Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Gene Expression Profile of Active HE4 Stimulation in Epithelial Ovarian Cancer Cells: Microarray Study and Comprehensive Bioinformatics Analysis

Liancheng Zhu
Shengjing Hospital of China Medical University https://orcid.org/0000-0002-0814-5816
Mingzi Tan
Liaoning Cancer Institute and Hospital
Haoya Xu
Shengjing Hospital of China Medical University
Bei Lin (linbei88@hotmail.com)
Shengjing Hospital of China Medical University

Research

Keywords: HE4 stimulation, microarray, bioinformatics, epithelial ovarian cancer, MAPK pathway

Posted Date: July 23rd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-46014/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background.

Human Epididymis Protein 4 (HE4) is a novel serum biomarker for diagnosis of epithelial ovarian cancer (EOC) with high specificity and sensitivity compared with CA125, and the increasing researches have been carried out on its roles in promoting carcinogenesis and chemoresistance in EOC in recent years, however, its underlying molecular mechanisms remain poorly understood. The aim of this study was to elucidate the molecular mechanisms of HE4 stimulation and to identify the key genes and pathways mediating carcinogenesis in EOC using microarray and bioinformatics analysis.

Methods.

We established a stable HE4-silence ES-2 ovarian cancer cell line labeled as “S”, and its active HE4 protein stimulated cells labeled as “S4”. Human whole genome microarray analysis was used to identify deferentially expressed genes (DEGs) from triplicate samples of S4 and S cells. “clusterProfiler” package in R, DAVID, Metascape, and Gene Set Enrichment Analysis (GSEA) were used to perform gene ontology (GO) and pathway enrichment analysis, and cBioPortal for WFDC2 coexpression analysis. GEO dataset (GSE51088) and quantitative real-time polymerase chain reaction (qRT-PCR) was applied for validation. The protein–protein interaction (PPI) network and modular analyses were performed using Metascape and Cytoscape.

Results.

In total, 713 DEGs were found (164 up regulated and 549 down regulated) and further analyzed by GO, pathway enrichment and PPI analyses. We found that MAPK pathway accounted for a significant portion of the enriched terms. WFDC2 coexpression analysis revealed ten WFDC2 coexpressed genes (TMEM220A, SEC23A, FRMD6, PMP22, APBB2, DNAJB4, ERLIN1, ZEB1, RAB6B, and PLEKHF1) that were also dramatically changed in S4 cells and validated by dataset GSE51088. Kaplan–Meier survival statistics revealed clinical significance for all of the 10 target genes. Finally, PPI was constructed, sixteen hub genes and eight molecular complex detections (MCODEs) were identified, the seeds of five most significant MCODEs were subjected to GO and KEGG enrichment analysis and their clinical significance was evaluated.

Conclusions.

By applying microarray and bioinformatics analyses, we identified DEGs and determined a comprehensive gene network of active HE4 stimulation in EOC cells. We offered several possible mechanisms and identified therapeutic and prognostic targets of HE4 in EOC.

Introduction

Human epididymis protein 4 (HE4) is a member of the WFDC domain family and encoded by the WFDC2 gene, features the characteristic WAP motif consisting of 8 cysteine-formed disulfide bonds[1]. It was initially discovered in human distal epididymal epithelial cells by Kirchhoff et al. in 1991[2]. In physiological conditions, HE4 is secreted into the blood to act as a protease inhibitor and is involved in the maturation of sperm cells[2]. Several types of cancers are associated with HE4 overexpression both in serum and tissues, and it is a relatively promising and useful biomarker for the diagnosis of ovarian cancer[3–5], primary fallopian tube carcinoma[6], endometrial cancer(combined with CA125)[7], lung cancer[8], breast cancer(combined with miR-127)[9], gastric cancer[1], colorectal cancer[10] and pancreatic adenocarcinomas[11]. In 2008, HE4 was cleared as a serum marker to monitor disease recurrence or progression in patients with epithelial ovarian cancer (EOC) by Food and Drug Administration of USA. Since then, more and more investigators are paying attention to it. Multiple studies found that HE4 is a useful biomarker possessing higher sensitivity and specificity than CA125 in the early confirmatory diagnosis for EOC and differentiation of pelvic masses, especially in combination with the risk of ovarian malignancy algorithm (ROMA)[12, 13]; it seems to be a good predictive factor for the ideal tumor cytoreductive surgery[3], pre-operative prediction of residual disease after interval cytoreduction[4], adjuvant chemotherapy resistance[14] and the possibility of ascites formation[15].

Nevertheless, most of the investigations on HE4 are focusing on its clinical application mentioned above, and there are few studies on its mechanism or function in EOC, whereas the results have not yet reached a consensus. Previous studies have reported that HE4 overexpression significantly promotes tumor cell apoptosis, adhesion and inhibits cell proliferation, migration, and invasiveness[16, 17], but other researchers addressed that high expression of HE4 promotes cell migration, spreading and proliferation[18]. Furtherly, it was found in vitro that HE4 may regulate the mitogen-activated protein kinase (MAPK) and phosphoinositide 3-kinase/AKT signal transduction (PI3K/AKT) pathways to produce tumor-suppressing effect [18, 19]. Recently, emerging studies have been carried out to investigate the association between HE4 and tumorigenesis as well as chemotherapeutic resistance in EOC, whereas the studies were inconclusive due to inconsistencies in results[20–24].

Using gene expression profile detection and bioinformatic analysis to retrieve a large amount of biological information accumulated to a specific gene is a research tool that can help to provide fundamental data for molecular mechanism investigations and to identify new interaction targets. Until now, no microarray profiling study of active HE4 stimulation within epithelial ovarian cancer cells has been performed. In this work, we performed microarray analyses to comprehensively analyze the expression profile of active HE4 stimulation. In total, 713 DEGs, 10 HE4 coexpressed genes, 16 hub genes and 8 MCODEs were found and further analyzed by GO, pathway enrichment and PPI analyses. We found that MAPK signaling pathway, TNF signaling pathway, PI3K-AKt signaling pathway, p53 signaling pathway and cell cycle may be crucial in HE4 stimulation. These verified coexpressed genes and hub genes may help us identify novel biomarkers and treatment targets synergic with HE4 in ovarian cancer in the future.

Materials And Methods

Cell culture, gene transfection and identification

ES-2 cells were purchased from the American Type Culture Collection (ATCC; Manassas, VA, USA) and maintained in RPMI-1640 medium with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin, at 37℃ in a humidified atmosphere with 5% CO_2. Cell culture, shRNA expression vectors construction of HE4, gene transfection and identification were prepared as previously described[25]. Stable cell lines: HE4 shRNAs low-expressing and its empty-plasmid transfected cell lines were labeled as “S” and “S_Mock”, respectively. The untreated cells were labeled as “S_Untreated”. Quantitative real-time polymerase chain reaction (qRT-PCR) and Western blotting were performed as preciously described[25]. The active HE4 protein (recombinant human HE4, rHE4, catalog: MBS355616, MyBioSource) was applied to stimulate S cells (serum-free medium that contained the recombinant HE4 protein 0.2 μg/ml for 24 h)[26] and labeled as “S4”.

Microarrays and bioinformatics analysis

Microarray analysis was performed using triplicate samples of S cells and S4 cells. Total RNA extraction and RNA quantity control were applied and assessed as previously described[25]. The RNA purity and integrity pass criteria were established as A260/A280 ≥ 1.8, A260/A230 ≥ 1 and RIN ≥ 6. gDNA contamination was evaluated by agarose gel electrophoresis. Target preparation and hybridization were performed as previously described[25], the pass criteria for CyDye incorporation efficiency at > 10 dye molecular/1000nt.

Purified RNA samples were subjected to Human Whole Genome OneArray® (Array Version: HOA6.1) with Phalanx hybridization buffer using Phalanx Hybridization System for microarray analysis, the hybridization process was described previously[25]. Fold-changes (FC) were calculated by Rosetta Resolver 7.2 with error model adjusted by Amersham Pairwise Ration Builder for signal comparison of sample. DEGs were identified through volcano plot filtering, and the thresholds for DEGs were |log2FC|≥1 and p value < 0.05, or log2 ratios is “NA” and the differences in intensity between the two samples ≥ 1, 000. Hierarchical clustering was performed using “pheatmap” package in R while the threshold of $\log_2|FC|$ was defined as ≥ 1.5. Gene Ontology (GO) analysis and pathway enrichment were performed using multiple databases, including “clusterProfiler” package in R, DAVID(https://david.ncifcrf.gov), and Metascape[27] (https://metascape.org), using p value < 0.05 as the cut-off thresthold. Gene Set Enrichment Analysis (GSEA, Version 4.0.3) was performed as per software instructions on the comprehensive microarray datasets showing differentially expressed to determine differences and enriched gene sets in Group S4 versus with S [28, 29]. In this study, we focused on the GO biological processes and pathway processes, so the gene sets “c5.bp.v7.1.symbols.gmt”, “c2.cp.biocarta.v7.1.symbols.gmt”, “c2.cp.pid.v7.1.symbols.gmt”, and “c2.cp.reactome.v7.1.symbols.gmt”, which were downloaded from the Molecular Signatures Database (MSigDB) (http://www.broadinstitute.org/gsea/), were used. Enrichment analysis was performed using 1000 phenotype permutations, gene sets with nominal p-value < 0.05, and selecting the weighted scoring scheme with a signal to statistical noise metric to rank genes and complete the GSEA analysis(31).

HE4 (WFDC2) coexpression analysis and validation by bioinformatics

Assessment of the coexpression genes of WFDC2 was performed using the cBioPortal database (http://www. cbioportal.org). The data obtained were RNA-Seq V2 RSEMdata from TCGA PanCancer Atlas that included 585 ovarian serous cystadenocarcinoma tissues. Spearman’s correlation score (≥ 0.2 was considered positively correlated and ≤ − 0.2 was considered negatively correlated with WFDC2) were used to select WFDC2 coexpressed genes. To predict the target genes that were also changed in our microarray, we use Venny online (https://bioinfogp.cnb.csic.es/tools/venny/) to identify the overlapping genes between DEGs and WFDC2 coexpressed genes.

For validation of target genes, the gene expression profile result, GSE51088, submitted by Slamon D, et al.[30] was used. This gene expression profile contains 152 epithelial ovarian cancer patients, 5 benign epithelial ovarian tumor patients and 15 normal healthy ovarian tissues. Based on this data, we calculated Pearson correlation coefficient between target genes and WFDC2, and compared the expression of target genes among malignant, benign and normal ovarian tissues (t test, p < 0.05 as cut-off criterion) by “ggplot2” and “ggpubr” packages in R. The clinical significance of the target genes were evaluated by online Kaplan-Meier survival analysis (http://www.kmplot.com/). A total of 1436 mRNA data samples for PFS and 1657 mRNA data samples for OS of epithelial ovarian cancer were interrogated. The patients were split into 2 groups (high vs. low) based on the expression level.

Validation of target genes by qRT-PCR

For validation, the qRT-PCR were conducted as previously described[25]. 10 target genes were selected, the primers were designed and purchased as previously described, they were presented in Table 1. All reactions were performed in triplicates, and the specificity of PCR amplification was determined by melting point curve analysis.

Table 1
Primers for 10 target genes in qRT-PCR validation

Gene name	Sense primer	Anti-sense primer	Product size(bp)
TMEM200A	ATTGGCAGCAAATACGAT	AAGCACTGATGGACGATG	251
SEC23A	AACACTGGTGCTCGTATC	TTGTAGCAGCTCGATTAG	181
FRMD6	GGACACTCTGGGTTGATT	GTCTTTGGTTTCCGACAT	125
PMP22	CCTCAGGAAATGTCCACCAC	CGCACAGACCAGCAAGAA	192
DNAJB4	TTTGGAAGACGAATGGGTG	TCTTGTTTGAGGCGGGAT	137
APBB2	CGAGCCTAATGCTGGTAA	CAAAGTGTCAATGAGGGATA	184
ERLIN1	ACCGAATAGAAGTGGTTAA	GCACAGCCTGTATAGTGA	246
ZEB1	AAGTGGCGGTAGATGGTA	TGTTGTATGGGTGAAGCA	103
RAB6B	CGGTGGCTGTGGTGGTGTA	CCAGGTCCGTCTTGTTGC	132
PLEKHF1	CAAGTGCGGCTTCGTGGTC	CCTCGTCGGAGTCATCGTCA	227

Protein–protein interaction (PPI) network construction and Identification of Key Module.

Metascape was used to establish a PPI network, and proteins with degree > 1 were selected. The network analyzers “CentiScape” of Cytoscape software was used to analyze the topology property of the network. Genes with a degree of connectivity ≥ 30 were defined as hub genes. “Molecular Complex Detection” (MCODE) in Metascape was used to analyze modules of the PPI network, with the degree cut-off set to 2. The seeds of key modules were identified, and GO analysis and KEGG pathway analysis were performed, finally, their clinical significance was evaluated.

Statistics

Statistical analyses were performed using the SPSS program (Version 24 for Mac; SPSS Inc., Chicago, IL, USA) and Graph Pad Prism 8 (version 8.21. for Mac; Graph Pad Prism Software Inc. San Diego, CA, USA). Quantitative data are presented as Mean ± SD. T test was used for comparison between two groups. “ggplot2”, “ggscatterstats”, “limma”, “pheatmap”, “clusterProfiler” and “enrichplot” packages were used in R language (R, Version 3.6.1; RStudio, Version 1.2.5019). A P-value < 0.05 was considered statistically significant.

Results

HE4 gene transfection identification and RNA quantity assessment

As detected by qRT-PCR and Western blot, the gene and protein expression levels were significantly lower in the HE4 shRNA transfection cells than the Untreated and Mock cells (Fig. 1A and B, all P < 0.01), and there was no statistical difference of HE4 in latter two groups of cells (P > 0.05). The RNA quantity and purity assessment showed that both of the two samples passed the criteria of amplification yield and labeling efficiency (Table 2).

Table 2
The cell line samples description and RNA qualification

Sample	OD260/280	OD260/230	RIN	Results
S	2.02	2.24	10.0	Pass
S4	2.03	2.08	9.4	Pass

Gene expression analysis and clustering

After chip hybridization and data obtaining, volcano analysis displayed the distribution of the 18398 expressed genes (Fig. 1C). Setting log2|FC|≥1 and P-value < 0.05 as cut-off criteria, 713 DEGs were identified, in which 549 genes down-regulated and 164 genes up-regulated (Fig. 1D, the raw data is available in Supplementary Table 1, all gene list is available in Supplementary Table 2). To the DEGs showed log2|FC|>2 differentially expressed, heatmap analysis revealed that 5 DEGs such as EPS15, MSMO1, TMPO, ECT2, ZMYND11 had higher expression levels in S4 cells relative to S cells, and 21 DEGs such as PABPC1, AP3S1, TMX2, PHF6, NR1D2, RAB23, NEK7 had lower expression levels in S4 cells relative to S cells (Fig. 1F).

Gene oncology function analysis of DEGs

We performed gene oncology (GO) enrichment analysis by uploading all the DEGs to “clusterProfiler” package in R to get the biological function. The DEGs were classified into three functional groups: biological process (BP), cellular component (CC) and molecular function (MF). The most enriched BP functions were coenzyme metabolic process and cofactor biosynthetic process. For CC, nuclear speck and nuclear inner membrane were the most enriched. In the clusters of MF, single-stranded RNA binding and translation factor activity were the most enriched (Fig. 2A). For gaining more biological

insight, we applied Metascape[27] to identify BP in which the DEGs participated, the enrichment analysis related to the significant GO terms selected for DEGs was shown as heatmap (Fig. 2B, C), among the diverse pathways highlighted, various are related to oncogenetic, such as: “cell division”, “DNA repair”, “regulation of growth”, “regulation of DNA metabolic process”, “regulation of mitotic cell cycle”, “phosphatidylinositol phosphorylation”, and “signal transduction by p53 class mediator”, etc. These results confirmed the function of HE4 to participate in tumorigenesis and development in epithelial ovarian carcinomas. Interestingly, the “response to wounding” and “cellular response to glucose starvation” pathway found to be modulated by HE4 was associated with immune response regulation and autophagy, which adumbrates the possible capacity of HE4 to drive in the immune mediators’ production and autophagy regulation. For GESA analysis setting GO biological process as gene set, a total of 132 items were enriched with NOM p-val < 0.05, we noticed that the “GO_MAP_KINASE_KINASE_KINASE_ACTIVITY” was the most significantly enriched one, with the highest NES score (NES = 2.166, NOM p-value = 0)(Supplementary Table 3), this indicates that the MAP kinase may participate in the tumorigenesis induced by HE4 activation (Fig. 2D).

Pathway enrichment analysis of DEGs

KEGG pathway enrichment analysis of the 713 DEGs was conducted by using online Metascape[27]. Nineteen KEGG pathways were enriched of the total DEGs with a criterion of Minimal overlap ≥ 3, p value cutoff < 0.01, and Minimal enrichment = 1.5. MAPK signaling pathway, TNF signaling pathway, PI3K-AKt signaling pathway, p53 signaling pathway and cell cycle were highly enriched in the DEGs, in which MAPK signaling pathway was the most significantly enriched (Fig. 2E (1)). In order to analyze and integrate the pathways involved in different gene lists, we divided the DEGs into up-regulation and down-regulation groups and uploaded them in Metascape to conduct a new pathway analysis. The current pathway enrichment analysis includes the pathways currently covered by Metascape: KEGG, Hallmark Gene Sets, Reactome Gene Sets, Canonical, and BioCarta Gene Sets, and the criterion is the same as previously shown. Finally, we found that, MAPK pathway, cell cycle, PI3K AKT mTOR pathways were highly enriched, in which MAPK signaling pathway was also the most significantly enriched in S4 cells compared with S cells (Fig. 2E (2)). To further explore the possibility of MAPK signaling pathway expression in different gene data sets of comprehensive microarray datasets, we used GSEA analysis, and found that MAPK signaling pathway has been significantly enriched in BioCarta, PID, and Reactome Gene Sets in S4 cells compared with S cells (all Nominal p-val < 0.05, Fig. 2F). This analysis suggests that MAPK pathway may be critical in the oncogenesis of HE4 in epithelial ovarian cancer.

HE4 coexpression analysis and validation

HE4 (encoded gene name is “WFDC2”) coexpression genes within epithelial ovarian cancer were identified from cBioPortal which is based on the TCGA Pancancer atlas, including 585 ovarian serous cystadenocarcinoma tissues. In total, 870 WFDC2 coexpressed genes were selected with |Spearman’s Correlation score|>0.2 and p-value < 0.05 as criteria (Supplementary Table 4). In total, 26 genes overlapped between DEGs and WFDC2 coexpressed genes, including 19 HE4 positively correlated and 7 negatively correlated genes(Table 3). To validate these target genes, GEO dataset GSE51088[30], which included 152 epithelial ovarian cancer patients (including 11 epithelial ovarian borderline tumors), 5 normal ovarian tumor patients and 15 normal healthy ovarian tissues, was used to calculate correlation coefficient. Ten genes were dramatically correlated with HE4 expression, including 8 HE4 negatively correlated genes (TMEM220A, SEC23A, FRMD6, PMP22, APBB2, DNAJB4, ERLIN1, and ZEB1), and 2 HE4 positively correlated genes (RAB6B, and PLEKHF1) (Fig. 3A). To further validate microarray data, the ten target genes were subjected to qRT-PCR detection for their differential expressions in S and S4 cells (Fig. 3B). As to these genes, their qRT-PCR data were entirely consistent with gene chip results (all P < 0.05). On the whole, qRT-PCR and gene chip data were correlated to each other in this study.

Table 3
The genes overlapped between DEGs and WFDC2 coexpressed genes generated from cBioPortal.

	DEGs		cBioPortal		
Genesymbol	LogFC	P.value	Spearman's Correlation	p-Value	q-Value
ETV1	-1.74112926	9.21829E-11	-0.25137639	6.07672E-06	0.000513469
TMEM200A	-1.67330712	9.37E-18	-0.23841013	1.84227E-05	0.001075433
SEC23A	-1.66999494	2.15E-24	-0.21612883	0.000107666	0.003532546
MATR3	-1.65897541	1.52E-37	-0.23665619	2.1305E-05	0.001187558
FBXW7	-1.57957489	6.79E-22	-0.2267981	4.72523E-05	0.002057419
EMC2	-1.49126921	6.49E-15	-0.22135379	7.22909E-05	0.002792783
FRMD6	-1.39540806	1.27E-20	-0.24639449	9.37318E-06	0.000667118
GTF2H3	-1.36037738	3.36E-14	-0.20451604	0.000252306	0.006047828
PMP22	-1.34490578	1.40243E-11	-0.20348488	0.00027152	0.006304242
IREB2	-1.29329478	1.96E-18	-0.25242374	5.54111E-06	0.000488278
TM4SF18	-1.26331644	0.001843692	-0.20174481	0.000307069	0.006797924
BLOC1S6	-1.2311467	2.83E-13	-0.23064443	3.4772E-05	0.001653979
APBB2	-1.21372596	3.68573E-05	-0.25358365	5.00049E-06	0.000462079
ZFAND6	-1.20997926	1.34621E-05	-0.20588664	0.000228727	0.005643468
DNAJB4	-1.18965026	4.7754E-11	-0.22039349	7.78373E-05	0.002944791
RANBP2	-1.16466642	1.6059E-09	-0.2361569	2.22005E-05	0.001214792
ERLIN1	-1.14961785	8.98E-17	-0.20737172	0.000205512	0.005308028
SLK	-1.08761539	3.19422E-09	-0.23495397	2.45069E-05	0.001288354
ZEB1	-1.01529781	1.18109E-09	-0.24887479	7.56265E-06	0.000576294
MKNK2	1.03904918	2.28E-16	0.22219649	6.77323E-05	0.00265548
SMARCD3	1.04293745	4.82E-16	0.26206354	2.32529E-06	0.00027583
RAB6B	1.06900496	2.2E-14	0.20381363	0.000265251	0.006227877
NOXA1	1.07467178	7.95927E-06	0.2024282	0.000292619	0.006587495
IQSEC2	1.10423888	9.16904E-08	0.24086495	1.50035E-05	0.000929112
PLEKHF1	1.11966989	0.001304667	0.38517331	1.29E-12	2.98E-09
RAPGEF3	1.30138476	3.89773E-11	0.24899116	7.48646E-06	0.000574842

To evaluate the clinical significance of 10 target genes, we screened the correlation of tumor type with these ten genes in GESE51088, we find that all of these ten genes were significantly correlated with tumor types, in which TMEM220A, SEC23A, FRMD6, PMP22, APBB2, DNAJB4, ERLIN1, and ZEB1were obviously downregulated in ovarian cancers compared with ovarian benign tumors and normal tissues, and RAB6B, PLEKHF1 were obviously upregulated in ovarian cancers compared with ovarian benign tumors and normal tissues (Fig. 4A). Kaplan-meier survival analysis were generated for a large cohort of ovarian cancer. In total, data from 1657 epithelial ovarian cancer patients were interrogated and hazard ratios (HR) and p-values for statistical significance were determined. The data are summarized in Table 4. Interestingly, all of the 10 target genes correlated with overall prognosis (Fig. 4B). These 10 target genes should be further investigated to explore their association with HE4 expression, which might expose the mechanism of HE4 during the development of epithelial ovarian cancer.

Table 4
Validation of 10 WFDC2 co-expressed genes among DEGs

Gene symbol	Microarray analysis		cBioPortal (585 Epithelial ovarian tissues)		GSE51088		Survival Curve (1657 EOC tissues)	
	LogFC	-log10(P.value)	Spearman's Correlation	-log10(P.value)	Pearson score R	-log10(P.value)	HR	LogRank p-value
TMEM200A	-1.6733071	17.0282604	-0.2384101	4.73463965	-0.5306474	13.1530447	1.50 (1.20–1.88)	3e-04
SEC23A	-1.6699949	23.6675615	-0.2161288	3.96790529	-0.3638438	6.03198429	1.25 (1.09–1.42)	0.0011
FRMD6	-1.3954081	19.8961963	-0.2463945	5.02811212	-0.2713641	3.49854362	1.59 (1.29–1.95)	7.5e-06
PMP22	-1.3449058	10.8531188	-0.2034849	3.56619818	-0.4983522	11.4571746	1.45 (1.27–1.65)	3e-08
APBB2	-1.213726	4.43347648	-0.2535836	5.30098657	-0.2996121	4.18508682	1.27 (1.09–1.48)	0.0022
DNAJB4	-1.1896503	10.3209902	-0.2203935	4.10881391	-0.4288876	8.36066295	1.31 (1.14–1.51)	0.00017
ERLIN1	-1.1496178	16.0467237	-0.2073717	3.68716704	-0.2856046	3.83564714	0.87 (0.76–0.99)	0.034
ZEB1	-1.0152978	8.92771701	-0.2488748	5.12132313	-0.4347875	8.59749789	1.60 (1.28–1.99)	2.2e-05
RAB6B	1.06900496	13.6575773	0.20381363	3.57634461	0.2067661	2.18706059	1.34 (1.05-.170)	0.017
PLEKHF1	1.11966989	2.88450032	0.38517331	11.8894103	0.2926423	4.00899113	1.17 (1.02–1.35)	0.03

Protein-protein interaction (PPI) network and modular analysis

The PPI enrichment analysis was carried out in Metascape online. The resultant network contains the subset of proteins that form physical interactions with at least one other member in the gene list. The “centiscape” plug-in in Cytoscape was used to find out the hub genes. Proteins with degree > 1 were selected. In total, 289 nodes (40.4% of all 713 DEGs) and 2942 PPI relation- ships were obtained (Fig. 5A). Sixteen genes with a degree of connectivity > 30 were defined as hub genes for HE4 activation (Table 5). According to the degree rank, the sixteen hub genes included HSPA1B, HSPA1A, SUMO1, CDK1, MAX, PABPC1, MAGOH, HNRNPU, YWHAG, RANBP2, SRSF1, CNBP, U2AF2, RNPS1, SMAD3, and POLR2D. These hub genes could interact with each other, which suggest that these hub genes might play an important role in HE4 activation and should be further studied in EOC.

Table 5
16 hub genes (degree ≥ 30) selected from PPI network according to the degree

Gene Symbol	Eccentricity	Closeness	Betweenness	Degree	Stress
HSPA1B	0.25	0.00198413	12720.52992	93	91258
HSPA1A	0.25	0.00193798	9887.649848	87	77174
SUMO1	0.25	0.00182482	12090.55822	74	77170
CDK1	0.25	0.00166945	7451.126297	59	44320
MAX	0.2	0.00168634	11030.17839	57	61586
PABPC1	0.25	0.00167224	3561.121821	46	26550
MAGOH	0.25	0.00162602	1812.074639	39	16986
HNRNPU	0.25	0.00162866	2527.645484	37	17974
YWHAG	0.25	0.0016129	2749.807044	36	20970
RANBP2	0.25	0.00158479	2066.126067	35	18706
SRSF1	0.25	0.00141844	1366.01731	35	10966
CNBP	0.25	0.0015748	1558.185614	34	14340
U2AF2	0.2	0.00144718	901.9806448	33	7256
RNPS1	0.25	0.00152207	514.6918118	31	6640
SMAD3	0.2	0.00160514	3586.354138	31	23532
POLR2D	0.25	0.0015674	760.9759902	30	8880

Here in this paper, the MCODE algorithm[31] has been implemented in Metascape to distinguish densely connected network components, and finally, 8 modules were identified and sequenced with the descending of Score from 8.65 to 1.00 (Supplementary Table 5). Five most significant MCODEs were extracted when the Score > 2 (Fig. 5B).

To further explore the biological function of the seeds in the five key modules (SMNDC1, HSPA1A, FNBP1L, GAR1, and SKA2), functional enrichment analysis was performed based on the Metascape by setting as Min Overlap = 2 and P value cutoff = 0.05. Regarding the GO terms, the main enriched ones were regulation of microtubule polymerization or depolymerization, and positive regulation of cellular component biogenesis. The pathway signaling analysis showed marked enrichment of cell cycle, and spliceosome (Fig. 5C).

To determine the clinical significance of 5 seed genes of the key modules, Kaplan-meier survival analysis were generated for a large cohort of ovarian cancer. In total, data from 1657 epithelial ovarian cancer patients for OS and 1436 patients for PFS were interrogated and hazard ratios (HR) and p-values for statistical significance were determined. As shown in Fig. 5D, all of the 5 seed genes obviously correlated with OS and PFS prognosis.

Discussion

Ovarian cancer is the seventh most common malignant tumor in the world. In 2012, it was estimated that there were 238,719 incident cases and the age-standardized rate was 6.1/100,000[32]. In 2020, it is expected by the American Cancer Society that there will be approximately 21,750 new ovarian cancer cases in the United States, and 13,9400 women will die from it. Owing to its occult onset and innocuous symptoms, most of the patients with ovarian cancer are diagnosed in advanced stage. Although the development of new anti-tumor drugs and the improvement of surgical treatment, the survival rates decline dramatically from 92% for patients with Stage I to 17–28% for those with advanced disease (Stages III-IV)[33]; majority of advanced stages patients eventually relapse and chemotherapeutic resistance. Although serum biomarker CA125 was widely used in clinical practice for diagnosis and differentiation, population-based screening serum cancer antigen (CA125) assessment and use of risk for ovarian cancer algorithm (ROCA) did not identify significant mortality reduction and has been proved to be ineffective[34]. Thus, it is urgent to clarify the underlying mechanism of oncogenesis of EOC and find out tumor biomarkers to facilitate early diagnosis and targeted therapy or prevention.

As a new tumor biomarker, HE4 has aroused full attention in recent years. However, most of the research focuses on its clinical application of early and differential diagnosis, relapse, prognosis, chemotherapeutic resistance, as well as other clinical aspects for EOC[12, 35], and there are few studies on its mechanism in ovarian cancer yet; it may be the reason that HE4 has not to be anchored as a therapeutic target due to a unanimous conclusion on its roles in the tumorigenesis and progression of EOC. As early as in 2011, Gao L. et al. [16] reported that they found overexpression of HE4 obviously promoted ovarian cancer cell apoptosis and adhesion, they noticed HE4 may inhibite ovarian cancer cell proliferation, migration and invasiveness, as well as xenograft tumor formation in vivo; thus they concluded that HE4 might play a protecting role in the progression of EOC. Further, in 2014, Kong et al.[19] found in vitro that this protective infulence may be attained by regulating the MAPK and PI3K/AKT pathways. On the contrary, other researchers noted that HE4 high expression promotes cell migration, adhesion, proliferation, and spreading, which can be associated with its effects on the EGFR-MAPK signaling pathway[18, 36]. What is more, HE4 contains fucosylated modification (Lewis y antigen)[37], Lewis y overexpression can promote HE4-

mediated invasion and metastasis in ovarian cancer cells[38]. Impressively, overexpression of Lewis y antigen enhanced tyrosine phosphorylation of EGFR and HER/neu, which improved cell proliferation by the PI3K/Akt and Raf/MEK/MAPK pathways[39]; hence, Lewis y antigen and HE4 may affect alike signaling pathways that promote tumor growth and malignancy[40]. HE4 overexpression promotes ovarian cancer cell xenograft tumor growth in vivo, antisense target of HE4 can suppress this effect, HE4 interacts with tumor microenvironment constituents (EGFR, IGF1R, Insulin) and transcription factor HIF1α, these results provide some convincing proof that HE4 is tied to growth factor signal and the MAPK/ERK pathway[41]. Annexin A2 (ANXA2) was identified as a robust interacting partner of HE4 by mass spectrometry and co-immunoprecipitation, the HE4-ANXA2 complex can promote ovarian cancer cell invasion and migration in vitro and tumor distant metastasis of lung in vivo, downregulation of HE4 decreases expression of MKNK2 and LAMB2, which were associated with signaling pathways of MAPK and focal adhesion[5].

In recent years, a growing number of investigations have gradually found that HE4 promotes cell proliferation, adhesion, invasion, migration, and chemoresistance in ovarian cancer[20–26, 42–47]. It was found that HE4 overexpression or rHE4 treatment in EOC cells resulted in upregulation of many transcripts coding for extracellular matrix proteins, including LAMC2, LAMB3, SERPINB2 and GREM1; moreover, in cells overexpressing HE4 or exposed to rHE4 in culture medium, the protein levels of LAMC2 and LAMB3 were continously increased, and in the presence of fibronectin, the focal adhesions were elevated in cells treated with rHE4[22].

It was known that ovarian cancer participates in evading immunosurveillance and orchestrating a suppressive immune microenvironment, a series of studies by James NE, et al.[23, 44, 45] found that, upon exposure of purified human peripheral blood mononuclear cells(PBMCs) to HE4, osteopontin (OPN) and DUSP6 appeared as the most inhibited and upregulated genes; the proliferation of human ovarian carcinoma cells in conditioned media from HE4-exposed PBMCs was enhanced, while the effect was attenuated by adding recombinant OPN or OPN-inducible cytokines (IL-12 and IFN-γ); HE4 can compromise both OPN-mediated T cell activation[44] and cytotoxic $CD8^+/CD56^+$ cells through upregulation of self-produced DUSP6[45], thus promoting the tumorigenesis of ovarian cancer[23, 44, 45, 48]. Other researchers found that HE4 promotes carcinogenesis of ovarian cancer by combining with histone deacetylase 3 (HDAC3) to activate PI3K/AKT pathway[46], and that HE4 knockdown suppresses the invasive cell growth and malignant progress of ovarian cancer by inhibiting JAK/STAT3 pathway[24]. Until now, a few studies have begun to delineate HE4's role in chemoresistance of ovarian cancer. It was noted that overexpression of HE4 promotes the collateral resistance of ovarian cancer cells to cisplatin and paclitaxel, and down-regulation of HE4 partially reverses the resistance of multiple chemotherapeutic agents; the HE4-mediated chemoresistance might be related to a variety of factors, including deregulation of MAPK signaling (EGR1 and p38 inhibition), and alterations of tubulin levels or stability; recombinant HE4 could upregulate the levels of α-tubulin, β-tubulin and microtubule associated protein tau (MAPT) [41, 43]. Similarly, in vitro, HE4 represses apoptosis induced by carboplatin, and recombinant HE4 results in increased BCL-2 expression and decreased Bax (Bcl-2 associated X protein) expression in carboplatin treated ovarian cancer cells, which reduces the ratio of Bax/Bcl-2; in addition, HE4 also suppresses EGR1 expression, which may contribute to the overall reduction of pro-apoptotic factors that lead to EOC chemoresistance[26]. HE4 can enhance the regulation of DUSP6 and they are positively correlated, DUSP6 deactivates extracellular-signal-regulated kinase (ERK), the inhibition of DUSP6 can alter gene expression of ERK pathway response genes (EGR1 and c-JUN) and sensitize ovarian cancer cells to chemotherapeutic agents(paclitaxel or carboplatin)[23, 48]. The resensitization of ovarian cancer cells to cisplatin and paclitaxel caused by HE4 knockdown is due to the corresponding decreases of ERK and AKT during gene knockouts, and the activation of these pathways inhibits the apoptotic signal of tumor cells[21].

Conclusions

Token together, mechanism underlying HE4's contribution to tumorigenesis, progression and chemoresistance in EOC has not been sufficiently established. Therefore, the microarray analysis on HE4, which can provide high-throughput data for accurate molecular function research, has become particularly essential. In this study, we analyzed the gene expression profile change after active HE4 protein stimulation in EOC cells, we tried to decipher the cellular biological processes by using the online tools for delineating the pathways and interaction network enriched. We found 713 DEGs (164 up regulated and 549 down regulated), pathway enrichment showed that MAPK pathway accounted for a significant portion of the enriched terms. WFDC2 coexpression analysis revealed ten WFDC2 coexpressed genes (TMEM220A, SEC23A, FRMD6, PMP22, APBB2, DNAJB4, ERLIN1, ZEB1, RAB6B, and PLEKHF1) that were also dramatically changed in S4 cells and validated by dataset GSE51088. Kaplan–Meier survival statistics revealed clinical significance for all of the 10 target genes. Finally, PPI was constructed, sixteen hub genes and eight molecular complex detections (MCODEs) were identified, the seeds of five most significant MCODEs were subjected to GO and KEGG enrichment analysis and their clinical significance was evaluated. These data have not been mentioned in previous studies, which can offer a new approach to further clarify the mechanism of HE4 in the oncogenesis and chemoresistance of EOC.

Declarations

Availability of data and materials

The datasets used and/or analyzed in the present study are available from the supplementary files or the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Not applicable.

Funding

The present study was supported by the National Natural Science Foundation of China (grant nos. 81472437 and 81672590), National Natural Science Foundation of China Youth Science Foundation (No. 81602438), Doctoral start-up fund of Liaoning Province (Grant No. 201601133) and 345 Talent Project of Shengjing Hospital.

Authors' contributions

LC Z and B L conceived and designed the idea to this paper; LC Z and HY X collected and analyzed the data, and drafted the paper; LC Z and MZ T analyzed the data and revised the final paper. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank everyone who take part in this study.

References

1. Peng C, Liu G, Huang K, Zheng Q, Li Y, Yu C: **Hypoxia-Induced Upregulation of HE4 Is Responsible for Resistance to Radiation Therapy of Gastric Cancer**. *Mol Ther Oncolytics* 2019, **12**:49-55.
2. Kirchhoff C, Habben I, Ivell R, Krull N: **A major human epididymis-specific cDNA encodes a protein with sequence homology to extracellular proteinase inhibitors**. *Biol Reprod* 1991, **45**(2):350-357.
3. Chen WT, Gao X, Han XD, Zheng H, Guo L, Lu RQ: **HE4 as a serum biomarker for ROMA prediction and prognosis of epithelial ovarian cancer**. *Asian Pac J Cancer Prev* 2014, **15**(1):101-105.
4. Lof P, van de Vrie R, Korse CM, van Driel WJ, van Gent M, Karlsen MA, Amant F, Lok CAR: **Pre-operative prediction of residual disease after interval cytoreduction for epithelial ovarian cancer using HE4**. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* 2019, **29**(8):1304-1310.
5. Zhuang H, Tan M, Liu J, Hu Z, Liu D, Gao J, Zhu L, Lin B: **Human epididymis protein 4 in association with Annexin II promotes invasion and metastasis of ovarian cancer cells**. *Molecular cancer* 2014, **13**:243.
6. Mi D, Zhang YX, Wang CJ, Feng Q, Qi P, Chen SQ: **Diagnostic and prognostic value of serum human epididymis protein 4 in patients with primary fallopian tube carcinoma**. *J Obstet Gynaecol Res* 2016, **42**(10):1326-1335.
7. Huang GQ, Xi YY, Zhang CJ, Jiang X: **Serum Human Epididymis Protein 4 Combined with Carbohydrate Antigen 125 for Endometrial Carcinoma Diagnosis: A Meta-Analysis and Systematic Review**. *Genet Test Mol Biomarkers* 2019, **23**(8):580-588.
8. He YP, Li LX, Tang JX, Yi L, Zhao Y, Zhang HW, Wu ZJ, Lei HK, Yu HQ, Nian WQ *et al*: **HE4 as a biomarker for diagnosis of lung cancer: A meta-analysis**. *Medicine (Baltimore)* 2019, **98**(39):e17198.
9. Lu M, Ju S, Shen X, Wang X, Jing R, Yang C, Chu H, Cong H: **Combined detection of plasma miR-127-3p and HE4 improves the diagnostic efficacy of breast cancer**. *Cancer Biomark* 2017, **18**(2):143-148.
10. Shi LP, Guo HL, Su YB, Zheng ZH, Liu JR, Lai SH: **MicroRNA-149 sensitizes colorectal cancer to radiotherapy by downregulating human epididymis protein 4**. *Am J Cancer Res* 2018, **8**(1):30-38.
11. O'Neal RL, Nam KT, LaFleur BJ, Barlow B, Nozaki K, Lee HJ, Kim WH, Yang HK, Shi C, Maitra A *et al*: **Human epididymis protein 4 is up-regulated in gastric and pancreatic adenocarcinomas**. *Hum Pathol* 2013, **44**(5):734-742.
12. Zhang L, Chen Y, Wang K: **Comparison of CA125, HE4, and ROMA index for ovarian cancer diagnosis**. *Curr Probl Cancer* 2019, **43**(2):135-144.
13. Brenk A, Bodzek P, Balis M, Barbachowska A, Janosz I, Olejek A: **Usefulness of HE4 protein in differentiation of pelvic masses in woman**. *Prz Menopauzalny* 2019, **18**(1):27-32.
14. Aarenstrup Karlsen M, Hogdall C, Nedergaard L, Philipsen Prahm K, Schou Karlsen NM, Weng Ekmann-Gade A, Henrichsen Schnack T, Svenstrup Poulsen T, Jarle Christensen I, Hogdall E: **HE4 as a predictor of adjuvant chemotherapy resistance and survival in patients with epithelial ovarian cancer**. *APMIS* 2016, **124**(12):1038-1045.

15. Liu D, Kong D, Li J, Gao L, Wu D, Liu Y, Yang W, Zhang L, Zhu J, Jin X: **HE4 level in ascites may assess the ovarian cancer chemotherapeutic effect**. *J Ovarian Res* 2018, **11**(1):47.
16. Gao L, Cheng HY, Dong L, Ye X, Liu YN, Chang XH, Cheng YX, Chen J, Ma RQ, Cui H: **The role of HE4 in ovarian cancer: inhibiting tumour cell proliferation and metastasis**. *J Int Med Res* 2011, **39**(5):1645-1660.
17. Dai C, Zheng Y, Li Y, Tian T, Wang M, Xu P, Deng Y, Hao Q, Wu Y, Zhai Z *et al*: **Prognostic values of HE4 expression in patients with cancer: a meta-analysis**. *Cancer Manag Res* 2018, **10**:4491-4500.
18. Lu R, Sun X, Xiao R, Zhou L, Gao X, Guo L: **Human epididymis protein 4 (HE4) plays a key role in ovarian cancer cell adhesion and motility**. *Biochemical and biophysical research communications* 2012, **419**(2):274-280.
19. Kong X, Chang X, Cheng H, Ma R, Ye X, Cui H: **Human epididymis protein 4 inhibits proliferation of human ovarian cancer cells via the mitogen-activated protein kinase and phosphoinositide 3-kinase/AKT pathways**. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* 2014, **24**(3):427-436.
20. Zhu L, Zhuang H, Wang H, Tan M, Schwab CL, Deng L, Gao J, Hao Y, Li X, Gao S *et al*: **Overexpression of HE4 (human epididymis protein 4) enhances proliferation, invasion and metastasis of ovarian cancer**. *Oncotarget* 2016, **7**(1):729-744.
21. Lee S, Choi S, Lee Y, Chung D, Hong S, Park N: **Role of human epididymis protein 4 in chemoresistance and prognosis of epithelial ovarian cancer**. *J Obstet Gynaecol Res* 2017, **43**(1):220-227.
22. Ribeiro JR, Gaudet HM, Khan M, Schorl C, James NE, Oliver MT, DiSilvestro PA, Moore RG, Yano N: **Human Epididymis Protein 4 Promotes Events Associated with Metastatic Ovarian Cancer via Regulation of the Extracelluar Matrix**. *Front Oncol* 2017, **7**:332.
23. James NE, Beffa L, Oliver MT, Borgstadt AD, Emerson JB, Chichester CO, Yano N, Freiman RN, DiSilvestro PA, Ribeiro JR: **Inhibition of DUSP6 sensitizes ovarian cancer cells to chemotherapeutic agents via regulation of ERK signaling response genes**. *Oncotarget* 2019, **10**(36):3315-3327.
24. Wang A, Jin C, Tian X, Wang Y, Li H: **Knockdown of HE4 suppresses aggressive cell growth and malignant progression of ovarian cancer by inhibiting the JAK/STAT3 pathway**. *Biol Open* 2019, **8**(9).
25. Zhu L, Guo Q, Jin S, Feng H, Zhuang H, Liu C, Tan M, Liu J, Li X, Lin B: **Analysis of the gene expression profile in response to human epididymis protein 4 in epithelial ovarian cancer cells**. *Oncol Rep* 2016, **36**(3):1592-1604.
26. Wang H, Zhu L, Gao J, Hu Z, Lin B: **Promotive role of recombinant HE4 protein in proliferation and carboplatin resistance in ovarian cancer cells**. *Oncol Rep* 2015, **33**(1):403-412.
27. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK: **Metascape provides a biologist-oriented resource for the analysis of systems-level datasets**. *Nat Commun* 2019, **10**(1):1523.
28. Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstråle M, Laurila E *et al*: **PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes**. *Nature Genetics* 2003, **34**(3):267-273.
29. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al*: **Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles**. *Proceedings of the National Academy of Sciences* 2005, **102**(43):15545-15550.
30. Karlan BY, Dering J, Walsh C, Orsulic S, Lester J, Anderson LA, Ginther CL, Fejzo M, Slamon D: **POSTN/TGFBI-associated stromal signature predicts poor prognosis in serous epithelial ovarian cancer**. *Gynecol Oncol* 2014, **132**(2):334-342.
31. Bader GD, Hogue CW: **An automated method for finding molecular complexes in large protein interaction networks**. *BMC Bioinformatics* 2003, **4**:2.
32. Zhang Y, Luo G, Li M, Guo P, Xiao Y, Ji H, Hao Y: **Global patterns and trends in ovarian cancer incidence: age, period and birth cohort analysis**. *BMC Cancer* 2019, **19**(1):984.
33. Doubeni CA, Doubeni AR, Myers AE: **Diagnosis and Management of Ovarian Cancer**. *Am Fam Physician* 2016, **93**(11):937-944.
34. Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, Amso NN, Apostolidou S, Benjamin E, Cruickshank D *et al*: **Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial**. *Lancet* 2016, **387**(10022):945-956.
35. Dochez V, Randet M, Renaudeau C, Dimet J, Le Thuaut A, Winer N, Thubert T, Vaucel E, Caillon H, Ducarme G: **Efficacy of HE4, CA125, Risk of Malignancy Index and Risk of Ovarian Malignancy Index to Detect Ovarian Cancer in Women with Presumed Benign Ovarian Tumours: A Prospective, Multicentre Trial**. *J Clin Med* 2019, **8**(11).
36. Chen Y, Mu X, Wang S, Zhao L, Wu Y, Li J, Li M: **WAP four-disulfide core domain protein 2 mediates the proliferation of human ovarian cancer cells through the regulation of growth- and apoptosis-associated genes**. *Oncol Rep* 2013, **29**(1):288-296.
37. Zhuang H, Gao J, Hu Z, Liu J, Liu D, Lin B: **Co-expression of Lewis y antigen with human epididymis protein 4 in ovarian epithelial carcinoma**. *PloS one* 2013, **8**(7):e68994.
38. Zhuang H, Hu Z, Tan M, Zhu L, Liu J, Liu D, Yan L, Lin B: **Overexpression of Lewis y antigen promotes human epididymis protein 4-mediated invasion and metastasis of ovarian cancer cells**. *Biochimie* 2014, **105**:91-98.
39. Liu JJ, Lin B, Hao YY, Li FF, Liu DW, Qi Y, Zhu LC, Zhang SL, Iwamori M: **Lewis(y) antigen stimulates the growth of ovarian cancer cells via regulation of the epidermal growth factor receptor pathway**. *Oncol Rep* 2010, **23**(3):833-841.
40. James NE, Chichester C, Ribeiro JR: **Beyond the Biomarker: Understanding the Diverse Roles of Human Epididymis Protein 4 in the Pathogenesis of Epithelial Ovarian Cancer**. *Front Oncol* 2018, **8**:124.

41. Moore RG, Hill EK, Horan T, Yano N, Kim K, MacLaughlan S, Lambert-Messerlian G, Tseng YD, Padbury JF, Miller MC *et al*: **HE4 (WFDC2) gene overexpression promotes ovarian tumor growth**. *Scientific reports* 2014, **4**:3574.
42. Zhu L, Gao J, Hu Z, Schwab C, Zhuang H, Tan M, Yan L, Liu J, Zhang D, Lin B: **Membranous expressions of Lewis y and CAM-DR-related markers are independent factors of chemotherapy resistance and poor prognosis in epithelial ovarian cancer**. *Am J Cancer Res* 2015, **5**(2):830-843.
43. Ribeiro JR, Schorl C, Yano N, Romano N, Kim KK, Singh RK, Moore RG: **HE4 promotes collateral resistance to cisplatin and paclitaxel in ovarian cancer cells**. *J Ovarian Res* 2016, **9**(1):28.
44. James NE, Cantillo E, Oliver MT, Rowswell-Turner RB, Ribeiro JR, Kim KK, Chichester CO, 3rd, DiSilvestro PA, Moore RG, Singh RK *et al*: **HE4 suppresses the expression of osteopontin in mononuclear cells and compromises their cytotoxicity against ovarian cancer cells**. *Clin Exp Immunol* 2018, **193**(3):327-340.
45. James NE, Oliver MT, Ribeiro JR, Cantillo E, Rowswell-Turner RB, Kim KK, Chichester CO, 3rd, DiSilvestro PA, Moore RG, Singh RK *et al*: **Human Epididymis Secretory Protein 4 (HE4) Compromises Cytotoxic Mononuclear Cells via Inducing Dual Specificity Phosphatase 6**. *Front Pharmacol* 2019, **10**:216.
46. Lou T, Zhuang H, Liu C, Zhang Z: **HDAC3 positively regulates HE4 expression to promote ovarian carcinoma progression**. *Arch Biochem Biophys* 2019, **675**:108044.
47. Wang J, Deng L, Zhuang H, Liu J, Liu D, Li X, Jin S, Zhu L, Wang H, Lin B: **Interaction of HE4 and ANXA2 exists in various malignant cells-HE4-ANXA2-MMP2 protein complex promotes cell migration**. *Cancer Cell Int* 2019, **19**:161.
48. Sato S, Itamochi H: **Dual specificity phosphatase 6 as a new therapeutic target candidate for epithelial ovarian cancer**. *Ann Transl Med* 2019, **7**(Suppl 8):S373.

Figures

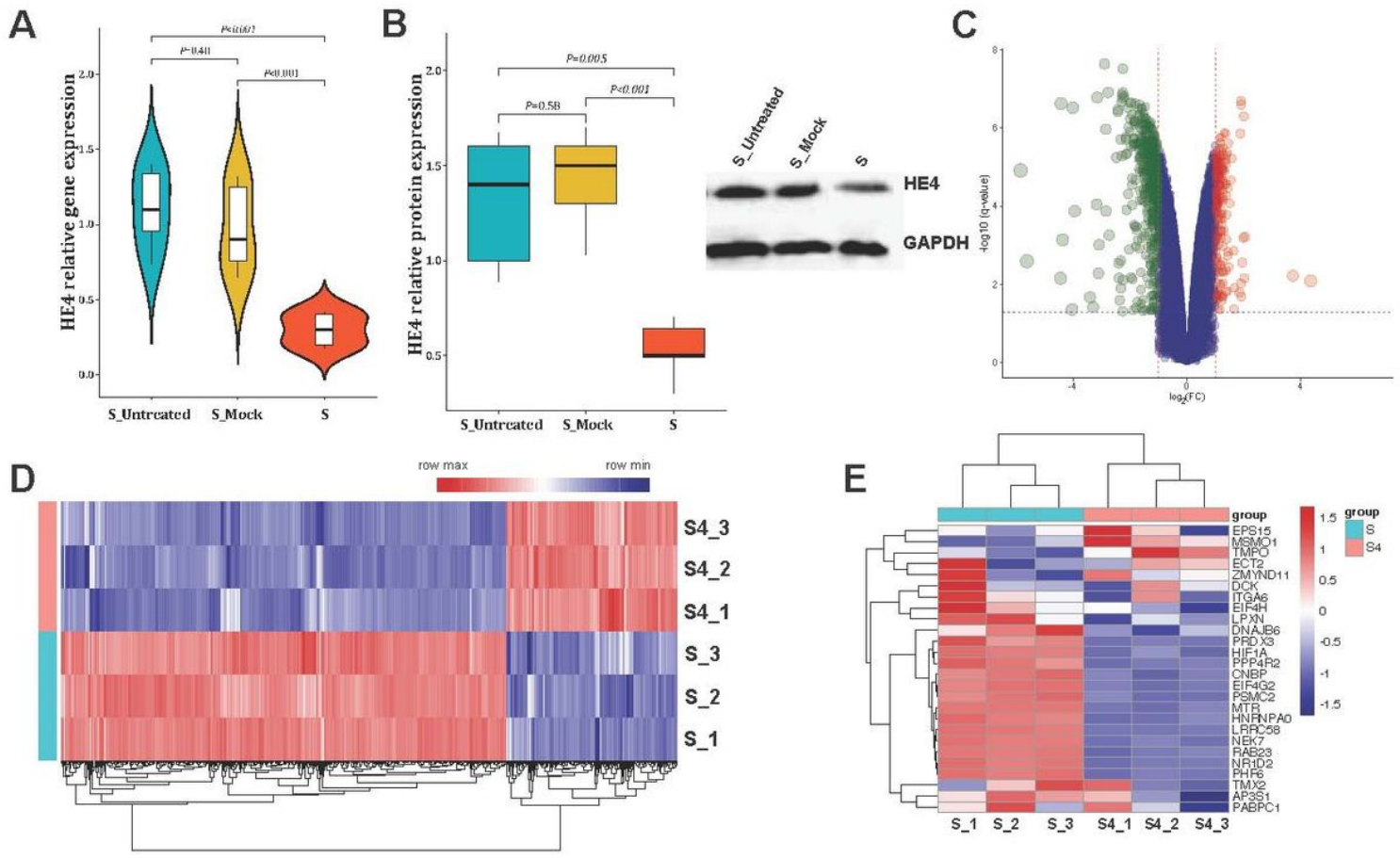

Figure 1

Confirmation of transfection in ovarian cancer cell line ES-2 and Identification of differentially expressed genes (DEGs) in response to active HE4 stimulation in ES-2 cells A. Quantitative real-time PCR results showing the gene expression of HE4 after HE4 shRNA transfection in ovarian cancer cell line ES-2. B. Western blot staining results showing the protein expression of HE4 after HE4 shRNA transfection. The plot is partly from our previously published paper[5]. C. Volcano plot of the 18398 expressed genes. Red color represented up-regulated genes in response to HE4 activation in ES-2 cells and green color represented down-regulated genes. The gene symbol names of the DEGs showing log2|Fold change|≥ 3 were listed in plot. D. Clustering

heatmap of the 713 genes exhibiting significantly differential expression, statistically significant DEGs were defined as log2|Foldchange|≥ 1 and p-value < 0.05. E. Heatmap showing the representative DEGs in ovarian cancer cells. The color key from blue to red indicates low to high gene expression, respectively.

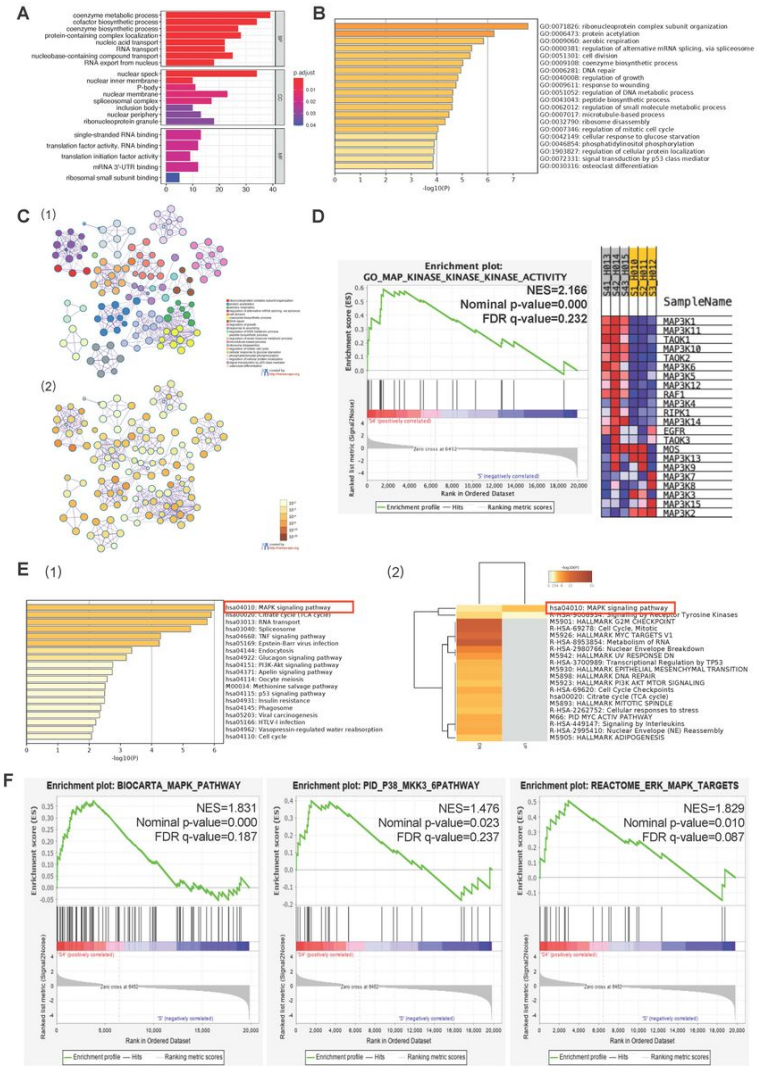

Figure 2

Gene ontology (GO) and pathway enrichment of 713DEGs A. GO enrichment analysis of 713 DEGs in “clusterProfiler” package of R, sorted by the ascending order of the adjusted p value of the GO list. BP, biological process; CC, cellular component; MF, molecular function. B. Heatmap showing GO enrichment of all the DEGs generated by Metascape and sorted by value of –Log10(p-value). The GO ID items were shown in the right panel. C. (1) Networks layout of the clusters generated with the complete DEGs regulated by HE4 in ovarian cancer cells, where each node represents an enriched term and is colored by its cluster ID, its size is proportional to the number of input genes falling into that term. One term from each cluster is selected to have its term description shown as label. (2) Network of enriched terms colored by P-value, where terms containing more genes tend to have a more significant P-value. D. GSEA shows that the ‘GO_MAP_KINASE_KINASE_KINASE_ACTIVITY’ gene-set comprising of genes specifically up-regulated in S4 cells is enriched as compared to the S cells. The heatmap of genes that are related to the analyzed GeneSet is showed in the right panel. E. (1)KEGG Pathway enriched by all the DEGs generated by Metascape and sorted by value of –Log10(p-value). (2) Pathway enrichment analysis (gene sets including: KEGG, Hallmark Gene Sets, Reactome Gene Sets, Canonical, and BioCarta Gene Sets) by up regulated and down regulated gene lists generated by Meatscape and sorted by value of –Log10(p-value). The pathway ID items were shown in the right panel. F. GSEA shows that the “BIOCARTA_MAPK_PATHWAY”, “PID_P38_MKK3_6ATHWAY”, and “REACTOME_ERK_MAPK_TARGETS” gene-sets comprising of genes specifically up-regulated in S4 cells are enriched as compared to the S cells.

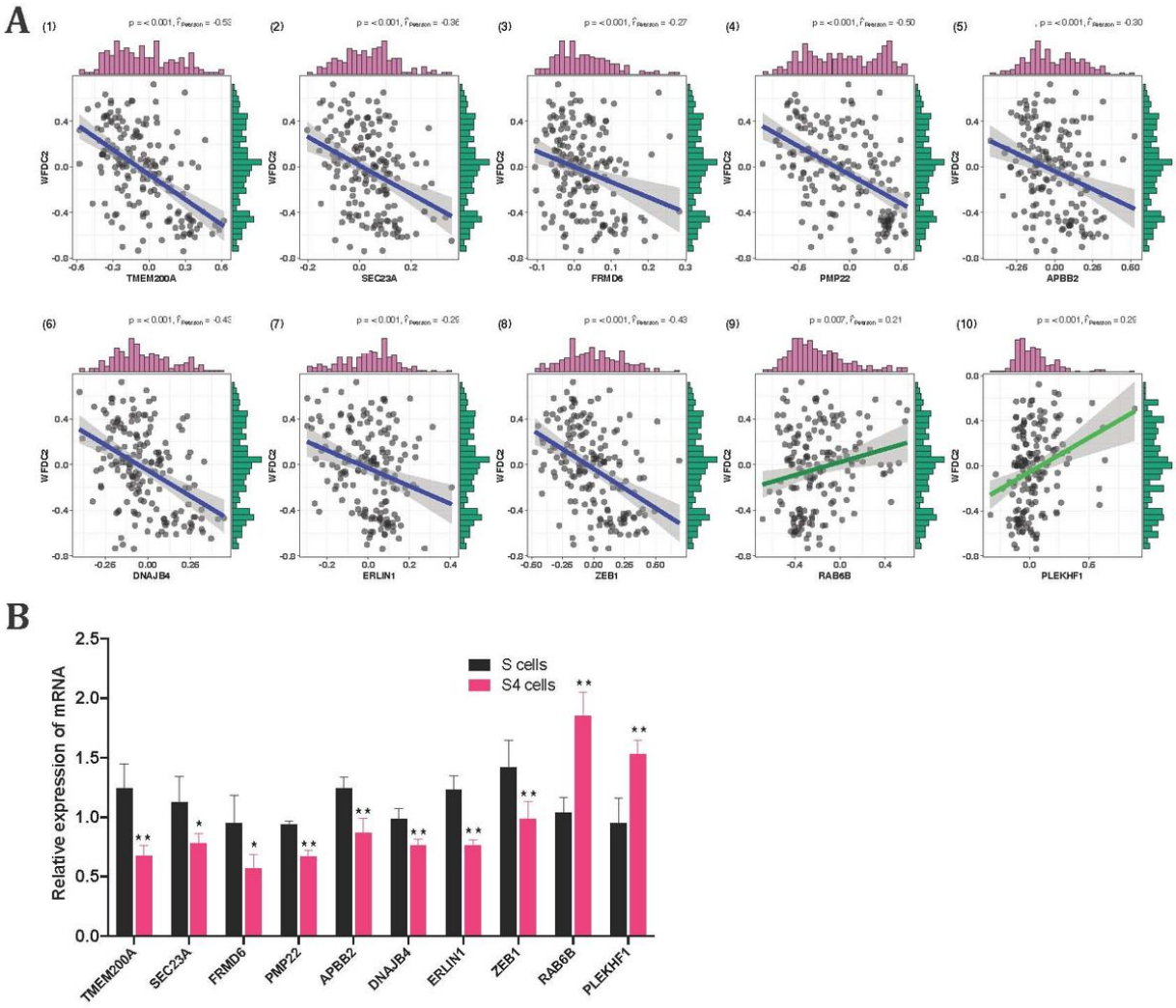

Figure 3

Correlation analysis of WFDC2 co-expressed genes and mRNA validation A. Pearson's correlation analysis using GSE51088 data. (1)–(8) were 8 WFDC2 negatively correlated genes (TMEM220A, SEC23A, FRMD6, PMP22, APBB2, DNAJB4, ERLIN1, and ZEB1), (9) and (10) were 2 WFDC2 positively correlated genes (RAB6B, and PLEKHF1). B. Quantitative real-time PCR revealed that the mRNA expression levels of 10 target genes in ovarian cancer cells (S4 vs. S). *indicates p<0.05, **indicates p <0.01.

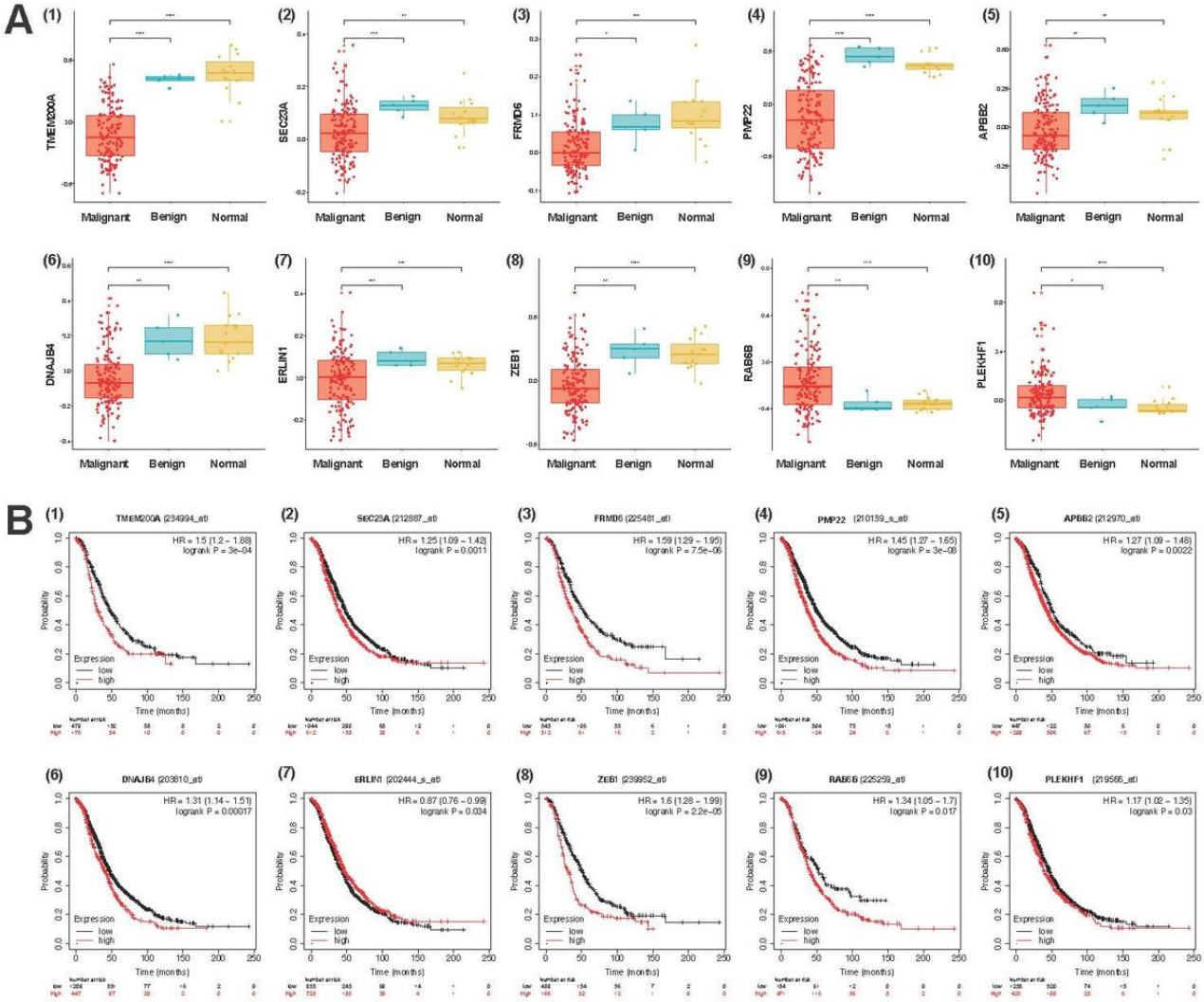

Figure 4

Clinical significance of 10 target genes A. The comparison of 10 target genes with tumor types by using GSE51088 which contains 152 epithelial ovarian cancers (including 11 epithelial ovarian borderline tumors), 5 normal ovarian tumors and 15 normal healthy ovarian tissues. (1)-(8) were 8 genes (TMEM220A, SEC23A, FRMD6, PMP22, APBB2, DNAJB4, ERLIN1, and ZEB1) that were significantly downregulated in ovarian cancers compared with ovarian benign tumors and normal tissues, and (9) and (10) were 2 genes (RAB6B, and PLEKHF1) were obviously upregulated in ovarian cancers compared with ovarian benign tumors and normal tissues. C. Online Kaplan–Meier survival statistics analysis revealed that all of these ten correlated genes were dramatically correlated with prognosis of overall survival.

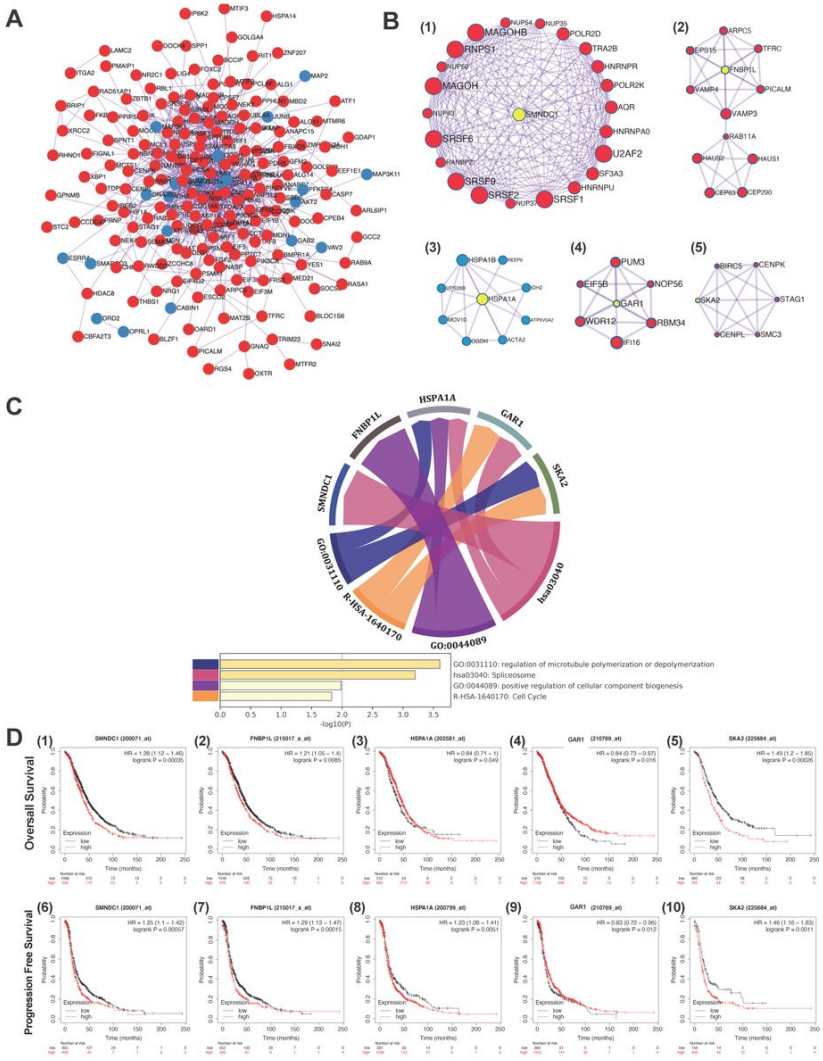

Figure 5

Protein-protein interaction (PPI) network construction, identification of MCODEs modules, and module seeds clinical significance evaluation A. PPI network of all the DEGs was constructed by Metascape. Red represented the down regulated genes, blue represented the up regulated genes. B. 5 MCODEs identification in the PPI network with MCODE score >2. Yellow represented the seed nodes of each MCODE modules, whereas the label size was determined according to the degree of connectivity. Red represented the down regulated genes, blue represented the up regulated genes. C. GO and KEGG pathway enrichment analysis of the seeds of the 5 MCODE modules generated by Metascape. D. Online Kaplan–Meier survival statistics analysis revealed that all of the five seed genes were significantly correlated with prognosis, both in overall survival ((1)-(5)) and in progression free survival ((6)-(10)).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable5.xlsx
- SupplementaryTable4.xlsx
- SupplementaryTable3.xlsx
- SupplementaryTable2.xlsx
- SupplementaryTable1.xlsx