Introduction

Ovarian cancer (OC) is a highly malignant tumor of the female reproductive system. Globally, there are nearly 240,000 new cases of OC diagnosed each year, and the death rate exceeds 50% (1,2). Patients with localized OC can be treated with standard treatments and have a good 5-year survival rate (3,4). However, detecting OC at an early stage is extremely challenging. Two-thirds of patients are diagnosed with an advanced stage of OC; in such cases, the disease has already spread throughout the peritoneal cavity and distantly metastasized at the time of diagnosis (5). Standard treatments have limited efficacy for advanced or
were obtained from GEO for analysis of DEGs in OC. The files of the platform and the series of matrix files were downloaded, then calibrated and log-transformed using the R version 3.2.0. DEGs in the 3 microarray datasets above were screened out with the limma package (log FC >1 and adjusted P<0.05), and common DEGs among the datasets were identified using RobustRankAggreg. Package (P<0.05) and the heatmap of these DEGs was generated by heatmap package.

**Functional enrichment analysis**

To investigate the gene functions of the DEGs obtained from the 3 GEO datasets, Metascape (http://metascape.org) (13) was used to perform GO function and KEGG pathway enrichment analyses, with P<0.01 set as the cutoff for significance. To investigate the function of modules, the DAVID bioinformatics resource was also used.

**PPI network construction**

The PPI network of the DEGs was constructed using the STRING database (14) and Cytoscape v 3.6.0. The combined score of PPI was 0.468, the PPI enrichment P value was <0.01. Molecular Complex Detection (MCODE) clustering analysis was applied to find clusters of genes in the PPI network. The cutoff parameters were: degree cutoff =2; node score cutoff =0.2; k-core =2; and max. depth =100.

**ONCOMINE analysis**

The ONCOMINE microarray database (http://www.oncomine.org) is a website containing data on gene expressions in multiple cancers (15). We used ONCOMINE to compare the gene expression level of structural maintenance of chromosomes protein 4 (SMC4) in OC tissues with that in normal tissues. The screening conditions were set to log FC >2; P value <0.01; and top 10% gene rank.

**KM plotter analysis**

The prognostic values of the DEGs in OC were determined using the KM survival plotter (www.kmplot.com), an online database containing gene expression profiles and clinical data from public databases (16). The hub genes in the PPI network were input into the database to examine their associations with the survival prognosis and of patients with OC. KM survival plots were used to compare the overall
survival (OS) and progression-free survival (PFS) of the low and high expression groups. The threshold for follow-up was 120 months.

**Expression of SMC4 in OC tissues**

The Human Protein Atlas v19 (http://www.proteinatlas.org/) is a database containing huge numbers of histological images of multiple tumors (17). We selected immunohistochemical (IHC) staining images from the database to compare the protein expression of SMC4 between OC tissues and normal tissues.

**Mutations of SMC4 gene in OC**

The cBio Cancer Genomics Portal (cBioPortal) database (http://cBioportal.org) explores and visualizes a large amount of genomic data of patients with cancer from sources including the Cancer Genome Atlas (TCGA) (18). We used cBioPortal to analyze somatic mutations and DNA copy number alterations (CNAs) of SMC4 in OC. We further analyzed the genetic alterations of SMC4 using data from UCSC Xena browser (http://xena.ucsc.edu/) (19), and copy number data were categorized into 4 groups as follows: shallow deletions, diploids, gains, and amplifications. UALCAN (http://ualcan.path.uab.edu) is a comprehensive interactive web tool based on level 3 RNA sequencing and clinical data of 31 cancer types in the TCGA database (20). We used UALCAN to identify the genes positively correlated with SMC4 and performed a functional pathway analysis of these genes using the DAVID bioinformatics resource.

**GSEA**

The Gene set enrichment analysis (GSEA) v3.0 software (www.broadinstitute.org/gsea/) was used to perform the functional pathway analysis of the SMC4 gene (21). Expression data were classified into high and low expression subgroups. Then, the enrichment analysis was done by employing the default weighted enrichment statistical method. The random combinatorial count was set to 1,000 times.

**Statistical analysis**

Several packages in R software were used to finish statistical calculations and graphs. The Kaplan-Meier plotter was employed to generate survival curves, and the threshold for follow-up was 120 months. P<0.05 and logFC ≥1 were considered as statistically significant for screening DEGs. log FC ≥2 and P<0.01 was set as statistically significant for the ONCOMINE analysis. P<0.01 was set to construct PPI network and perform enrichment analysis.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

**Results**

**Identification of DEGs in the GEO database**

From the GSE14407 dataset, 2,756 DEGs (1,137 upregulated genes and 1,619 downregulated genes) were screened out. From the GSE18520 dataset, 2,623 DEGs (1,037 upregulated genes and 1,586 downregulated genes) were identified. From the GSE26712 dataset, 1,200 DEGs (444 upregulated genes and 756 downregulated genes) were screened out. The common DEGs identified from the 3 datasets were analyzed with RobustRankAggreg, and finally, a total of 879 common DEGs were obtained. The top 20 upregulated DEGs and downregulated DEGs are shown in a heatmap in Figure 1.

**Enrichment analysis of DEGs**

Metascape was used to perform function enrichment analyses and the top 20 Go and KEGG pathways were plotted. For GO term enrichment analysis, the results showed that the upregulated DEGs were predominantly enriched in cell division, cell cycle phase transition, DNA conformation change. KEGG analysis revealed that the upregulated DEGs were mostly enriched in cell cycle, G2/M checkpoints (Figure 2A). For GO term enrichment analysis, the downregulated DEGs were significantly enriched in developmental growth, cell part morphogenesis, inflammatory response (Figure 2B).

**PPI network**

The PPI network had 626 nodes and 6,353 interactions (Figure 3A). Further, an important module containing 76 nodes and 2,656 interactions was identified from the PPI network (Figure 3B). The top 10 hub genes showing significant interactions were ZWINT, SMC4, NDC80, NEK2, AURKB, CENPF, KIF20A, KIF11, FAM83D, and CENPE. The GO analyses showed that the genes in the
module were mainly enriched in protein binding, ATP binding (Figure 3C). The KEGG pathway analysis showed that these genes were enriched in cell cycle, oocyte meiosis, progesterone-mediated oocyte maturation and DNA replication (Figure 3D).

Expression of SMC4 in common tumor types

Using ONCOMINE, we found 425 studies that had investigated the mRNA level of SMC4 in human cancers compared with normal tissues. Of them, 67 studies reported that the expression levels of SMC4 were increased in human cancers including breast cancer, colorectal cancer, cervical cancer and OC, while 5 studies reported the downregulation of SMC4 (Figure 4A). Studies showed that SMC4 was highly expressed in OC tissues compared to normal tissues (Figure 4B,C). Subsequently, we validated the protein expression levels of SMC4 by performing analysis using the Human Protein Atlas. The results showed SMC4 to be upregulated in OC tissues and downregulated in normal tissues (Figure 4D,E).

Genetic alterations of SMC4 in OC

We further discussed the causes of SMC4 overexpression in OC. From cBioPortal database, SMC4 alterations accounted for 7 to 18% of genetic alterations in OC. And alterations included somatic mutations, amplifications, and multiple alterations (Figure 5A,B,C,D). Notably, we found that the majority of SMC4 alterations were copy number amplifications. Missense mutation and embedded deletion (Figure 5C,E) were also frequently detected. We also analyzed the genetic alterations of SMC4 using the UCSC Xena browser. The copy number data in the TCGA were downloaded and categorized into 4 groups (shallow deletions, diploids, gains, and amplifications). With the amplification of copy numbers, the corresponding SMC4 gene expression increased significantly (Figure 5F,G,H). We found that SMC4 copy number amplification was related to high SMC4 expression in OC. UALCAN, a tool for in-depth analysis of TCGA data, was utilized to verify the correlated hub genes of SMC4. Then, GO and KEGG pathway analyses of SMC4 and its related hub genes were performed. The GO analysis showed that those genes were significantly enriched in regulation of signal transduction by P53 class mediator, regulation cell cycle, cell division and so on. The KEGG pathway analysis showed that the genes were enriched in progesterone-mediated oocyte maturation, the P53 signaling pathway, oocyte meiosis, cell cycle and DNA replication (Figure 5I,J).

KM Plotter analysis

As shown in Figure 6, the prognostic value of SMC4 expression in OC was also analyzed. A high expression of SMC4 was correlated with significantly shorter OS and PFS in patients with OC (Figure 6A,B). Therefore, the SMC4 was subsequently screened as a putative prognostic marker for OC.

Enrichment analysis of SMC4 gene

SMC4 was found to be highly enriched in the cell cycle, spliceosome, ubiquitin mediated proteolysis, and adherens junctions (Figure 7A,B,C,D), which suggested...
that \( SMC4 \) might promote the development of OC via these signaling pathways.

**Discussion**

Novel biomarkers for OC could inform personalized treatment decisions and aid in the early prediction of prognosis for patients who are at high risk of disease recurrence and death. Unfortunately, prognostic biomarkers for OC are lacking. Here, we integrated gene expression data from the GEO datasets GSE26712, GSE18520, and GSE14407, and utilized muti-bioinformatics methods to explore potential prognostic biomarkers for OC. We identified 879 common DEGs from the 3 gene expression datasets. By constructing the PPI network, we screened out 10 hub genes in significant modules.

The roles of the hub genes in the progression of OC were confirmed by the results of functional analyses. Furthermore, we found that 1 hub gene in the PPI network, \( SMC4 \), had a high expression at both the mRNA and protein levels, and an increased expression of \( SMC4 \) was associated with an unfavorable prognosis of OC. The results of analysis using cBioPortal showed that \( SMC4 \) alterations accounted for 7% to 18% of genetic alterations in OC. The majority of \( SMC4 \) alterations were copy number amplifications. Increased copy numbers levels were correlated with higher \( SMC4 \) expression. UALCAN was utilized to verify the hub genes co-expressed with \( SMC4 \). Finally, GSEA analysis showed that \( SMC4 \) were mainly enriched in the cell cycle, spliceosome, ubiquitin mediated proteolysis, and adherens junctions. \( SMC4 \) was subsequently identified as a prognostic biomarker closely associated with biological function, which
Figure 3 Protein–protein interaction analysis and construction of module of network. (A) Regulatory network of common differentially expressed genes (DEGs) in ovarian cancer (OC). (B) A significant module enriched by upregulated genes in OC. (C,D) Results of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses of the module.
Figure 4 The expression levels of structural maintenance of chromosomes protein 4 (SMC4) in tumor tissues compared with normal tissues. (A) SMC4 expression in various types of cancer from the ONCOMINE database. (B,C) SMC4 expression in different ovarian cancer studies in the ONCOMINE database. Immunohistochemical (IHC) images of anti-SMC4 staining in human ovarian cancer (D) and normal tissue (E) taken from the Human Protein Atlas (http://www.proteinatlas.org).

might facilitate the prediction of prognosis in patients with OC.

SMC4 belongs to the structural maintenance of chromosomes (SMC) family, the members of which are involved in many physiological and pathological processes (22,23). SMC4 is highly expressed in multiple tumors, including hepatocellular carcinoma (24), colorectal cancer (25), prostate cancer (26), and lung cancer (27). Its high expression in hepatocellular carcinoma is responsible for tumor dedifferentiation and vascular invasion, and is correlated to an advanced disease stage (28). In Zhao et al.’s study, knockdown of SMC4 via RNA interference reduced prostate cancer cell migration and invasion (29). In another study, SMC4 overexpression was found to promote the proliferative, migratory, invasive capabilities of glioma cells (30). The above studies suggested that SMC4 played an important role in multiple tumors; however, its expression and function in OC were poorly understood. In the present study, we discovered abnormal expression and amplification of the SMC4 in OC. Our observations indicated that over expression SMC4 contributed to the tumorigenicity and prognosis of OC.

We also hypothesized that the aberrant expression of SMC4 affected the outcomes of patients with OC by regulating various signaling pathways. We found that SMC4 activates signaling pathways related to the cell cycle, spliceosome, ubiquitin mediated proteolysis, and adherens junctions, which were involved in the development of OC.

Dysregulation of the cell cycle pathway is important to tumorigenesis (31-34). SMC4 is associated with dysregulation of the cell cycle. For instance, Jiang et al. found that SMC4 upregulation drastically increased the proliferative capability of glioma cells by accelerating G1-S-phase transition (30). In A549 cells, knockdown of SMC4 was also observed to downregulate the cell cycle-related proteins cyclin B1 and cyclin-dependent kinase.
1 to suppress cell proliferation (27). Thus, we postulate that SMC4 affected cell proliferation in OC via cell cycle regulation.

Our study also found that upregulation of SMC4 plays an essential role in OC via spliceosome, ubiquitin mediated proteolysis and adherens junctions. Aberrations of the spliceosome pathway have been shown to be involved in multiple processes in cancer, such as invasion, metastasis, and angiogenesis (35,36). Spliceosomal pathway genes affected the proliferation and metastasis of hepatocellular carcinoma (37). Ubiquitin mediated proteolysis has emerged in multiple cancers, and it was responsible for oncogenic transformation.
Figure 6 KM plotter analysis of structural maintenance of chromosomes protein 4 (SMC4) for the overall survival (OS) and progression-free survival (PFS) in ovarian cancer (OC) patients. (A) The OS rate of OC patients with high SMC4 expression or low SMC4 expression. (B) The PFS rate of OC patients with high SMC4 expression or low SMC4 expression.

Figure 7 Structural maintenance of chromosomes protein 4 (SMC4) correlated gene set enrichment analysis (GSEA) was performed. The results showed that SMC4 was enriched in (A) cell cycle, (B) spliceosome, (C) ubiquitin mediated proteolysis, (D) adherens junction.
for glioma (38). Adherens junction controlled normal cell-cell adhesion, might play an vital role in driving cancer cell dissemination (39). However, these pathway has rarely been reported in OC. Further investigation of the relationship between the regulation of these pathways and development of OC should be also carried out.

Our preliminary study has confirmed SMC4 as a prognostic biomarker in OC and its’ activated signaling pathways. However our study had some limitations, the experiments in vitro and in vivo for validation should be conducted. We would further discuss the specific impacts of SMC4 on resistance of platinum-based chemotherapy and recurrence in OC, which are main reasons for treatment failure and main factor affecting outcome in future study.

Conclusions

In the present study, we used a comprehensive bioinformatics analysis to identify DEGs in patients with OC. We finally identified SMC4 to have vital involvement in OC, and its overexpression to be closely associated with a poor prognosis in patients with the disease. We also found that SMC4 plays important roles in the biological processes of OC. As a conclusion, we could tell that SMC4 is a useful and novel prognostic indicator for patients with OC. Further molecular biological experiments should be performed to illuminate the biological characteristics of SMC4 in OC.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at http://dx.doi.org/10.21037/tcr-21-380

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr-21-380). 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. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


(English Language Editor: J. Reynolds)