

RESEARCH

Open Access



# A bioinformatics analysis: ZFHX4 is associated with metastasis and poor survival in ovarian cancer

Shuai Zong<sup>1</sup>, Ping-ping Xu<sup>2</sup>, Yin-hai Xu<sup>1</sup> and Yi Guo<sup>1\*</sup>

## Abstract

**Background:** Metastasis was the major cause of the high mortality in ovarian cancer. Although some mechanisms of metastasis in ovarian cancer were proposed, few have been targeted in the clinical practice. In the study, we aimed to identify novel genes contributing to metastasis and poor clinical outcome in ovarian cancer from bioinformatics databases.

**Methods:** Studies collecting matched primary tumors and metastases from ovarian cancer patients were searched in the Gene Expression Omnibus (GEO) database. Differentially expressed genes (DEGs) were screened by software *R* language. Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for the DEGs were implemented by Metascape. Venn diagram was plotted to present overlapping DEGs. The associations between the overlapping DEGs and prognosis were tested by Cox proportional hazard regression model using a cohort of ovarian cancer patients from the TCGA database. Genes affecting patients' outcomes significantly were served as hub genes. The mechanisms of the hub genes in promoting ovarian cancer metastasis were then predicted by *R* software.

**Results:** Two gene expression profiles (GSE30587 and GSE73168) met the inclusion criteria and were finally analyzed. A total of 259 genes were significantly differentially expressed in GSE30587, whereas 712 genes were in GSE73168. In GSE30587, DEGs were mainly involved in extracellular matrix (ECM) organization; For GSE73168, most of DEGs showed ion trans-membrane transport activity. There were 9 overlapping genes between the two datasets (RUNX2, FABP4, CLDN20, SVEP1, FAM169A, PGM5, ZFHX4, DCN and TAS2R50). ZFHX4 was proved to be an independent adverse prognostic factor for ovarian cancer patients (HR = 1.44, 95%CI: 1.13–1.83,  $p = 0.003$ ). Mechanistically, ZFHX4 was positively significantly correlated with epithelial-mesenchymal transition (EMT) markers ( $r = 0.54$ ,  $p = 2.59 \times 10^{-29}$ ) and ECM-related genes ( $r = 0.52$ ,  $p = 2.86 \times 10^{-27}$ ).

**Conclusions:** ZFHX4 might promote metastasis in ovarian cancer by regulating EMT and reprogramming ECM. For clinical applications, ZFHX4 was expected to be a prognostic biomarker for ovarian cancer metastasis.

**Keywords:** ZFHX4, Ovarian cancer, Metastasis, Epithelial-mesenchymal transition, Extracellular matrix

## Introduction

Ovarian cancer is the leading cause of gynecological cancer death in the developed countries and the second mortal gynecologic malignancy in China [1, 2]. Extensive metastasis due to late diagnosis and refractory relapse, which always presented as widespread intra-peritoneal dissemination, is chiefly responsible for this

\*Correspondence: 962077959@qq.com

<sup>1</sup> Department of Laboratory Medicine, Affiliated Hospital of Xuzhou Medical University, No. 99 West Huaihai Road, Xuzhou, Jiangsu 221002, People's Republic of China

Full list of author information is available at the end of the article



discouraging statistic [3, 4]. Therefore, the exploration of the mechanism of metastasis has been put under the spotlight in the investigations for ovarian cancer. Rodriguez et al. found that canonical TGF- $\beta$  signaling could significantly increase metastatic potential of ovarian cancer cells as early as 2001 [5]; Extracellular matrix (ECM) remodeling, a recent research hotspot, was perceived as contributing to ovarian cancer metastasis as well [6]. Other mechanisms associated with ovarian cancer metastasis were also widely proposed [7], but few have been targeted or fully-validated in the clinical practice. For instance, the dual nature of TGF- $\beta$  in tumor suppressing and tumor promoting limited its role in judging the prognosis of ovarian cancer patients [8]; Matrix metalloproteinase inhibitors (MMPis) that were designed to control ECM degradation did not improve patient survival due to lack of specificity and severe side effects [9].

Currently, with the rapid development of molecular expression profiling technologies, such as gene microarray, next generation sequencing and protein chips, and bioinformatics tools, several research groups have successfully predicted clinical outcomes and guided treatment decisions for cancer patients using bioinformatics data. Mannaprint<sup>®</sup> assays identified poor outcome patients early in the course of breast cancer by results from transcriptome profiling [10]. Using machine-learning techniques on the SAMIT trials, Sundar et al. developed a gene signature, which could represent as predictive biomarker for paclitaxel benefit in gastric cancer [11]. Liu et al. screened out five genes as a panel of biomarkers from sequencing data, and demonstrated that the panel could stratified EGFR-mutant NSCLC patients and provided precise guidance for them in the ADJUVANT trial [12]. These researches provided directions for us to search for novel genes and pathways involved in ovarian cancer metastasis.

Therefore, in the study, we aimed to find out new hub genes which were fit for drug targets and biomarkers in ovarian cancer metastasis, and probe into underlying mechanisms by mining bioinformatics databases. For the purposes, we searched the studies including samples of primary tumors and metastatic lesion in the Gene Expression Omnibus (GEO) database, and differentially expressed genes (DEGs) between groups were, then, screened by programming software *R* language. This was followed by Gene Ontology (GO) functional analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment of DEGs. Next, the hub genes were identified by using Venn diagram and Cox regression analysis, and the functions of the hub genes in ovarian cancer metastasis were explored using samples from TCGA database with *R* software.

## Materials

### Strategies of search

We systematically searched for the relevant studies in the GEO database (<https://www.ncbi.nlm.nih.gov/gds/>). The search terms were shown as follow: [("Ovarian Neoplasms") OR ("Ovarian Neoplasm") OR ("Ovary Neoplasms") OR ("Ovary Neoplasm") OR ("Ovary Cancer") OR ("Ovary Cancers") OR ("Ovarian Cancer") OR ("Ovarian Cancers") OR ("Cancer of Ovary") OR ("Cancer of the Ovary")] AND [("Metastasis") OR ("Metastases") OR ("Metastatic")]. The inclusion criteria for eligible studies were that (1) the enrolled population was treatment-naive patients with ovarian cancer; (2) matched primary tumors and metastatic lesions from ovarian cancer patients were collected and analyzed; (3) platforms used for sequencing or gene chip were introduced. The exclusion criteria were that (1) ovarian cancer patients in the study were secondary to / combined with other malignancies; (2) gene probe ID in certain platforms could not be transformed into gene names; (3) the data in the downloaded files were incomplete or invalid.

### Selection of DEGs

The raw data were downloaded as Series Matrix Files. The data values in the microarray were log<sub>2</sub>-transformed, and normalized using function *normalize.quantiles* in *R* programming tools (version 4.2.0). Probes were converted to gene symbols according to the annotation information of the platform. Then, "limma" package in the *R* language (version 4.2.0) was used to screen DEGs. The thresholds for DEGs identification were set as *p* value < 0.05 and absolute fold change  $\geq 1.5$ . Volcano plots and heat maps were generated using TBtools (version 1.049), which is an integrative toolkit developed for interactive analyses of big biological data [13].

### Analysis of functions and pathways

To predict potential biological functions of DEGs and pathways they might be involved in, we using Metascape (<https://metascape.org/>) to perform annotation analysis. Metascape was a web-based tool, and has been designed for experimentalists to apply powerful computational analysis pipelines to analyze and interpret large-scale datasets [14]. Using Metascape, we implemented the GO enrichment for the lists of DEGs. Deregulated pathways were also ranked on the basis of  $-\log_{10}$  (*p* value) from KEGG analysis by Metascape.

### Identification of hub genes

Firstly, Venn diagram was plotted online to present common DEGs (<http://jvenn.toulouse.inra.fr/app/example.html>), and the overlapping DEGs were regarded as candidate genes for further analysis. Next, the associations

between candidate genes and prognosis were tested by Cox proportional hazard regression model on the cohort of ovarian cancer patients from the TCGA database. Univariate and multivariate cox regression analysis were performed using “forestplot” package from *R* software (version 4.2.0). A nomogram was developed to predict the 1, 3, and 5-year overall survival using “rms” package in *R* programming language (version 4.2.0). The nomogram was validated by calibration curves using the same *R* package. We served genes affecting patients’ outcome significantly ( $p < 0.05$ ) as hub genes. Finally, Kaplan–Meier analysis was used to verify the effect of hub genes on survival using Kaplan–Meier plotter (<http://kmplot.com/analysis/>).

### Exploration of mechanisms of hub genes in ovarian cancer metastasis

The functions of the hub genes and the signaling pathways they might be involved in were explored using samples from TCGA database. “GSVA” package in *R* software (version 4.2.0) was used to analyze the correlation between hub genes and pathway scores by Spearman correlation [15], and  $p$  values less than 0.05 were considered statistically significant. The associations of hub genes and other gene sets were also performed using the same *R* package.

## Results

### Data source

One hundred and ten studies were retrieved from the GEO database. After screening titles and abstracts, 101 studies were excluded. Among them, 82 studies did not select matched primary tumor and metastasis, 14 studies included patients combined with other malignant tumors, 3 studies uploaded the incomplete data, and 2 studies contained patients received neo-adjuvant chemotherapy before operations. In the remaining 9 studies, 5 did not provide enough platform annotation information for us to convert EntrezID into gene names, so 4 records, which contains GSE30587, GSE73064, GSE73091 and GSE73168, respectively, met the inclusion criteria. Because GSE73064 and GSE73091 were subgroups of GSE73168, 2 gene expression profiles (GSE30587 and GSE73168) were finally analyzed.

GSE30587 consisted of 9 pairs of matched primary and metastatic tumors from ovarian cancer patients, and metastatic lesions were located in the omenta. GSE73168 contained 8 pairs of matched lesions, and metastases were isolated from peritonea. The chip-based platform GPL6244 (HuGene-1\_0-st; Affymetrix Human Gene 1.0 ST Array) was applied for the GSE30587, and the data from GSE73168 was generated from GPL570 platform

(HG-U133\_Plus\_2; Affymetrix Human Genome U133 plus 2.0 Array).

### Gene Expression Changes in GSE30587 and GSE73168

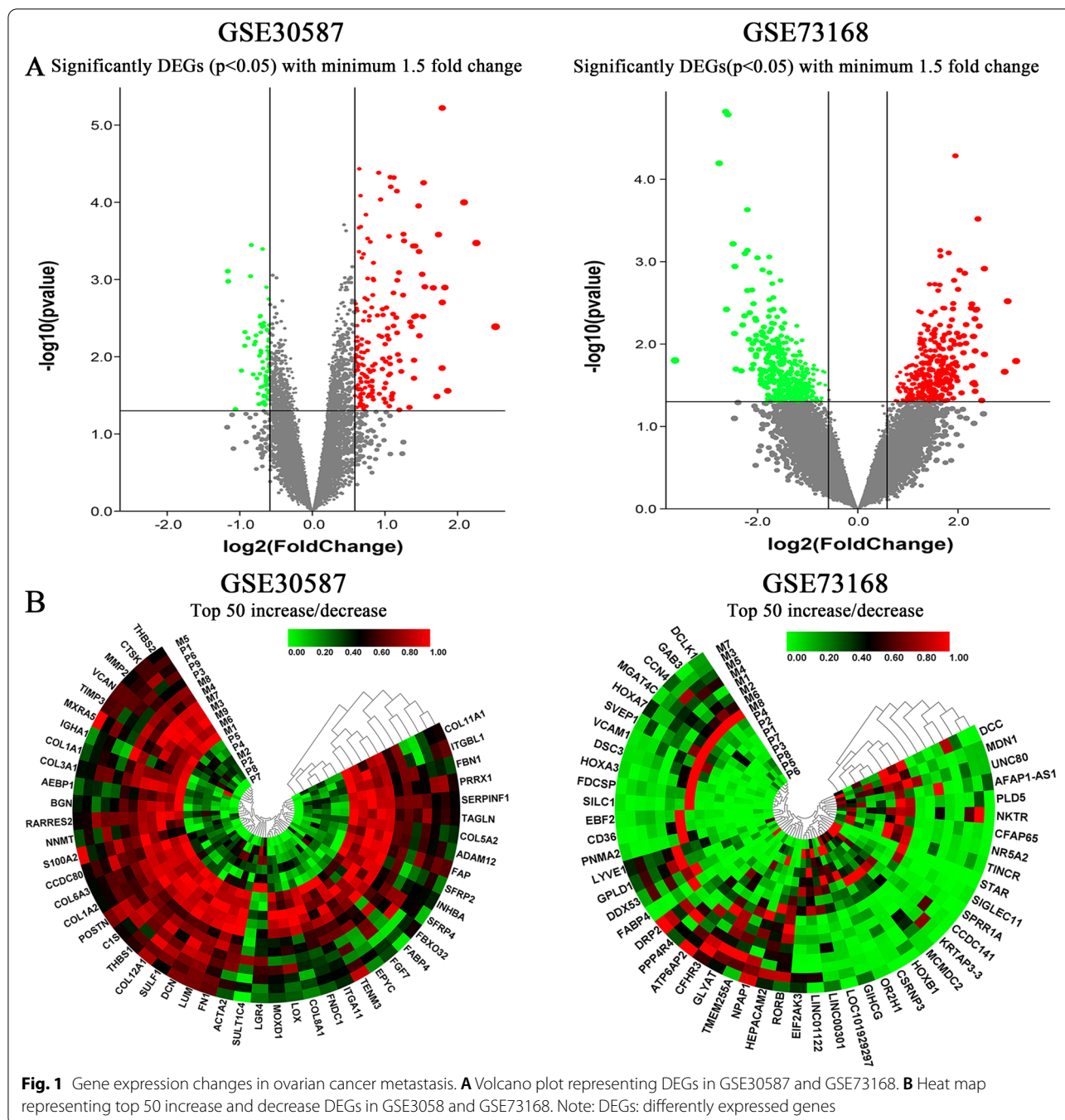
A total of 259 genes were significantly differentially expressed in GSE30587 with a significant level of  $p$  value  $< 0.05$  and absolute fold change  $\geq 1.5$  (Fig. 1A Left), whereas 712 genes were differentially expressed in GSE73168 (Fig. 1A Right). The green dots represented down-regulated genes, while the up-regulated genes were red-color coded. The top 50 up-regulated and down-regulated genes were then determined and showed as heat maps (Fig. 1B).

GO analysis was done to annotate the biological functions DEGs occurring during metastasis. In GSE30587, DEGs were mainly located in the extracellular matrix (ECM) of cellular components (CC), exhibited the molecular functions (MF) of ECM structural constituent, and were involved in the biological processes (BP), such as extracellular matrix organization, collagen binding, and integrin binding (Fig. 2A Left). For GSE73168, DEGs were primarily enriched in plasma membranes, showed functions of anion trans-membrane transporter activity, and participated in the processes of inorganic ion trans-membrane transport (Fig. 2A Right). Remarkably, terms related with development, for instance, vasculature development, heart development, cell fate commitment, skeletal system development, embryonic organ development, regulation of wound healing and tissue remodeling, were common in both of the two data sets (Fig. 2A).

KEGG analysis was done to find out deregulated pathways involved in metastasis. Complement and coagulation cascades and ECM-receptor interaction were the most significant pathways in the GSE30587 (Fig. 2B Left), while the most significantly correlated pathways in GSE73168 were Neuroactive ligand-receptor interaction (Fig. 2B Right). GO enrichment and KEGG pathway analysis offered guides for exploring the functions of key genes screened out for subsequent experiments.

### Overlapping Genes in GSE30587 and GSE73168

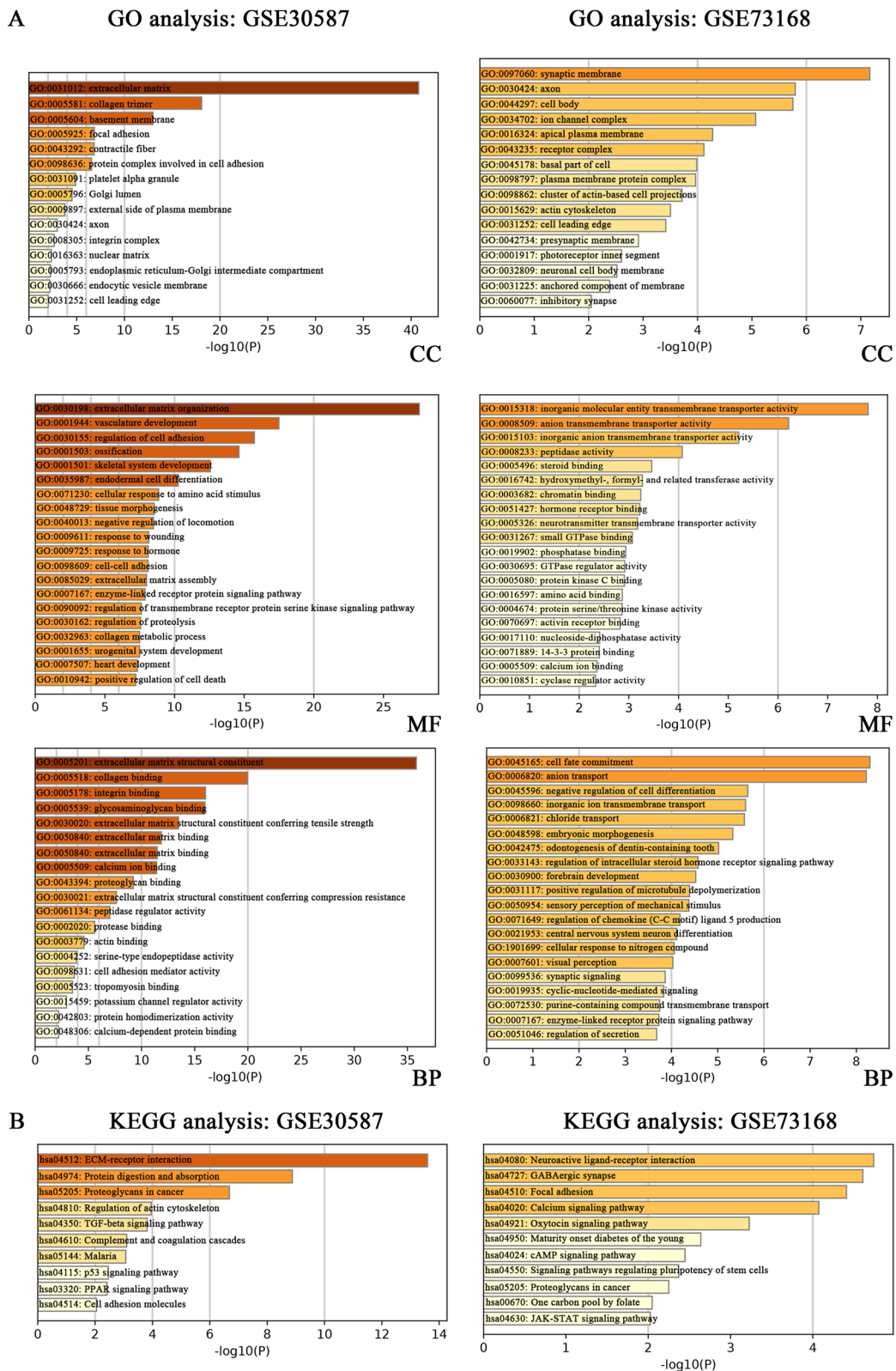
It is expected that the expressions of various genes altered as the tumors progressed from primary sites to metastasis loci. As mentioned above, GSE30587 showed that the process of metastasis resulted in deregulation of 259 mRNAs, while the dissemination formation was accompanied by changes of 712 mRNAs in GSE73168. In the next step, we characterized the overlap in the two datasets. Of the 259 and 712 DEGs in GSE30587 and GSE73168, respectively, 9 overlapping genes were identified (Fig. 3A). Four genes were commonly up-regulated, two genes were down-regulated in both of the datasets, whereas other three genes showed different directions of



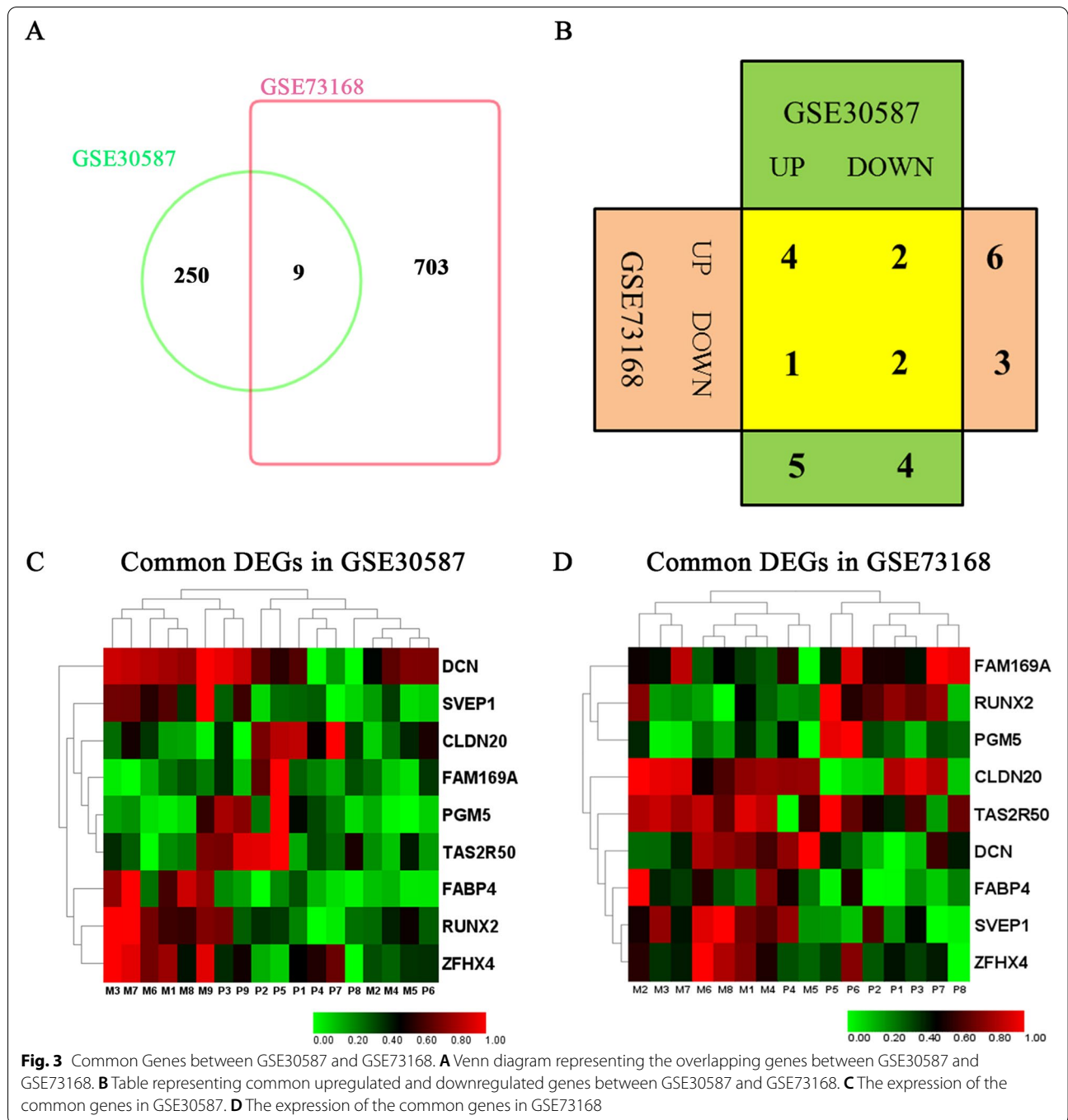
change in the two groups two (Fig. 3B). The co-upregulated DEGs included FABP4, SVEP1, ZFH4 and DCN, FAM169A and PGM5 were co-downregulated, CLDN20 and TAS2R50 were down-regulated in GSE30587 but up-regulated GSE73168, and RUNX2 was up-regulated in GSE30587 but down-regulated in GSE73168 (Fig. 3C and D). The nine overlapping genes were considered as candidate genes for further analysis.

### Univariate and multivariate analysis of ovarian cancer patients in the TCGA database

To validate the effects of candidate genes on the outcome of ovarian cancer, we conducted univariate and multivariate Cox proportional hazards analysis on 374 patients from the TCGA database. The characteristics of the patients in the cohort were included in Table 1. The median follow-up was 86.0 months. At the end of follow-up, a total of 230



**Fig. 2** Function analysis of deregulated genes. **A** GO enrichment analysis of DEGs in GSE30587 and GSE73168. **B** KEGG analysis of DEGs in GSE30587 and GSE73168. Note: GO: Gene Ontology; DEGs: differentially expressed genes; CC: cellular component; MF: molecular function; BP: biological process; KEGG: Kyoto Encyclopedia of Genes and Genomes



(61.5%) deaths occurred. The median age at diagnosis was 59 years (range: 30–87). Patients with stage III/IV ovarian cancer accounted for 94.1%.

Univariable analysis indicated that ZFH4 was the prognostic predictors among the candidate genes (Fig. 4A). In multivariable analysis, ZFH4 was still proved to be an independent adverse prognostic factor for ovarian cancer patients (HR=1.44, 95%CI: 1.13–1.83,  $p=0.003$ ) (Fig. 4B).

A nomogram was plotted to predict 1-, 3-, and 5-year OS of ovarian cancer patients. As shown in the Fig. 4C, different values of ZFH4 corresponded to various points. Total points could be calculated according to the patient's ZFH4, age and racial. Therefore, we could predict ovarian cancer patients' OS by corresponding total points to survival probability. The calibration curve certified that nomogram-predicted curves were close to the observed curves,

**Table 1** Clinical information of ovarian cancer patients in the TCGA database

	Characteristics	N (%)
<b>Status</b>	Alive	144 (38.50)
	Dead	230 (61.50)
<b>Age</b>	Median (Min, Max)	59 (30, 87)
<b>Gender</b>	Female	374 (100.00)
<b>Race</b>	American Indian	8 (2.14)
	Asian	11 (2.94)
	Black	25 (6.68)
	Islander	6 (1.61)
	White	324 (86.63)
<b>TNM stage</b>	IC	2 (0.53)
	IIA	3 (0.81)
	IIB	5 (1.34)
	IIC	15 (4.01)
	IIIA	7 (1.87)
	IIIB	14 (3.74)
	IIIC	271 (72.46)
	IV	57 (15.24)
<b>Grade</b>	G1	6 (1.60)
	G2	42 (11.24)
	G3	320 (85.56)
	G4	6 (1.60)

Age was presented using median and range

suggesting satisfactory agreement between the predicted results and the actual values (Fig. 4D).

We also checked the effect of ZFH4 on ovarian cancer patient's prognosis using Kaplan–Meier Plotter. Patients were split by median expression, and increased ZFH4 caused significantly decreased OS (HR=1.28, 95%CI: 1.11–1.47,  $p=0.00066$ ) and progress-free survival (PFS) (HR=1.86, 95%CI: 1.53–2.25,  $p=1.4 \times 10^{-10}$ ) (Fig. 4E and F). ZFH4, thus, was selected as the hub gene for the function exploration.

#### Mechanisms of ZFH4 in ovarian cancer metastasis

To better understand the potential mechanisms of the hub gene in promoting metastasis in ovarian cancer, we explored the common pathways in which ZFH4 might be involved in using ovarian cancer samples from TCGA database by R software GSVA package via ssGSEA algorithm. The most significantly positively correlated functions

included collagen formation ( $r=0.67$ ,  $p=4.72 \times 10^{-51}$ ) and degradation of ECM ( $r=0.63$ ,  $p=2.84 \times 10^{-43}$ ) (Fig. 5A and B). ECM-related genes ( $r=0.52$ ,  $p=2.86 \times 10^{-27}$ ) and EMT markers ( $r=0.54$ ,  $p=2.59 \times 10^{-29}$ ) were the most significantly positively correlated gene sets (Fig. 5C and D). There was a significant positive correlation between TGF- $\beta$  and the ZFH4 ( $r=0.60$ ,  $p=2.31 \times 10^{-37}$ ) (Fig. 5E).

As mentioned above, GO and KEGG analysis indicated that EMT and matrisome alteration were key features during ovarian cancer progression, which were in line with the predicted functions of ZFH4. As enrichment analysis also implied that ion channels might play an important role in ovarian cancer metastasis, we further analyzed the relationships between ZFH4 and ion trans-membrane transporters using gene sets from GO datasets (GO: 0,015,318). We found that TRPM3, BEST1, TRPV4, SLC39A8, SLC10A7, ABCC8, CACNB4, SLC6A15, TSPO2, KCNK1 and PON3 were significantly positively correlated with ZFH4 ( $p<0.05$ ), while COX4I1 was significantly negatively correlated with ZFH4 ( $p<0.05$ ) (Fig. 6). The results suggested that ZFH4 might be regulated by the activation of ion trans-membrane channels.

Notably, we also found that the expression of ZFH4 was significantly negatively correlated with DNA replication ( $r=-0.23$ ,  $p=8.5 \times 10^{-06}$ ), MYC targets ( $r=-0.27$ ,  $p=1.72 \times 10^{-07}$ ), DNA repair ( $r=-0.29$ ,  $p=6.44 \times 10^{-09}$ ) and tumor proliferation signature ( $r=-0.11$ ,  $p=0.032$ ) (Fig. 7A, B, C and D), but significantly positively correlated with apoptosis ( $r=0.41$ ,  $p=1.65 \times 10^{-16}$ ) (Fig. 7E). These data suggested that ZFH4 repressed the proliferative phenotype of ovarian cancer cells while it boosted their metastatic potentials.

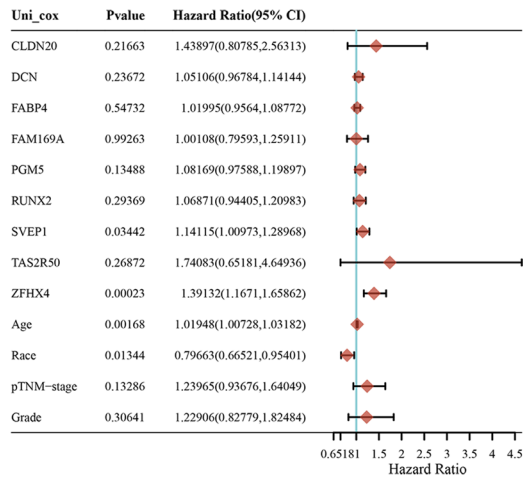
#### Discussion

In the present study, we did a bioinformatics analysis for primary tumors and metastatic lesions of ovarian cancer samples from GEO database to explore new genes and their mechanisms in ovarian cancer metastasis. While some information was available on the factors and pathways involved in the promoting invasiveness [16, 17], the progression of ovarian cancer was an intricate, multistage process along with numerous molecular changes, which further needed to be explored. Therefore, we systematically searched relevant studies in the GEO data library, and comprehensively analyzed the available datasets to

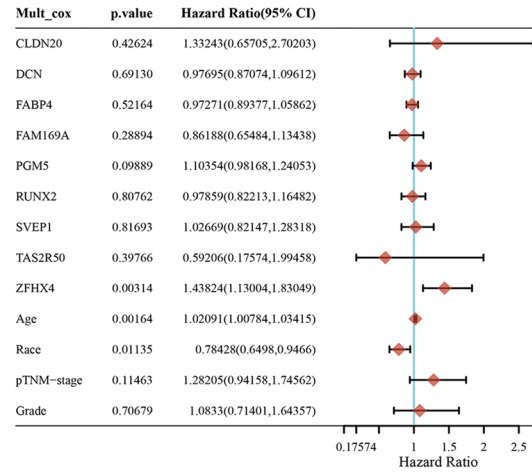
(See figure on next page.)

**Fig. 4** Hub genes selection using univariate and multivariate analysis. **A** Univariate analysis of prognostic factors of OS in ovarian cancer patients. **B** Multivariate analysis of prognostic factors of OS in ovarian cancer patients. **C** The nomogram for predicting probabilities of patients of 1, 3, and 5-year OS. **D** Evaluation of ZFH4 prediction model by calibration plots. **E** Kaplan–Meier plot for OS analysis for ZFH4. **F** Kaplan–Meier plot for PFS analysis for ZFH4. Notes: OS: overall survival; PFS: progress-free survival; Islander: Pacific Islander

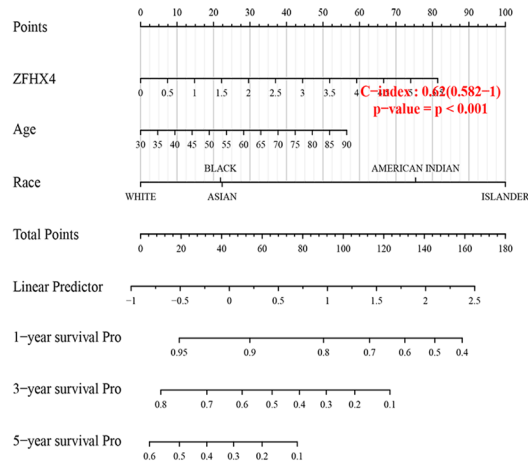
**A**



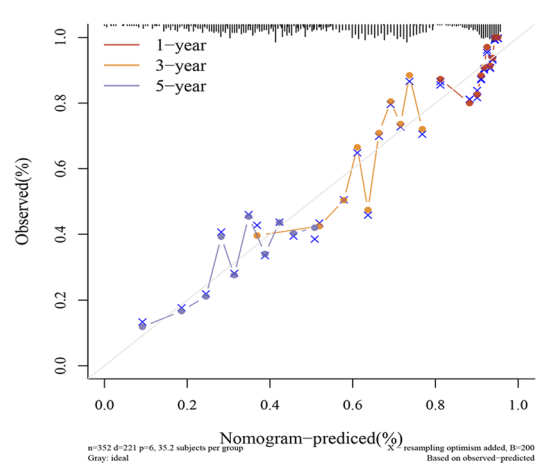
**B**



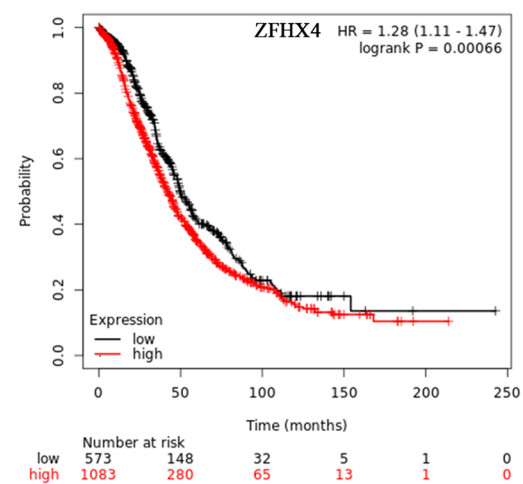
**C**



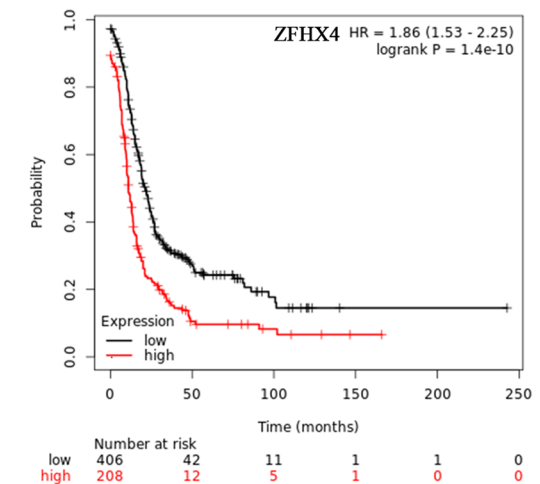
**D**



**E**



**F**



**Fig. 4** (See legend on previous page.)



find out common and unique signatures during the process of metastasis in ovarian cancer. Our results showed that the alterations of matrisome components and changes of gated channel activity were involved in this process. Moreover, we firstly identified ZFHX4 as the hub gene contributing to the ovarian cancer metastasis, and EMT regulation and ECM remodeling were considered as its potential mechanisms.

It has been well demonstrated that gene expression profiling of ECM components underwent deregulation during early colonization [18]. In turn, ECM remodeling could be conducive to cancer metastasis [19]. In the primary sites, basement membranes, which lined the basal surface of tumors and presented as dense collagen structures, were broken through proteolytic ECM degradation and non-proteolytic force by neoplastic cells and cancer-associated fibroblasts (CAFs) when tumor rapid growth led to diffusion-limited oxygen and nutrient supply [20–22]. Next, collagens in interstitial matrix were linearized by tumor-derived lysyl oxidase (LOX), which created migratory tracks used for the migration of the tumor cells [23, 24]. In the metastatic sites, however, CAFs deposited more collagens into matrix to support new lesions growth and help them evade immune surveillance [25, 26]. In accordance with the previous studies, we found that genes encoding alpha chains for fibrillar collagens, such as COL3A1, COL5A2, COL1A1 and COL1A2, were the top up-regulated DEGs in the metastatic samples compared to the primary ones. Besides collagens, altered glycoproteins, another kind of component in the ECM, might also contribute to the tumor metastasis. For instance, the expression of POSTN, which encoded a secreted extracellular matrix protein named Periostin, showed significant increase in the GSE30587. Yue et al reported that stromal Periostin could fuel the migration and invasion of ovarian cancer cells by binding to integrin  $\alpha\beta3$  and subsequently activating the PI3K/Akt pathway and inducing the EMT [27]. The regulation of TGFBI was disordered as well according to our results, and its encoding protein (transforming growth factor  $\beta$  induced protein, TGFBIp) was located in the extracellular matrix to bind to type I, II and IV collagens. As mentioned in the recent papers, macrophage-derived TGFBIp promoted ovarian cancer cell migration via inducing an immunosuppressive microenvironment [28, 29].

Ion channel disorder appeared to be responsible for the enhanced migration of cancer cells and ECM dysregulation [30, 31]. Aberrant activation of  $Ca^{2+}$ -permeable channels on the endoplasmic reticulum membrane and plasma membrane led to impulse  $Ca^{2+}$  influx [32]. The oscillatory  $Ca^{2+}$  signaling facilitated invadopodial precursor assembly via Src activation and increased proteolytic activity of invadopodia by recruitment of ECM-degrading membrane type 1 matrix metalloproteinase (MT1-MMP) [33]. Then,  $Ca^{2+}$ -activated  $K^+$  channels, such as SK3, enhanced the driving force for  $Ca^{2+}$  entry through hyperpolarization of resting membrane potential [34, 35]. Moreover, abnormal opens of sodium ion channels also contributed to invadopodia formation via F-actin polymerization [36].  $Na^+/H^+$  exchangers on the membrane caveola resulted in a per-imembrane acidification, which rendered secreted cathepsins and pH-dependent ECM-digesting proteases more active [37]. In our study, inorganic ion transmembrane transports for calcium (TRPM3), sodium (SLC family) and potassium (KCNK1) were significantly up-regulated in the metastatic ovarian cancer samples compared to the primary ones, which were in accord with previous researches.

ZFHX4 is one of the five members of zinc finger homobox family with molecular weights of 397 kDa [38]. The protein comprises 22 typical  $C_2H_2$  structures of zinc finger and 4 homeodomains, and its coding gene extends over 180 kb and is located on 8q13.3-q21.11 [38]. ZFHX4 is predicted to be active in nucleus, and enable DNA-binding transcription factor activity. Fèvre-Montange et al firstly reported that ZFHX4 was highly expressed in paraneoplastic tumors, a rare but aggressive type of pineal region tumor, using oligonucleotide arrays in 2006 [39]. Similarly, Wang et al found that ZFHX4 amplifications were more common in the metastatic colorectal cancer patients than those without metastases [40]. In ovarian cancer, elevated expression of ZFHX4 was also demonstrated to be correlated with poor outcomes, which was mainly influenced by metastasis, in bioinformatics databases and real-world cohort samples [41, 42]. Our results were in accord with these studies, which indicated that ZFHX4 reduced the survival times of ovarian cancer patients through promoting metastasis.

Currently, researches for the functions of ZFHX4 largely focused on the therapy resistance and

(See figure on next page.)

**Fig. 5** Functions and mechanisms of ZFHX4 in ovarian cancer metastasis. **A** The correlation between ZFHX4 and collagen-related genes. **B** The correlation between ZFHX4 and genes related to ECM degradation. **C** The correlation between ZFHX4 and ECM-related genes. **D** The correlation between ZFHX4 and EMT-related genes. **E** The correlation between ZFHX4 and TGF- $\beta$ . Notes: ECM: extracellular matrix; EMT: epithelial-mesenchymal transition

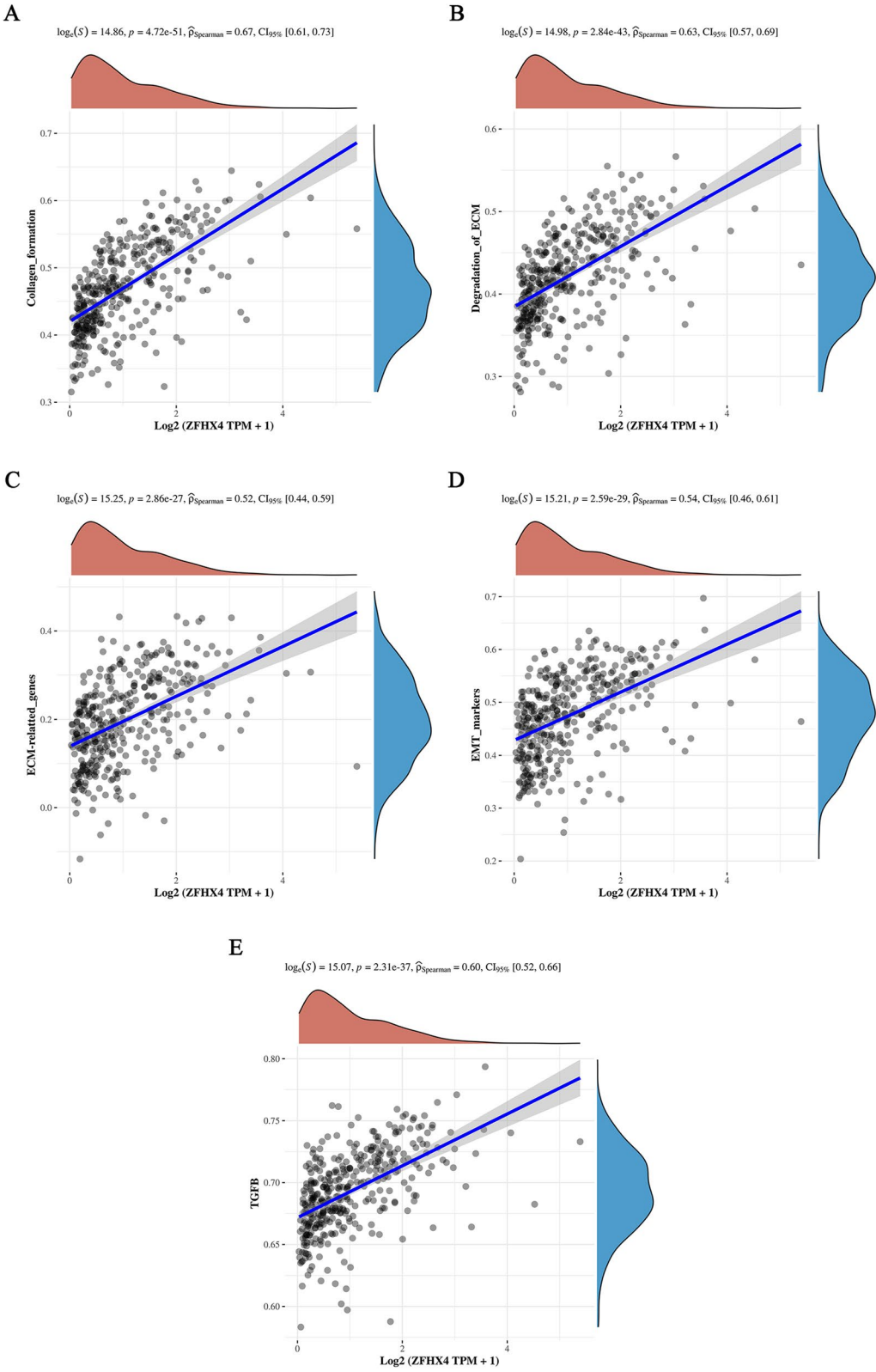
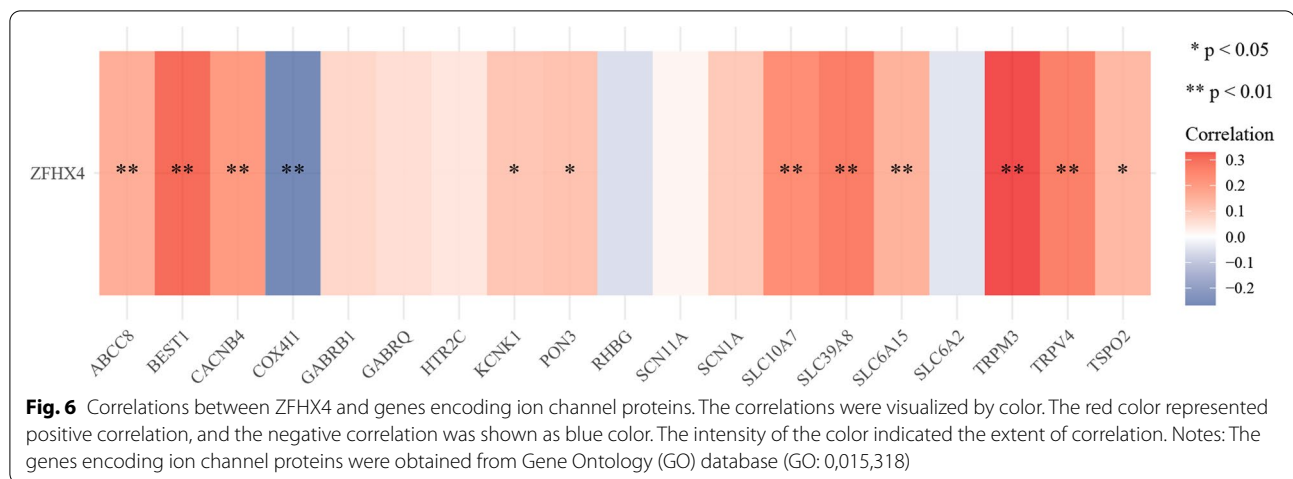


Fig. 5 (See legend on previous page.)



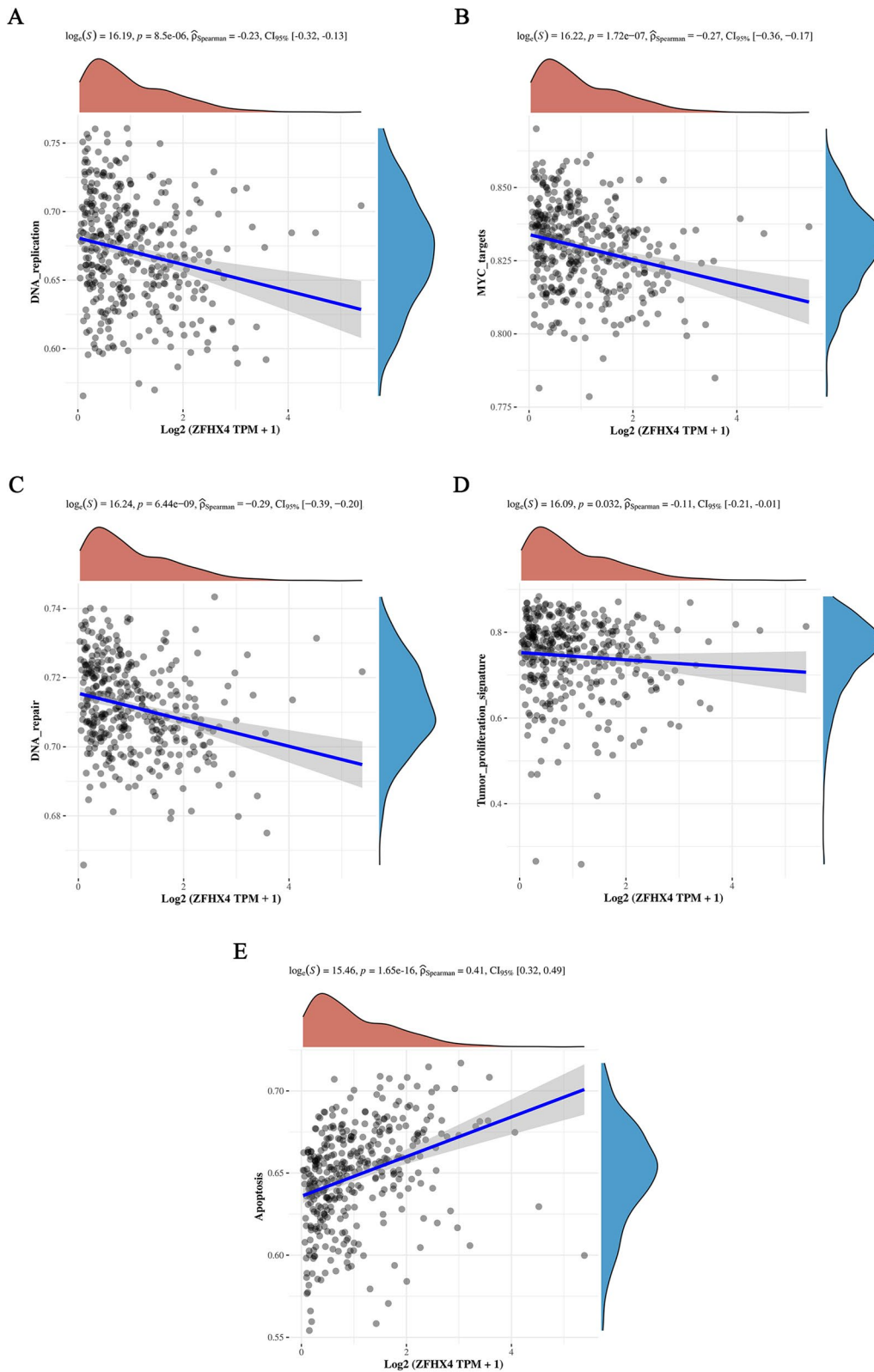
consequent recurrence. Chudnovsky et al. suggested that ZFHx4 interacted with the CHD4, an essential member of the NuRD (nucleosome remodeling and deacetylase) complex, to regulate the therapy-resistant tumor initiating cells in glioblastomas [43]. However, the detailed mechanism by which ZFHx4 regulated metastasis of ovarian cancer remained unknown. According to our bioinformatics analysis, we, for the first time, linked ZFHx4 with EMT-associated genes. In 2020, “the EMT International Association (TEMTIA)” issued a consensus statement on the famous academic journal “NATURE REVIEWS MOLECULAR CELL BIOLOGY” to provide guidelines and definitions for EMT [44]. In the statement, experts emphasized central roles of transcription factors belonging to the ZEB families (Zeb1 and Zeb2) in the execution of EMT in cancer metastasis [44]. Coincidentally, Zeb1 and Zeb2 were also two members of zinc finger homobox family, and they shared similar molecular structures with ZFHx4 [45], which rendered the function of ZFHx4 as an EMT inducer reasonable. Besides, ZFHx4 could be modulated by canonical pathways for EMT regulation, such as TGF- $\beta$  signaling pathway [46, 47], which further supported the contributions of ZFHx4 to EMT. Simultaneously, we found the positive correlation between ZFHx4 and ECM-related genes. ZFHx4, on the one hand, was conducive to the collagen formation, but on the other hand, induced degradation of ECM. The results might be explained by the fact that individual cancer cell migrated faster in a mesenchymal phenotypes, while collective cell migration depended on the cell–matrix junctions [48]. Next, significant correlations were found between ZFHx4 and lots of ion channels. Eskandari et al. reported that Kv11.1 channel reprogrammed EMT in colon cancer via regulating zinc

finger homobox proteins [49], suggesting that ZFHx4 might also act as a mediator converting the signals from ion transporter activation to EMT occurrence and ECM reprogramming in ovarian cancer. Although ZFHx4 took an active part in promoting cancer metastasis, it functioned as a tumor suppressor in tumor proliferation. This rather interesting finding might be explained by the opinion that expression changes of certain genes necessary for the initial tumorigenesis were not essential for subsequent metastasis, which was proposed by Mitra et al. [18]. In their study, they observed that most of shared deregulated genes metastasis versus primary tumors were altered in the different directions [18].

There were surely some limitations in the study. Firstly, DEGs were screened from primary tumors and metastatic lesions in GEO database, while the candidate genes were generated from ovarian cancer patients with various stages in TCGA database. Secondly, laboratory-based experiments have not been conducted to validate the mechanisms of ZFHx4 in metastasis in ovarian cancer. Thirdly, the interpretation of these data was without consideration of which cellular compartment expresses ZFHx4, since whole tumor samples were sequenced in GSE73168. In the future studies, clinical samples from ovarian cancer patients would be collected to detect the expressions of ZFHx4. Then, the relationships between ZFHx4 and EMT- and ECM-related proteins would be further confirmed by *in vitro* or (and) *in vivo* experiments.

## Conclusion

In summary, based on the bioinformatics analysis, we hypothesized that ZFHx4 might promote metastasis in ovarian cancer by regulating EMT and reprogramming



**Fig. 7** Functions of ZFHX4 in ovarian cancer proliferation. **A** The correlation between ZFHX4 and genes related to DNA replication. **B** The correlation between ZFHX4 and genes targeting MYC. **C** The correlation between ZFHX4 and genes related to DNA repair **D** The correlation between ZFHX4 and genes related to tumor proliferation. **E** The correlation between ZFHX4 and apoptosis-related genes

matrisome. For clinical applications, ZFH4 was expected to be a prognostic biomarker for ovarian cancer metastasis. On the one hand, ZFH4 could be used to predict occurrence of metastasis in ovarian cancer patients, especially for those have completed therapy and were undergoing follow-up; On the other hand, the expression of ZFH4 is prognostic for survival outcomes of ovarian cancer patients with metastasis.

#### Abbreviations

GEO: Gene Expression Omnibus; DEGs: Differentially expressed genes; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; ECM: Extracellular matrix; EMT: Epithelial-mesenchymal transition; CC: Cellular components; MF: Molecular functions; BP: Biological processes; CAFs: Cancer-associated fibroblasts; LOX: Lysyl oxidase; MT1-MMP: Matrix 1-metalloproteinase.

#### Acknowledgements

We would like to extend our thanks to the generous researchers who have uploaded their experimental data onto online databases.

#### Authors' contributions

Shuai Zong and Ping-ping Xu performed the data mining and drafted the manuscript, and they were contributed equally to the whole study. Yin-hai Xu participated in the data analysis and interpretation. Yi Guo was involved in the study design and critical revision of the manuscript. All authors have read and approved the final manuscript. As the corresponding author, Yi Guo had the final responsibility for the decision to submit the article for publication.

#### Funding

This work was supported by the Foundation Research Project (Natural Science Foundation) of Jiangsu Province in China (Grant No. BK20201013), Key Research and Development Program (Social Development) of Xuzhou City in China (Grant No. KC20158) and "Shuang-Chuang Doctor" Project of Jiangsu Province in China (Grant No. 2020204003).

#### Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### Declarations

##### Ethics approval and consent to participate

Not applicable.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

##### Author details

<sup>1</sup>Department of Laboratory Medicine, Affiliated Hospital of Xuzhou Medical University, No. 99 West Huaihai Road, Xuzhou, Jiangsu 221002, People's Republic of China. <sup>2</sup>Department of Laboratory Medicine, Xuzhou Central Hospital, Jiangsu 221006, China.

Received: 29 May 2022 Accepted: 23 July 2022

Published online: 01 August 2022

#### References

- Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA: Cancer J Clin.* 2021;71(1):7–33.
- Wang Z, Guo E, Yang B, Xiao R, Lu F, You L, Chen G. Trends and age-period-cohort effects on mortality of the three major gynecologic cancers in China from 1990 to 2019: Cervical, ovarian and uterine cancer. *Gynecol Oncol.* 2021;163(2):358–63.
- Roberts CM, Cardenas C, Tedja R. The role of intra-tumoral heterogeneity and its clinical relevance in epithelial ovarian cancer recurrence and metastasis. *Cancers (Basel).* 2019;11(8):1083.
- Bregenzler ME, Horst EN, Mehta P, Novak CM, Repetto T, Mehta G. The role of cancer stem cells and mechanical forces in ovarian cancer metastasis. *Cancers (Basel).* 2019;11(7):1008.
- Rodriguez GC, Haisley C, Hurteau J, Moser TL, Whitaker R, Bast RC Jr, Stack MS. Regulation of invasion of epithelial ovarian cancer by transforming growth factor-beta. *Gynecol Oncol.* 2001;80(2):245–53.
- Dhaliwal D, Shepherd TG. Molecular and cellular mechanisms controlling integrin-mediated cell adhesion and tumor progression in ovarian cancer metastasis: a review. *Clin Exp Metastasis.* 2022;39(2):291–301.
- Yousefi M, Dehghani S, Nosrati R, Ghanei M, Salmaninejad A, Rajaie S, Hasanzadeh M, Pasdar A. Current insights into the metastasis of epithelial ovarian cancer - hopes and hurdles. *Cell Oncol (Dordr).* 2020;43(4):515–38.
- Bachman KE, Park BH. Duel nature of TGF-beta signaling: tumor suppressor vs. tumor promoter. *Curr Opin Oncol.* 2005;17(1):49–54.
- Winer A, Adams S, Mignatti P. Matrix metalloproteinase inhibitors in cancer therapy: turning past failures into future successes. *Mol Cancer Ther.* 2018;17(6):1147–55.
- van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *New England J Med.* 2002;347(25):1999–2009.
- Sundar R, Barr Kumarakulasinghe N, Huak Chan Y, Yoshida K, Yoshikawa T, Miyagi Y, Rino Y, Masuda M, Guan J, Sakamoto J, et al. Machine-learning model derived gene signature predictive of paclitaxel survival benefit in gastric cancer: results from the randomised phase III SAMIT trial. *Gut.* 2022;71(4):676–85.
- Liu SY, Bao H, Wang Q, Mao WM, Chen Y, Tong X, Xu ST, Wu L, Wei YC, Liu YY, et al. Genomic signatures define three subtypes of EGFR-mutant stage II-III non-small-cell lung cancer with distinct adjuvant therapy outcomes. *Nat Commun.* 2021;12(1):6450.
- Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13(8):1194–202.
- 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.
- Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14:7.
- Sestito R, Tocci P, Roman C, Di Castro V, Bagnato A. Functional interaction between endothelin-1 and ZEB1/YAP signaling regulates cellular plasticity and metastasis in high-grade serous ovarian cancer. *J Exp Clin Cancer Res.* 2022;41(1):157.
- Shi X, Yu X, Wang J, Bian S, Li Q, Fu F, Zou X, Zhang L, Bast RC, Jr, Lu Z et al. SIK2 promotes ovarian cancer cell motility and metastasis by phosphorylating MYLK. *Mol Oncol.* 2022;16(13):2558–74.
- Mitra S, Tiwari K, Podicheti R, Pandhiri T, Rusch DB, Bonetto A, Zhang C, Mitra AK. Transcriptome profiling reveals matrisome alteration as a key feature of ovarian cancer progression. *Cancers (Basel).* 2019;11(10):1513.
- Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat Commun.* 2020;11(1):5120.
- Popova NV, Jucker M. The functional role of extracellular matrix proteins in cancer. *Cancers (Basel).* 2022;14(1):238.
- Grossen A, Smith K, Coulibaly N, Arbuckle B, Evans A, Wilhelm S, Jones K, Dunn I, Townner R, Wu D, et al. Physical forces in glioblastoma migration: a systematic review. *Int J Mol Sci.* 2022;23(7):4055.
- Biffi G, Tuveson DA. Diversity and biology of cancer-associated fibroblasts. *Physiol Rev.* 2021;101(1):147–76.
- Glentis A, Oertle P, Mariani P, Chikina A, El Marjou F, Attieh Y, Zaccarini F, Lae M, Loew D, Dingli F, et al. Cancer-associated fibroblasts induce metalloprotease-independent cancer cell invasion of the basement membrane. *Nat Commun.* 2017;8(1):924.
- Setargew YFI, Wyllie K, Grant RD, Chitty JL, Cox TR. Targeting lysyl oxidase family mediated matrix cross-linking as an anti-stromal therapy in solid tumours. *Cancers (Basel).* 2021;13(3):491.

25. Fang M, Yuan J, Peng C, Li Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol.* 2014;35(4):2871–82.
26. Romer AMA, Thorseth ML, Madsen DH. Immune modulatory properties of collagen in cancer. *Front Immunol.* 2021;12:791453.
27. Yue H, Li W, Chen R, Wang J, Lu X, Li J. Stromal POSTN induced by TGF-beta1 facilitates the migration and invasion of ovarian cancer. *Gynecol Oncol.* 2021;160(2):530–8.
28. Steitz AM, Steffes A, Finkernagel F, Unger A, Sommerfeld L, Jansen JM, Wagner U, Graumann J, Muller R, Reinartz S. Tumor-associated macrophages promote ovarian cancer cell migration by secreting transforming growth factor beta induced (TGFB1) and tenascin C. *Cell Death Dis.* 2020;11(4):249.
29. Lecker LSM, Berlato C, Maniati E, Delaine-Smith R, Pearce OMT, Heath O, Nichols SJ, Trevisan C, Novak M, McDermott J, et al. TGFB1 production by macrophages contributes to an immunosuppressive microenvironment in ovarian cancer. *Cancer Res.* 2021;81(22):5706–19.
30. Parthasarathi KTS, Mandal S, Singh S, Gundimeda S, Jolly MK, Pandey A, Sharma J. In silico analysis of ion channels and their correlation with epithelial to mesenchymal transition in breast cancer. *Cancers (Basel).* 2022;14(6):1444.
31. Prevarskaya N, Skryma R, Shuba Y. Ion Channels in cancer: are cancer hallmarks oncochannelopathies? *Physiol Rev.* 2018;98(2):559–621.
32. Liang X, Zhang N, Pan H, Xie J, Han W. Development of store-operated calcium entry-targeted compounds in cancer. *Front Pharmacol.* 2021;12:688244.
33. Sun J, Lu F, He H, Shen J, Messina J, Mathew R, Wang D, Sarnaik AA, Chang WC, Kim M, et al. STIM1- and Orai1-mediated Ca(2+) oscillation orchestrates invadopodium formation and melanoma invasion. *J Cell Biol.* 2014;207(4):535–48.
34. Tiffner A, Hopf V, Schober R, Sallinger M, Grabmayr H, Hoglinger C, Fahrner M, Lunz V, Maltan L, Frischauf I, et al. Orai1 Boosts SK3 Channel Activation. *Cancers (Basel).* 2021;13(24):6357.
35. Gueguinou M, Crottes D, Chantome A, Rapetti-Mauss R, Potier-Cartereau M, Clarysse L, Girault A, Fourbon Y, Jezequel P, Guerin-Charbonnel C, et al. The SigmaR1 chaperone drives breast and colorectal cancer cell migration by tuning SK3-dependent Ca(2+) homeostasis. *Oncogene.* 2017;36(25):3640–7.
36. James AD, Leslie TK, Kaggie JD, Wiggins L, Patten L, Murphy O'Duinn J, Langer S, Labarthe MC, Riemer F, Baxter G, et al. Sodium accumulation in breast cancer predicts malignancy and treatment response. *Br J Cancer.* 2022;127(2):337–49.
37. Brisson L, Gillet L, Calaghan S, Besson P, Le Guennec JY, Roger S, Gore J. Na(V)15 enhances breast cancer cell invasiveness by increasing NHE1-dependent H(+) efflux in caveolae. *Oncogene.* 2011;30(17):2070–6.
38. Hemmi K, Ma D, Miura Y, Kawaguchi M, Sasahara M, Hashimoto-Tamaoki T, Tamaoki T, Sakata N, Tsuchiya K. A homeodomain-zinc finger protein, ZFHx4, is expressed in neuronal differentiation manner and suppressed in muscle differentiation manner. *Biol Pharm Bull.* 2006;29(9):1830–5.
39. Fevre-Montange M, Champier J, Szathmari A, Wierinckx A, Mottolose C, Guyotat J, Figarella-Branger D, Jouvret A, Lachuer J. Microarray analysis reveals differential gene expression patterns in tumors of the pineal region. *J Neuropathol Exp Neurol.* 2006;65(7):675–84.
40. Wang Z, Zheng X, Wang X, Chen Y, Li Z, Yu J, Yang W, Mao B, Zhang H, Li J, et al. Genetic differences between lung metastases and liver metastases from left-sided microsatellite stable colorectal cancer: next generation sequencing and clinical implications. *Ann Transl Med.* 2021;9(12):967.
41. Ha M, Kim J, Park SM, Hong CM, Han ME, Song P, Kang CD, Lee D, Kim YH, Hur J, et al. Prognostic role of zinc finger homeobox 4 in ovarian serous cystadenocarcinoma. *Genet Test Mol Biomarkers.* 2020;24(3):145–9.
42. Millstein J, Budden T, Goode EL, Anglesio MS, Talhouk A, Intermaggio MP, Leong HS, Chen S, Elatir W, Gilks B, et al. Prognostic gene expression signature for high-grade serous ovarian cancer. *Ann Oncol.* 2020;31(9):1240–50.
43. Chudnovsky Y, Kim D, Zheng S, Whyte WA, Bansal M, Bray MA, Gopal S, Theisen MA, Bilodeau S, Thiru P, et al. ZFHx4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor-initiating cell state. *Cell Rep.* 2014;6(2):313–24.
44. Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M, Campbell K, Cano A, Casanova J, Christofori G, et al. Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* 2020;21(6):341–52.
45. Perez-Oquendo M, Gibbons DL. Regulation of ZEB1 function and molecular associations in tumor progression and metastasis. *Cancers (Basel).* 2022;14(8):1864.
46. Xu J, Lamouille S, Derynck R. TGF-beta-induced epithelial to mesenchymal transition. *Cell Res.* 2009;19(2):156–72.
47. Peng D, Fu M, Wang M, Wei Y, Wei X. Targeting TGF-beta signal transduction for fibrosis and cancer therapy. *Mol Cancer.* 2022;21(1):104.
48. Mayor R, Etienne-Manneville S. The front and rear of collective cell migration. *Nat Rev Mol Cell Biol.* 2016;17(2):97–109.
49. Eskandari N, Senyuk V, Moore J, Kalik Z, Luan Q, Papautsky I, Moshiri A, Bocchetta M, Salami SA, Oryan S, et al. Molecular activation of the Kv111 channel reprograms EMT in colon cancer by inhibiting TGFbeta signaling via activation of calcineurin. *Cancers (Basel).* 2021;13(23):6025.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

