deaminases, age at diagnosis, the prevalence of foldback inversion structural variations, the prevalence of duplications or deletion rearrangements and homologous recombination deficiency. Among the signatures, the prevalence of foldback inversion structural variations was also identified as a prognostically significant high-grade serous cancer group associated with poor survival compared with homologous recombination deficiency (Wang et al., 2017). Furthermore, Manuela Tumiati et al. developed the HR score, which is calculated as the percentage of RAD51-positive cells also positive for both cyclinA2 and CK7, to predict platinum sensitivity and overall survival. The authors reported that low HR scores correlated with platinum sensitivity and improved overall survival (Tumiati et al., 2018). To the best of our knowledge, we herein present is the first high-accuracy prognostic model of ovarian cancer constructed using multiple DNA repair genes. The 14 DNA repair genes in the signature that were evaluated not only show high prediction accuracy but also provide potential targets, including but not

limited to *BRAC1/2*, for synthetic lethality. Although the 14-gene signature has significant prognostic value, some limitations remain; for example, we did not evaluate the synergistic effect of each DNA repair gene in ovarian cancer due to a lack of data. We will further evaluate the synergistic effect of the DNA repair genes in the signature.

CONCLUSION

In conclusion, our study profiled DNA repair genes that are consistently altered between ovarian cancer and normal control samples from the GEO and TCGA databases, and these genes were also evaluated in terms of prognostic prediction based on our samples. The combination of these biomarkers may serve as a signature to stratify ovarian cancer patients into low-risk and high-risk groups for assessing overall survival, which should be helpful for precision and personalized treatment.

DATA AVAILABILITY

All datasets analyzed for this study are cited in the manuscript and the **Supplementary files**.

ETHICS STATEMENT

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments. The project was approved by the institutional ethics committee of Peking Union Medical College Hospital, CAMS Chinese Academy of Medical Sciences, No. S-825 2018. Informed consent was obtained from all participants included in the study.

AUTHOR CONTRIBUTIONS

XH: Study design and data analysis, KS: Study design, funding and clinical data provider, HS: Data collection, data analysis, manuscript writing, follow-up, DC, XM, JY, PP, MY, HZ, YZ, LL: Clinical data provider. All authors have read, edited, and approved the final version of the manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00839/full#supplementary-material

REFERENCES

Abdel-Fatah, T. M., Albarakati, N., Bowell, L., Agarwal, D., Moseley, P., Hawkes, C., et al. (2013). Single-strand selective monofunctional uracil-DNA glycosylase (SMUG1) deficiency is linked to aggressive breast cancer and predicts response to adjuvant therapy. *Breast Cancer Res. Treat.* 142 (3), 515–527. doi: 10.1007/s10549-013-2769-6

An, J. S., Huang, M. N., Song, Y. M., Li, N., Wu, L. Y., and Zhan, Q. M. (2013). A preliminary study of genes related to concomitant chemoradiotherapy resistance in advanced uterine cervical squamous cell carcinoma. *Chin. Med. J.* 126 (21), 4109–4115. doi: 10.3760/cma.j.issn.0366-6999.20131320

Bell, D. A. (2005). Origins and molecular pathology of ovarian cancer. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc.* 18 Suppl 2, S19–S32. doi: 10.1038/modpathol.3800306

Brawley, O. W. (2017). Accepting the existence of breast cancer overdiagnosis. *Ann. Intern. Med.* 166 (5), 364–365. doi: 10.7326/M16-2850

Brown, J. S., O'Carrigan, B., Jackson, S. P., and Yap, T. A. (2017). Targeting DNA repair in cancer: beyond PARP inhibitors. *Cancer Discov.* 7 (1), 20–37. doi: 10.1158/2159-8290.CD-16-0860

Carlson, J. W., Miron, A., Jarboe, E. A., Parast, M. M., Hirsch, M. S., Lee, Y., et al. (2008). Serous tubal intraepithelial carcinoma: its potential role in primary peritoneal serous carcinoma and serous cancer prevention. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 26 (25), 4160–4165. doi: 10.1200/JCO.2008.16.4814

Dinh, P., Harnett, P., Piccart-Gebhart, M. J., and Awada, A. (2008). New therapies for ovarian cancer: cytotoxics and molecularly targeted agents. *Crit. Rev. Oncol. Hematol.* 67 (2), 103–112. doi: 10.1016/j.critrevonc.2008.01.012

Esserman, L. J., Thompson, I. M., Jr., and Reid, B. (2013). Overdiagnosis and overtreatment in cancer: an opportunity for improvement. *Jama* 310 (8), 797–798. doi: 10.1001/jama.2013.108415

Feeley, K. M., and Wells, M. (2001). Precursor lesions of ovarian epithelial malignancy. *Histopathology* 38 (2), 87–95. doi: 10.1046/j.1365-2559.2001.01042.x

Fleming, N. D., Agadjanian, H., Nassanian, H., Miller, C. W., Orsulic, S., Karlan, B. Y., et al. (2012). Xeroderma pigmentosum complementation group C single-nucleotide polymorphisms in the nucleotide excision repair pathway correlate with prolonged progression-free survival in advanced ovarian cancer. *Cancer* 118 (3), 689–697. doi: 10.1002/cncr.26329

Flesken-Nikitin, A., Choi, K. C., Eng, J. P., Shmidt, E. N., and Nikitin, A. Y. (2003). Induction of carcinogenesis by concurrent inactivation of p53 and Rb1 in the mouse ovarian surface epithelium. *Cancer Res.* 63 (13), 3459–3463.

Gayarre, J., Kamieniak, M. M., Cazorla-Jimenez, A., Munoz-Repeto, I., Borrego, S., Garcia-Donas, J., et al. (2016). The NER-related gene GTF2H5 predicts survival in high-grade serous ovarian cancer patients. *J. Gynecol. Oncol.* 27 (1), e7. doi: 10.3802/jgo.2016.27.e7

Harada, H., Nagai, H., Tsuneizumi, M., Mikami, I., Sugano, S., and Emi, M. (2001). Identification of DMC1, a novel gene in the TOC region on 17q25.1 that shows loss of expression in multiple human cancers. *J. Hum. Genet.* 46 (2), 90–95. doi: 10.1007/s100380170115

Howitt, B. E., Hanamornroongruang, S., Lin, D. I., Conner, J. E., Schulte, S., Horowitz, N., et al. (2015). Evidence for a dualistic model of high-grade serous carcinoma: BRCA mutation status, histology, and tubal intraepithelial carcinoma. *Am. J. Surg. Pathol.* 39 (3), 287–293. doi: 10.1097/PAS.0000000000000369

Jarboe, E., Folkins, A., Nucci, M. R., Kindelberger, D., Drapkin, R., Miron, A., et al. (2008). Serous carcinogenesis in the fallopian tube: a descriptive classification. *Int. J. Gynecol. Pathol. Off. J. Int. Soc. Gynecol. Pathol.* 27 (1), 1–9. doi: 10.1097/pgp.0b013e31814b191f

Kaelin, W. G., Jr. (2005). The concept of synthetic lethality in the context of anticancer therapy. *Nat. Rev. Cancer* 5 (9), 689–698. doi: 10.1038/nrc1691

Kang, J., and D'Andrea, A. D. (2012). Kozono D. A DNA repair pathway-focused score for prediction of outcomes in ovarian cancer treated with platinum-based chemotherapy. *J. Natl. Cancer Inst.* 104 (9), 670–681. doi: 10.1093/jnci/djs177

Kim, J., Park, E. Y., Kim, O., Schilder, J. M., Coffey, D. M., Cho, C. H., et al. (2018). Cell origins of high-grade serous ovarian cancer. *Cancers* 10 (11), 433. doi: 10.3390/cancers10110433

Klinck, R., Bramard, A., Inkel, L., Dufresne-Martin, G., Gervais-Bird, J., Madden, R., et al. (2008). Multiple alternative splicing markers for ovarian cancer. *Cancer Res.* 68 (3), 657–663. doi: 10.1158/0008-5472.CAN-07-2580

Konstantinopoulos, P. A., Ceccaldi, R., Shapiro, G. I., and D'Andrea, A. D. (2015). Homologous recombination deficiency: exploiting the fundamental vulnerability of ovarian cancer. *Cancer Discov.* 5 (11), 1137–1154. doi: 10.1158/2159-8290.CD-15-0714

Korourian, A., Roudi, R., Shariftabrizi, A., Kalantari, E., Sotoodeh, K., and Madjd, Z. (2017). Differential role of Wnt signaling and base excision repair pathways in gastric adenocarcinoma aggressiveness. *Clin. Exp. Med.* 17 (4), 505–517. doi: 10.1007/s10238-016-0443-0

Li, Y. L., Ye, F., Cheng, X. D., Hu, Y., Zhou, C. Y., Lu, W. G., et al. (2010). Identification of glia maturation factor beta as an independent prognostic predictor for serous ovarian cancer. *Eur. J. Cancer (Oxford, England: 1990)* 46 (11), 2104–2118. doi: 10.1016/j.ejca.2010.04.015

Li, D., Moughan, J., Crane, C., Hoffman, J. P., Regine, W. F., Abrams, R. A., et al. (2016). RECQ1 A159C polymorphism is associated with overall survival of patients with resected pancreatic cancer: a replication study in NRG Oncology Radiation Therapy Oncology Group 9704. *Int. J. Radiat. Oncol. Biol. Phys.* 94 (3), 554–560. doi: 10.1016/j.ijrobp.2015.10.062

Lim, D., and Oliva, E. (2013). Precursors and pathogenesis of ovarian carcinoma. *Pathology* 45 (3), 229–242. doi: 10.1097/PAT.0b013e32835f2264

Martinez, J. S., von Nicolai, C., Kim, T., Ehlen, A., Mazin, A. V., Kowalczykowski, S. C., et al. (2016). BRCA2 regulates DMC1-mediated recombination through the BRC repeats. *Proc. Nat. Acad. Sci. U. S. A.* 113 (13), 3515–3520. doi: 10.1073/pnas.1601691113

Minchom, A., Aversa, C., and Lopez, J. (2018). Dancing with the DNA damage response: next-generation anti-cancer therapeutic strategies. *Ther. Adv. Med. Oncol.* 10, 1758835918786658. doi: 10.1177/1758835918786658

Moeller, B. J., Pasqualini, R., and Arap, W. (2009). Targeting cancer-specific synthetic lethality in double-strand DNA break repair. *Cell cycle (Georgetown, Tex)* 8 (12), 1872–1876. doi: 10.4161/cc.8.12.8743

Moschetta, M., George, A., Kaye, S. B., and Banerjee, S. (2016). BRCA somatic mutations and epigenetic BRCA modifications in serous ovarian cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* 27 (8), 1449–1455. doi: 10.1093/annonc/mdw142

Nikiforov, Y. E., Seethala, R. R., Tallini, G., Baloch, Z. W., Basolo, F., Thompson, L. D., et al. (2016). Nomenclature revision for encapsulated follicular variant of papillary thyroid carcinoma: a paradigm shift to reduce overtreatment of indolent tumors. *JAMA Oncol.* 2 (8), 1023–1029. doi: 10.1001/jamaoncol.2016.0386

O'Connor, M. J. (2015). Targeting the DNA damage response in cancer. *Mol. Cell* 60 (4), 547–560. doi: 10.1016/j.molcel.2015.10.040

Oza, A. M., Cibula, D., Benzaquen, A. O., Poole, C., Mathijssen, R. H., Sonke, G. S., et al. (2015). Olaparib combined with chemotherapy for recurrent platinum-sensitive ovarian cancer: a randomised phase 2 trial. *Lancet Oncol.* 16 (1), 87–97. doi: 10.1016/S1470-2045(14)71135-0

Pardini, B., Rosa, F., Barone, E., Di Gaetano, C., Slyskova, J., Novotny, J., et al. (2013). Variation within 3'-UTRs of base excision repair genes and response to therapy in colorectal cancer patients: a potential modulation of microRNAs binding. *Clin. Cancer Res.* 19 (21), 6044–6056. doi: 10.1158/1078-0432.CCR-13-0314

Perets, R., Wyant, G. A., Muto, K. W., Bijron, J. G., Poole, B. B., Chin, K. T., et al. (2013). Transformation of the fallopian tube secretory epithelium leads to high-grade serous ovarian cancer in Brca;Tp53;Pten models. *Cancer Cell* 24 (6), 751–765. doi: 10.1016/j.ccr.2013.10.013

Powell, C. B., Chen, L. M., McLennan, J., Crawford, B., Zaloudek, C., Rabban, J. T., et al. (2011). Risk-reducing salpingo-oophorectomy (RRSO) in BRCA mutation carriers: experience with a consecutive series of 111 patients using a standardized surgical-pathological protocol. *Int. J. Gynecol. Cancer Off. J. Int. Gynecol. Cancer Soc.* 21 (5), 846–851. doi: 10.1097/IGC.0b013e31821bc7e3

Scott, M., and McCluggage, W. G. (2006). Current concepts in ovarian epithelial tumorigenesis: correlation between morphological and molecular data. *Histol. Histopathol.* 21 (1), 81–92. doi: 10.14670/hh-21.81

Scully, R. E. (1995). Pathology of ovarian cancer precursors. *J. Cell. Biochem. Suppl.* 23, 208–218. doi: 10.1002/jcb.240590928

Shih Ie, M., and Kurman, R. J. (2004). Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. *Am. J. Pathol.* 164 (5), 1511–1518. doi: 10.1016/S0002-9440(10)63708-X

Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer statistics, 2018. *CA Cancer J. Clin.* 68 (1), 7–30. doi: 10.3322/caac.21442

Spriggs, D. R., and Longo, D. L. (2018). Progress in BRCA-mutated ovarian cancer. *N. Engl. J. Med.* 379 (26), 2567–2568. doi: 10.1056/NEJMe1812644

Sugasawa, K. (2016). Molecular mechanisms of DNA damage recognition for mammalian nucleotide excision repair. *DNA Repair* 44, 110–117. doi: 10.1016/j.dnarep.2016.05.015

Szabova, L., Yin, C., Bupp, S., Guerin, T. M., Schlomer, J. J., Householder, D. B., et al. (2012). Perturbation of Rb, p53, and Brca1 or Brca2 cooperate in inducing metastatic serous epithelial ovarian cancer. *Cancer Res.* 72 (16), 4141–4153. doi: 10.1158/0008-5472.CAN-11-3834

Tanwar, P. S., Mohapatra, G., Chiang, S., Engler, D. A., Zhang, L., Kaneko-Tarui, T., et al. (2014). Loss of LKB1 and PTEN tumor suppressor genes in the ovarian surface epithelium induces papillary serous ovarian cancer. *Carcinogenesis* 35 (3), 546–553. doi: 10.1093/carcin/bgt357

The Cancer Genome Atlas Research Network (2011). Integrated genomic analyses of ovarian carcinoma. *Nature* 474 (7353), 609–615. doi: 10.1038/nature10166

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. *CA Cancer J. Clin.* 65 (2), 87–108. doi: 10.3322/caac.21262

Torre, L. A., Trabert, B., DeSantis, C. E., Miller, K. D., Samimi, G., Runowicz, C. D., et al. (2018). Ovarian cancer statistics, 2018. *CA Cancer J. Clin.* 68 (4), 284–296. doi: 10.3322/caac.21456

Tumiati, M., Hietanen, S., Hynninen, J., Pietila, E., Farkkila, A., Kaipio, K., et al. (2018). A functional homologous recombination assay predicts primary chemotherapy response and long-term survival in ovarian cancer patients. *Clin. Cancer Res.* 24 (18), 4482–4493. doi: 10.1158/1078-0432.CCR-17-3770

Vittori, M., Breznik, B., Hrovat, K., Kenig, S., and Lah, T. T. (2017). RECQ1 helicase silencing decreases the tumour growth rate of U87 glioblastoma cell xenografts in zebrafish embryos. *Genes* 8 (9), 222. doi: 10.3390/genes8090222

Viziteu, E., Klein, B., Basbous, J., Lin, Y. L., Hirtz, C., Gourzones, C., et al. (2017). RECQ1 helicase is involved in replication stress survival and drug resistance in multiple myeloma. *Leukemia* 31 (10), 2104–2113. doi: 10.1038/leu.2017.54

Wang, Y. K., Bashashati, A., Anglesio, M. S., Cochrane, D. R., Grewal, D. S., Ha, G., et al. (2017). Genomic consequences of aberrant DNA repair mechanisms stratify ovarian cancer histotypes. *Nat. Genet.* 49 (6), 856–865. doi: 10.1038/ng.3849

Zhai, Y., Wu, R., Kuick, R., Sessine, M. S., Schulman, S., Green, M., et al. (2017). High-grade serous carcinomas arise in the mouse oviduct *via* defects linked to the human disease. *J. Pathol.* 243 (1), 16–25. doi: 10.1002/path.4927

Zhang, S. F., Wang, X. Y., Fu, Z. Q., Peng, Q. H., Zhang, J. Y., Ye, F., et al. (2015). TXNDC17 promotes paclitaxel resistance *via* inducing autophagy in ovarian cancer. *Autophagy* 11 (2), 225–238. doi: 10.1080/15548627.2014.998931

Zhao, Z., Zhang, A., Zhao, Y., Xiang, J., Yu, D., Liang, Z., et al. (2018). The association of polymorphisms in nucleotide excision repair genes with ovarian cancer susceptibility. *Biosci. Rep.* 38 (3), BSR20180114. doi: 10.1042/BSR20180114

Zhu, Y., Yang, H., Chen, Q., Lin, J., Grossman, H. B., Dinney, C. P., et al. (2008). Modulation of DNA damage/DNA repair capacity by XPC polymorphisms. *DNA Repair* 7 (2), 141–148. doi: 10.1016/j.dnarep.2007.08.006