RESEARCH

An EMT-based gene signature enhances the clinical understanding and prognostic prediction of patients with ovarian cancers

Qi-jia Li[1†], Zi-liang Wu[1†], Juan Wang[2], Jing Jiang[1*] and Bing Lin[1*]

Abstract

Background Ovarian cancer (OC) is one of the most common gynecological cancers with malignant metastasis and poor prognosis. Current evidence substantiates that epithelial-mesenchymal transition (EMT) is a critical mechanism that drives OC progression. In this study, we aspire to identify pivotal EMT-related genes (EMTG) in OC development, and establish an EMT gene-based model for prognosis prediction.

Methods We constructed the risk score model by screening EMT genes via univariate/LASSO/step multivariate Cox regressions in the OC cohort from TCGA database. The efficacy of the EMTG model was tested in external GEO cohort, and quantified by the nomogram. Moreover, the immune infiltration and chemotherapy sensitivity were analyzed in different risk score groups.

Results We established a 11-EMTGs risk score model to predict the prognosis of OC patients. Based on the model, OC patients were split into high- and low- risk score groups, and the high-risk score group had an inevitably poor survival. The predictive power of the model was verified by external OC cohort. The nomogram showed that the model was an independent factor for prognosis prediction. Moreover, immune infiltration analysis revealed the immunosuppressive microenvironment in the high-risk score group. Finally, the EMTG model can be used to predict the sensitivity to chemotherapy drugs.

Conclusions This study demonstrated that EMTG model was a powerful tool for prognostic prediction of OC patients. Our work not only provide a novel insight into the etiology of OC tumorigenesis, but also can be used in the clinical decisions on OC treatment.

Keywords Ovarian cancer, EMT, Prognosis, Immune infiltration, Chemotherapy

[†]Qi-jia Li and Zi-liang Wu contributed equally to this work.

*Correspondence:
Jing Jiang
13550060871@163.com
Bing Lin
linbingcdutcm@163.com
[1] Hospital of Chengdu University of Traditional Chinese Medicine, No.39 Shi-er-qiao Road, Chengdu 610072, Sichuan Province, China
[2] Department of Public Health, School of Clinical Medicine, Chengdu University of Traditional Chinese Medicine, Chengdu 610072, China

Background

Ovarian cancer (OC) is one of the most fatal and lethal gynecologic cancers for women [1, 2]. There are more than 295,000 newly diagnosed OC patients and 184,000 deaths per year [3]. Histologically, OC can be split into epithelial, germ cell, and sex-cord-stromal subtypes, among which the epithelial subtype occupies the vast majority (> 90%) [4, 5]. Epithelial OC originates from the ovarian surface epithelium and fallopian tubes [6], with the 5-year survival rate < 50% after diagnosis [7]. A wide range of genetic and environmental factors can increase the risk of OC, including aging, genomic mutation [8],

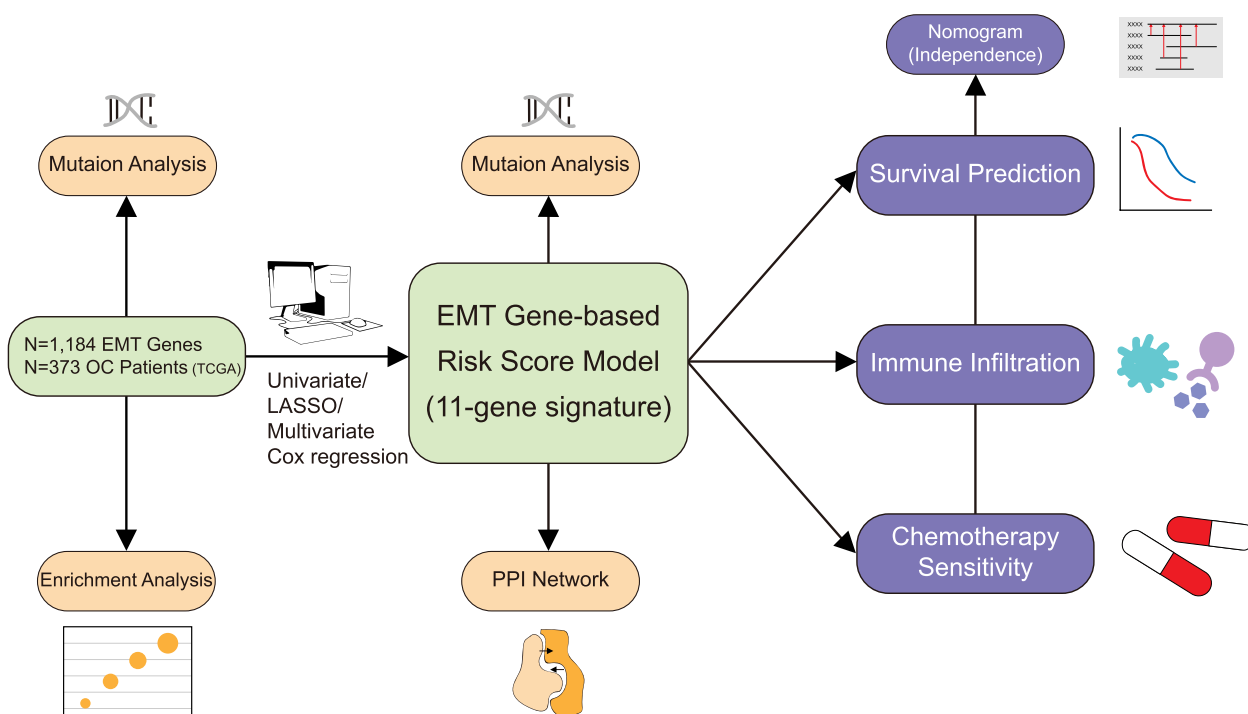Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 2 of 13

Fig. 1 Graphic workflow of this study. A flow chart of this study showing the procedure of identification and exploration of 11-EMT gene-based risk score model in OC

ovulation [9], postmenopausal hormonal therapy use [10], infertility and nulliparity [11]. Despite advanced medical and surgical treatments, including debulking surgery, neoadjuvant therapy, and chemotherapy [12], the long-term survival for OC patients remains not satisfied, mainly due to the failure to detect the tumor at early-stages. In fact, over 70% OC are diagnosed at Stage III or IV [13], with inevitably poor prognosis. Accordingly, there is increasing pressure on developing prognostic models for the survival prediction of OC.

Over decades, preclinical studies have revealed the pathological origins, cellular heterogeneities, and molecular features of OC, with implications in clinical benefit [14]. The invasive characteristic of OC is driven by multiple cellular events, including epithelial-mesenchymal transition (EMT) [15]. EMT is defined as a cellular event characterized by epithelial cells acquiring mesenchymal phenotypes in parallel with loss of epithelial features [16, 17]. Once EMT programs are activated, tumor epithelial cells abolish their polarity and cell-cell adhesion, and gain invasive abilities [18]. Emerging evidence has suggested that EMT not only promotes tumor metastasis, but also increases tumor stemness, cytokine release, cancer-associated angiogenesis, immune escapes, and chemoresistance [18]. Therefore, EMT is widely accepted as a hallmark of cancer [19]. Current understandings have described

multiple factors regulating EMT in OC, including LEFTY1 [20], TRAP1 [21], MDM2 [22], and Rac1 [23]. All these experimental data provide the possibility to build EMT gene (EMTG) models for prognosis prediction of OC.

In this study, we performed comprehensive analysis to establish an EMTG-based risk score model to predict the survival status of OC patients, using the EMTG expression files from public database. Moreover, we applied our EMTG model to evaluate the immune infiltration and chemotherapy sensitivities (Fig. 1). Our work demonstrates the prognostic power of EMTG in OC, and affords a potential tool for improving OS management.

Methods
Dataset and clinical information acquisition
The EMT gene set was obtained from dbEMT database with a total of 1184 EMT-related genes [24, 25]. The transcriptomic files and clinical information were collected from the Cancer Genome Atlas (TCGA) portal (https://portal.gdc.cancer.gov/) and GEO database (https://www.ncbi.nlm.nih.gov/geo/). The TCGA-OC cohort included 373 patients was assigned as the training set, while the GSE26193 testing set contained 107 OC patients.

Functional enrichment and genomic mutation analysis

The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted to demonstrate the EMT gene related biological functions using the "clusterprofiler" package of R software. For mutation analysis of the whole genome or EMT genes, somatic mutation data of OC were downloaded from the TCGA and analyzed by the "maftools" R package. For protein-protein interaction (PPI) analysis, the GeneMANIA database was applied to construct PPI network [26].

Identification and verification of the risk score model

To establish the EMT gene-based model, 1184 EMT-related genes were screened by the univariate Cox regression analysis using "survival" and "survminer" R packages ($p < 0.05$) [27], and 96 genes were retained. The 96 genes were crossed with the differentially expressed genes (DEGs, $n = 5520$) between OC patients and controls, to get a set of 23 genes. Next, the LASSO Cox analysis (by the "glmnet" R package [28]) was conducted to get 16 genes for step wise multivariate Cox regression analysis (by "survival" and "survminer" R packages [27]). Finally, 11 genes were obtained to construct the EMT gene risk score model, and the formula to calculate the probability of inferior survival for each sample was, Risk score = expression of gene $1 \times \beta1$ + expression of gene $2 \times \beta2$ + ⋯ + expression of gene n × βn, (β indicating the coefficients).

Then, the TCGA-OC cohort were divided into low- and high-risk groups according to the median risk score. The Kaplan–Meier (K-M) survival analysis was adopted to compare the survival difference between the two groups. The time-dependent receiver operating characteristic (ROC) analysis (by "timeROC" R package) was used to assess the discrimination of the risk score model. Moreover, the univariate and multivariate Cox analyses were performed to evaluate the independence of risk scores and the prognosis of OC patients. Finally, to verify the efficacy of the risk score model, the above process was applied in additional testing cohort from GSE26193.

Construction of the risk score-based nomogram

An EMT gene-based nomogram was conducted to predict the survival rate of individual OC patient. Clinical factors including the age, disease grade, histological stage and risk score were incorporated using the "rms" package of R software.

Evaluation of the immune infiltrations

The CIBERSORT algorithm was used to evaluate infiltrations of 22 immune cells for each OC patient in TCGA-OC cohort, using the "CIBERSORT" R package [29]. All the TCGA-OC patients were divided into the low- and high-risk score groups to compare the difference in immune cell compositions. The p value of each sample deconvolution was determined by CIBERSORT, and samples with $p < 0.05$ was screened out for box-plot graphing.

Correlation analysis of the risk score with drug sensitivity

To study the relationship between the risk score and drug sensitivity, the CellMiner database was applied (v2.7, https://discover.nci.nih.gov/cellminer/home.do) [30]. The transcriptomic data of 60 cancer cell lines and the IC50 of over 20, 000 drugs (approved by FDA or under clinical trials) were interrogated to determine the correlative significance with the |cor| > 0.3 and $p < 0.01$, and visualized by scatter plot using the "ggplot2" R package, which is performed by a well-established pipeline in [31–33].

Statistical analysis

All statistical analyses were performed by the R software (version 4.0.3). The statistical significance was established with the p value < 0.05 without additional statement.

Results

Functional enrichment and genomic mutation of EMT genes related to OC

To deeply explore the EMT gene (EMTG) alterations in OC, we investigated their biological functions and genomic features. First, we applied GO and KEGG analysis to detect the functional enrichment of 1184 EMTG. Results of the GO analysis showed that the top terms were enriched in tissue development and cell proliferation, including mesenchyme development, epithelial cell proliferation and mesenchymal cell differentiation (Fig. 2A). KEGG analysis further demonstrated that the enriched pathways were consisted of multiple cancer items, focal adhesion, oncogenic PI3K/AKT, Hippo and FoxO signalings (Fig. 2B), supporting the notion that EMT is correlated with multiple human cancers and interconnected with oncogenic pathways.

Next, we examined the mutation frequency of 1184 EMTG in TCGA-OC cohort, to address whether genomic mutation is a leading cause of the EMT in OC. It's found that EMTG mutation occurred in 95.56% OC patients, and the top 20 mutation frequency ranging from 2 to 92% (Fig. 2C). The mutation frequency of *TP53* reached 92%, which may be enriched in the Type II of epithelial ovarian cancers [34, 35]. However, the mutation frequency of other genes was lower than 10%, including *BRCA1* (4%), *ATM* (2%), *MTOR* (2%), and *LATS1* (2%) (Fig. 2C and Supplementary Fig. S1). Similarly, results of the cBioPortal database showed that the major mutation
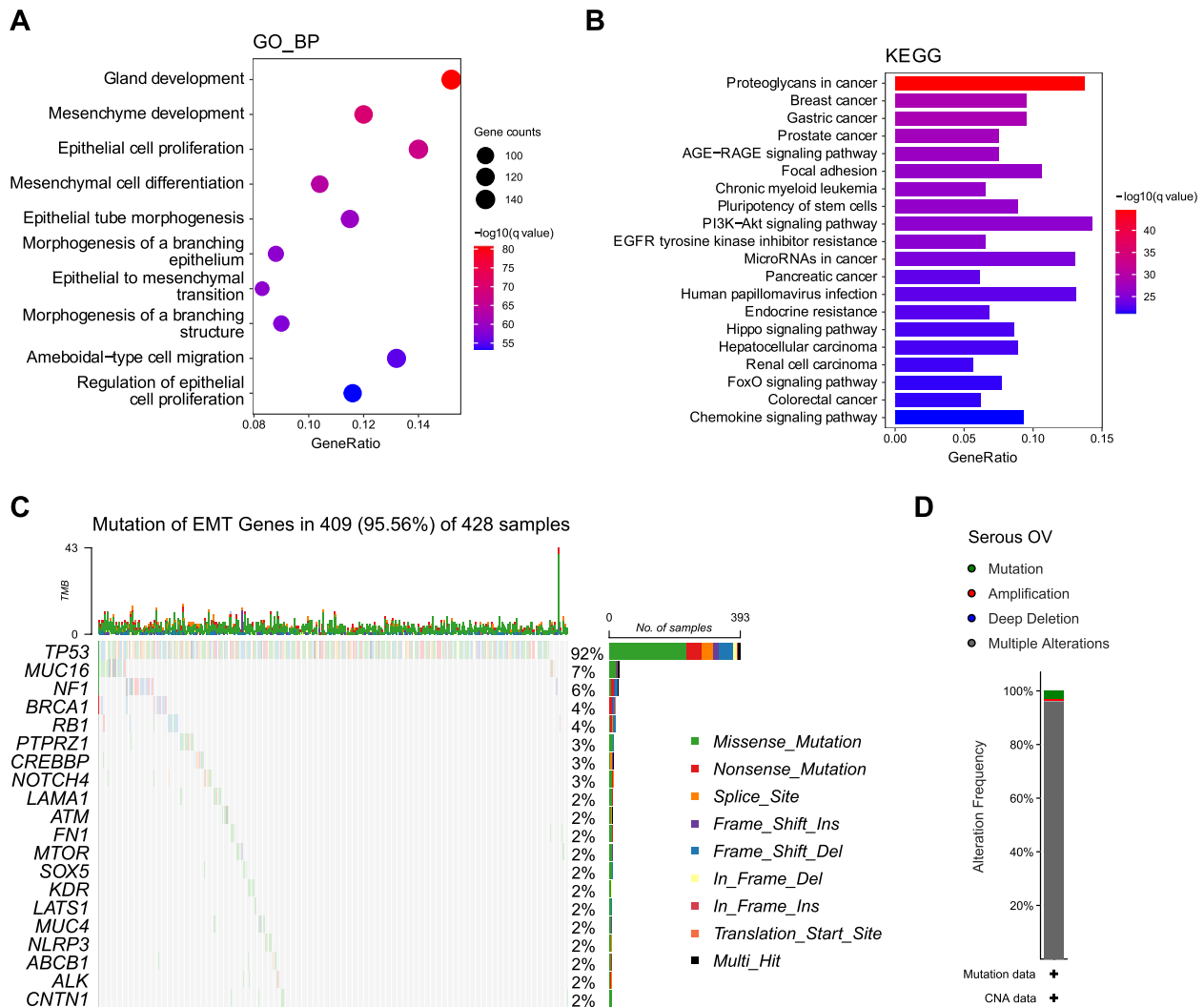
Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 4 of 13

Fig. 2 Functional enrichment and genomic mutation of EMT genes related to OC. **A-B** GO (**A**) and KEGG (**B**) analysis showing the functional enrichment of EMT genes. **C** The waterfall plots showing the top 20 mutated EMT genes in the OC cohort of TCGA. Each row represents a mutated gene, and each column represents a patient. **D** The alteration frequency of EMT genes in OC patients from cBioPortal database

style of the EMTG was Multiple Alterations, not Amplification or Deep Deletion (Fig. 2D and Supplementary Fig. S1). Thus, most EMTG (except *TP53*) tend to preserve genomic stabilities. All these results suggest that altered EMTG expression, but not genomic mutation, may be the major cause of dysregulation of EMTG in OC.

Identification and validation of the EMTG risk score model in OC

Based on above analysis, we constructed the risk score model, by importing EMTGs into a multiple step screening pipeline including the univariate Cox regression, LASSO Cox regression and multivariate Cox analysis with stepwise regression in TCGA-OC cohort (Fig. 1 and

Supplementary Fig. S2). Afterwards, the 11-gene EMTG signature were obtained comprised of *ESR2, PKD1, CXCL9, TNFSF11, FOXN1, LEFTY1, MYL2, MMP7, SFRP2, CXCR4,* and *HMGB3* (Fig. 3A). All the 11 EMTG were differentially expressed between tumor tissues and normal controls. Evidently, expressions of *HMGB3, FOXN1, TNFSF11, CXCR4, MMP7, CXCL9* and *SFRP2* were higher in tumors, while *PKD1, MYL2, ESR2* and *LEFTY1* were higher in the normal controls (Fig. 3B).

Next, the Kaplan-Meier (K-M) survival analysis revealed that the overall survival (OS) of the low-risk score group was much better than the high-risk group (Fig. 3C). The ROC analysis confirmed the predictive ability of the EMTG signature, showing the AUC values
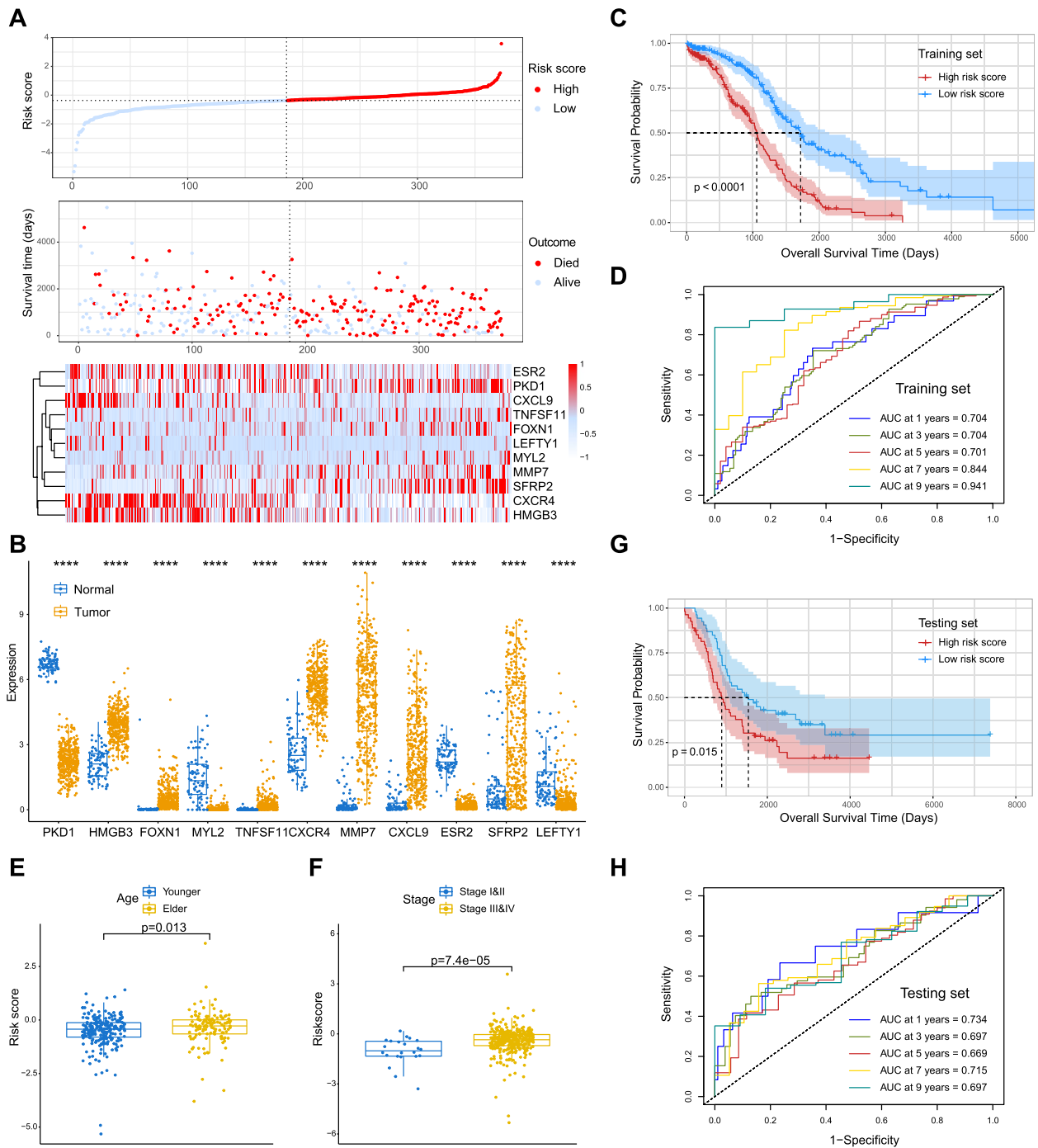
Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 5 of 13

Fig. 3 Identification and validation of the EMTG risk score model in OC. **A** The heatmap showing the distribution of 11 EMTG genes and survival status for each patient of OC in TCGA (training set). **B** The box plots showing the 11 EMTG expression levels in tumors and controls in TCGA. **C** Kaplan–Meier analysis showing the distinct survival tendency in low- and high-risk score groups in TCGA (training set). **D** ROC curves showing the predictive efficiency of EMTG model for the survival rate in TCGA (training set). **E-F** Diagrams showing the correlations of risk score with patient age (**E**) and histological stages (**F**) in TCGA cohort. **G** Kaplan–Meier analysis showing the distinct survival tendency in low- and high-risk score groups in GSE26193 (testing set). **H** ROC curves showing the predictive efficiency of EMTG model for the survival rate in GSE26193 set

Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 6 of 13

of 0.704 (1-year), 0.704 (3-year), 0.701 (5-year), 0.844 (7-year), and 0.941 (9-year) respectively (Fig. 3D). Moreover, we split the OS patients into distinct subgroups as per the patient age, WHO grade and histological stage, and found that the risk score model could still pinpoint different outcome in subgroups. Particularly, the low-risk score group consistently represented better OS than the high-risk score group (Supplementary Fig. 2B-G). All these data clearly demonstrate the robustness of our EMTG risk score model.

To verify the accuracy of the EMTG model, we further validated it in external GEO cohort (GSE26193). Likewise, the OS patients ($n = 107$) were stratified into low- and high- risk score groups cut off by the median risk score. The K-M survival analysis revealed that the low-risk score group had better OS than the high-risk score group (Fig. 3G). The ROC analysis showed that the EMTG signature performed well in the prognostic prediction, with the AUC values of 0.734 (1-year), 0.697 (3-year), 0.669 (5-year), 0.715 (7-year), and 0.697 (9-year) respectively (Fig. 3H). Collectively, all these findings indicate that our EMTG risk score model could predict the survival status of OC patients with sufficient power.

Assessment of the prognostic independency of the EMTG risk score model in OC

In view of the predictive ability of EMTG model in OC, we attempted to investigate whether the risk score model was an independent factor for the survival prediction of OC. Accordingly, we performed univariate and multivariate Cox regression analyses to evaluate the predictive power of OS-related parameters including age, stage, grade, and risk score. The univariate Cox analysis demonstrated that the age, stage, and risk score were statistically associated with the prognosis ($p < 0.05$) (Fig. 4A). The multivariate Cox analysis further indicated that only the age and risk score were consistently associated with the prognosis ($p < 0.05$) (Fig. 4B), and apparently, the risk score had the highest predictive power among all the parameters (Fig. 4C-D). Next, we tried to facilitate the clinical use of the EMTG model, and conducted a nomogram to predict the 3-, 5-, and 7- year survival status of OS patients (Fig. 4E). The calibration charts revealed the consistency of predictive and actual survival status for all the 3-, 5-, and 7- years (Fig. 4F-H). Thus, the EMTG signature is proved to be an independent factor for survival prediction, and its effectiveness is verified by multiple approaches.

Investigation of molecular features in distinct EMTG risk score groups of OC

To deeply explore the molecular difference between low- and high- risk score groups, we next studied genomic

alterations by assessment of the DNA mutations. We found that genomic mutations existed in 96.97% of the low-risk score group and 94.33% of the high-risk score group. The top enriched genes were *TP53*, *TTN*, *CSMD3*, *MUC16*, which displayed no obvious difference between low- and high- risk score groups (Fig. 5A-B). However, it's noted that the high-risk score group had more mutations on *TOP2A* (6%), *CDK12* (6%), and *BRCA1* (4%), by which may contribute to its poor prognosis.

To further understand how the 11 screened EMTGs contributed to OC development, we investigated their functions by protein-protein interaction (PPI) analysis. As shown, the 11 EMTGs were located in the center and surrounded by 20 hub proteins (Fig. 5C). Interestingly, we found that members of Wnt family (Wnt4, PAX2, Wnt1, SFRP5 etc.) were dramatically enriched in PPI network, which is widely known as a master regulator of EMT programs [36]. All these analyses suggest that genomic mutation and interactive network involving Wnt may contribute to the EMT regulation and OC progression.

Identification of the immune landscape in distinct EMTG risk score groups of OC

The heterogenous immune microenvironment is an important orchestrator of OC development, and determines the prognosis [37]. Thus, we examined whether the different EMTG risk score groups displayed different immune microenvironment. Using the CIBERSORT algorithm, we evaluated alterations of 22 immune cells in TCGA-OC cohort. The distribution of each type immune cells was displayed in the diagram with variant abundant ratios for each sample (Fig. 6A). Apparently, there were 7 types of immune cells showed statistical difference between the low- and high- risk score groups ($p < 0.05$) (Fig. 6B). Compared with the low-risk score group, the high-risk score group held decreased levels of M1 macrophages, CD8 T cells, CD4 memory activated T cells and T follicular helper cells, together with increased M2 macrophages, resting NK cells, and CD4 memory resting T cells (Fig. 6B). These results suggests that the microenvironment of the high-risk score group is more immunosuppressive, which may contribute to its poor prognosis.

Prediction of drug sensitivity in distinct EMTG risk score groups of OC

Chemotherapy is an indispensable OC management, however, chemoresistance results in treatment failure and tumor relapse [38]. To better understand the molecular mechanisms of chemoresistance in OC, we examined whether the EMTG model could predict drug sensitivity to chemotherapeutic agents. Thereby, we calculated the correlations of risk score to IC50 of drugs from the CellMiner database.
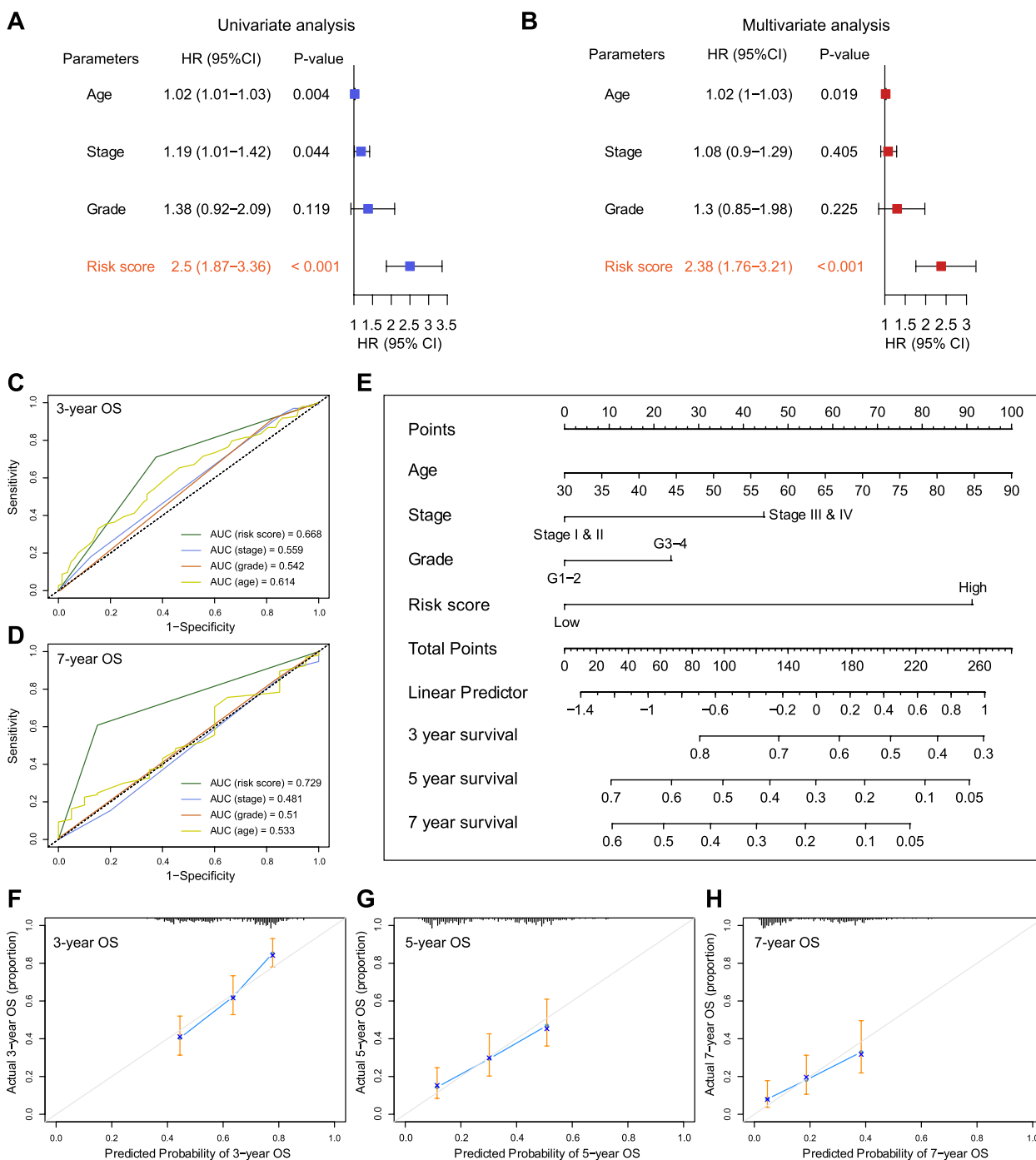
Fig. 4 Assessment of the prognostic independency of EMTG risk score model in OC. **A-B** Univariate (**A**) and multivariate (**B**) Cox regression analysis showing that EMTG risk score was an independent factor for OS prediction in OC patients. **C-D** ROC curves showing the predictive efficiency of the EMTG risk score and clinical factors for 3- (**C**) and 5- (**D**) year survival rate in TCGA. **E** A nomogram showing the predicting probabilities of OC patients with 3-, 5-, and 7-year OS in TCGA. **F-H** Calibration plots showing the prediction of OC patients with 3- (**F**), 5- (**G**), and 7- (**H**) year OS in TCGA. X-axis indicating nomogram-predicted probability of survival; Y-axis indicating the actual survival

Results revealed different tendencies of the risk score to chemotherapy drugs. For instance, the IC50 of XAV-939, Bleomycin, Staurosporine and AT-9283 were increased as the elevated risk scores (Fig. 7A-B), while the IC50 of Lapatinib, Nilotinib, Luminespib and Volasertib were decreased as the elevated risk scores (Fig. 7E-H).
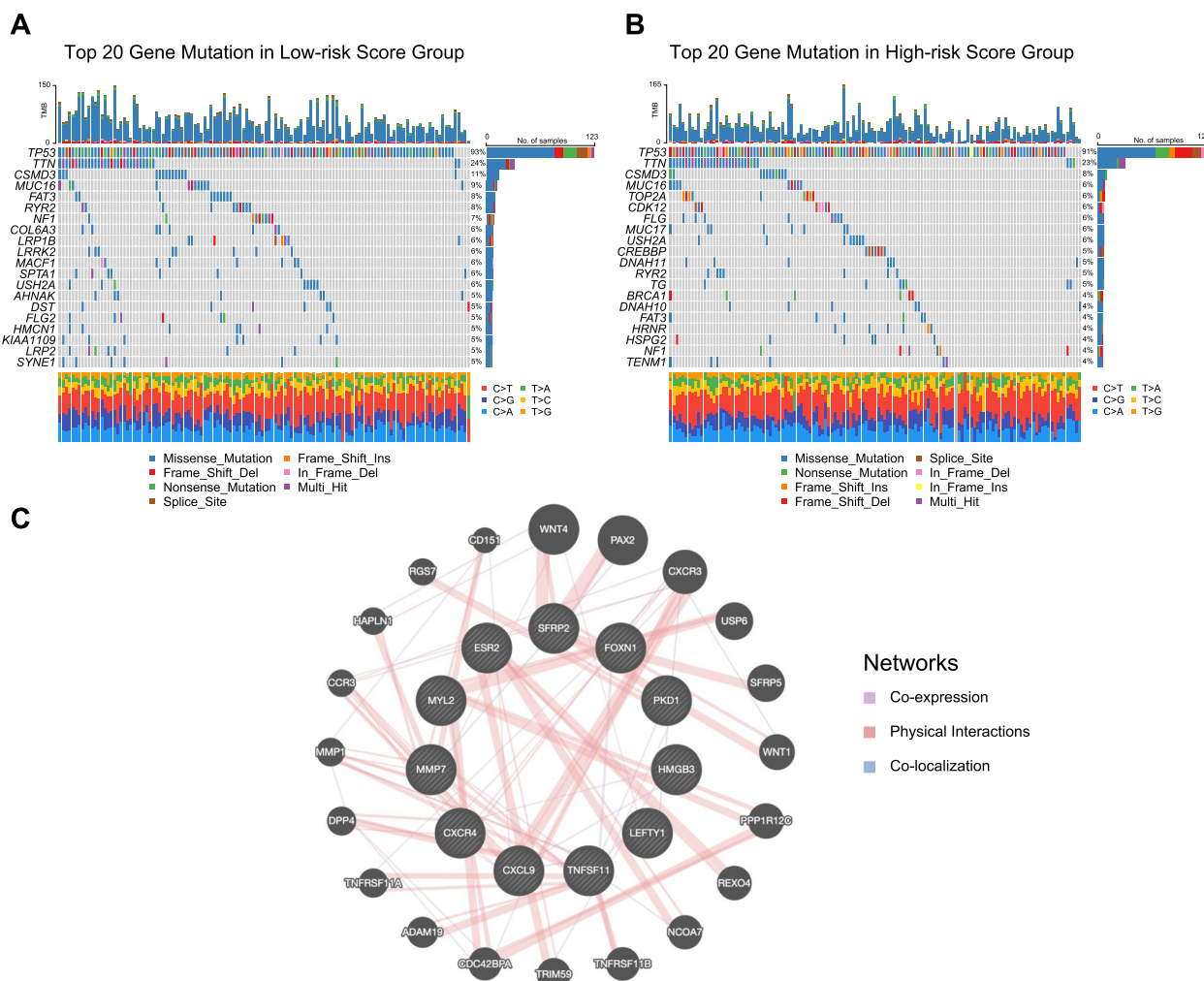
Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 8 of 13

Fig. 5 Investigation of molecular features in distinct EMTG risk score groups of OC. **A** The waterfall plots showing the top 20 gene mutations in the low-risk score group of OC. **B** The waterfall plots showing the top 20 gene mutations in the high-risk score group of OC. **C** A PPI diagram showing the protein-protein interaction of the 11-EMTG genes. The inner circle represents the 11 EMTG genes, and the outer circle represents the hub interactive genes calculated by GeneMANIA

Therefore, the former set of drugs were impactful for the low-risk score patients, and the later may be more effective for the high-risk score patients. Nevertheless, our EMTG risk score model is potential to play a role in the selection of chemotherapy drugs for individual OC patient.

Discussion

OC is one of the most prevalent cancers for women in the world, characterized by its metastatic aggressiveness and unfavorable prognosis [13]. The rapid growth and invasion of OC is strongly regulated by epithelial-mesenchymal transition (EMT) [15]. In this study, we established a robust EMT-related gene model to predict the survival status of OC patients, which can be used in clinical management for prognosis surveillance.

EMT was originally identified by Betty Hay et al. in vertebrate embryonic development in 1960s [39, 40]. Over decades, current studies have established the underlying mechanisms by which EMT regulates OC malignant behaviors [41]. These experimental works provide the rationale for prognostic prediction of OC using EMT related genes. Actually, manipulating a series of genes to predict the prognosis is more accurate and definitive than only one single gene model [42]. Based on this assumption, previous studies have employed EMT genes to predict the prognosis of human cancers, including head and neck squamous cell cancer (HNSCC) [43], hepatocellular carcinoma [44], endometrial carcinoma [45], and colon adenocarcinoma [46]. But the predictive power of EMT genes in OC prognosis has not yet been evaluated. Therefore, we acquired the EMT gene expression profiles from
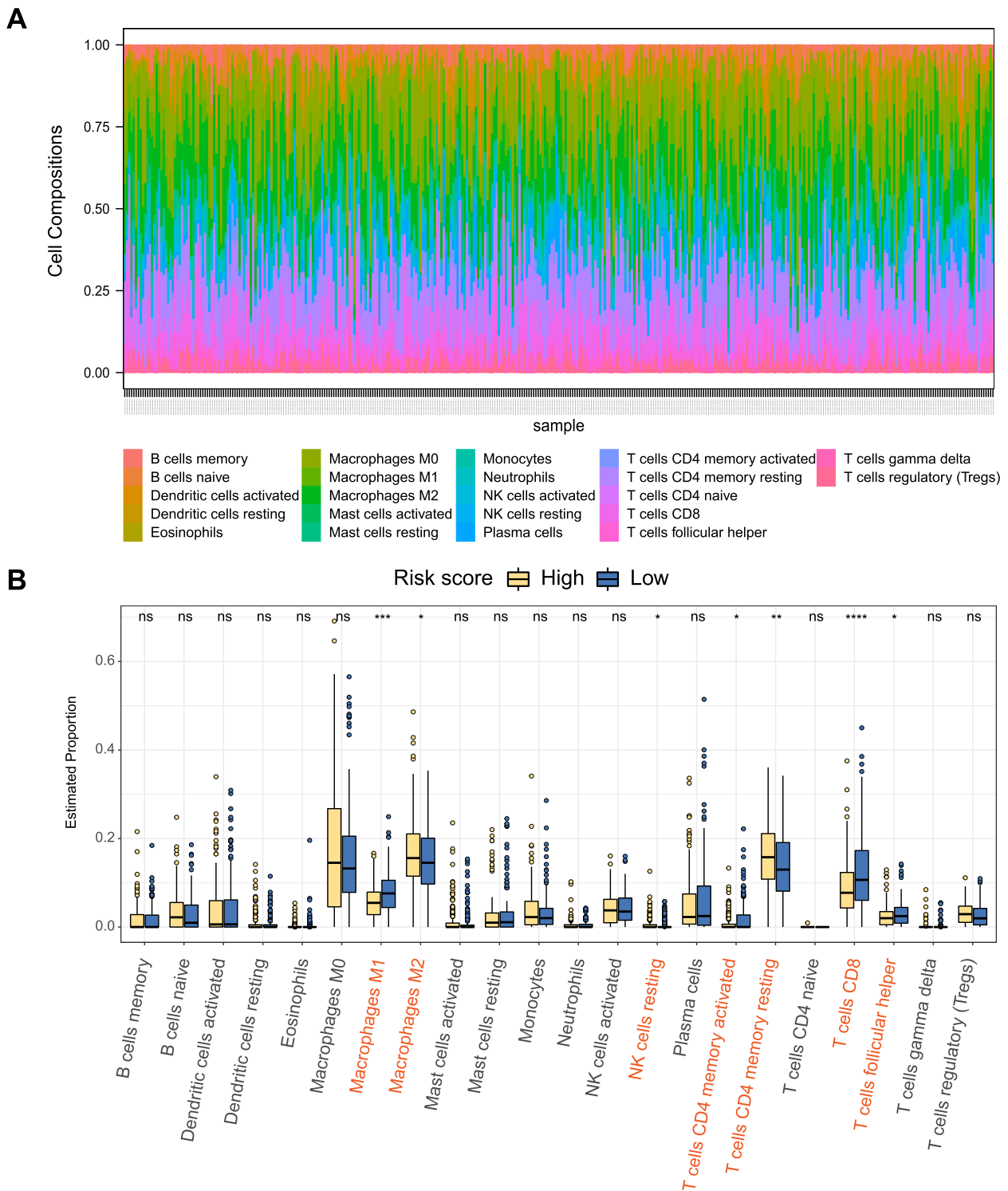
Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 9 of 13

Fig. 6 Identification of immune landscape in distinct EMTG risk score groups of OC. **A** A diagram showing the composition of 22 immune cells in OC cohort of TCGA. Colorful bars with different lengths represent individual immune cell proportions. **B** Box plots showing the difference of immune cell populations in low- and high-risk score groups in OC cohort of TCGA
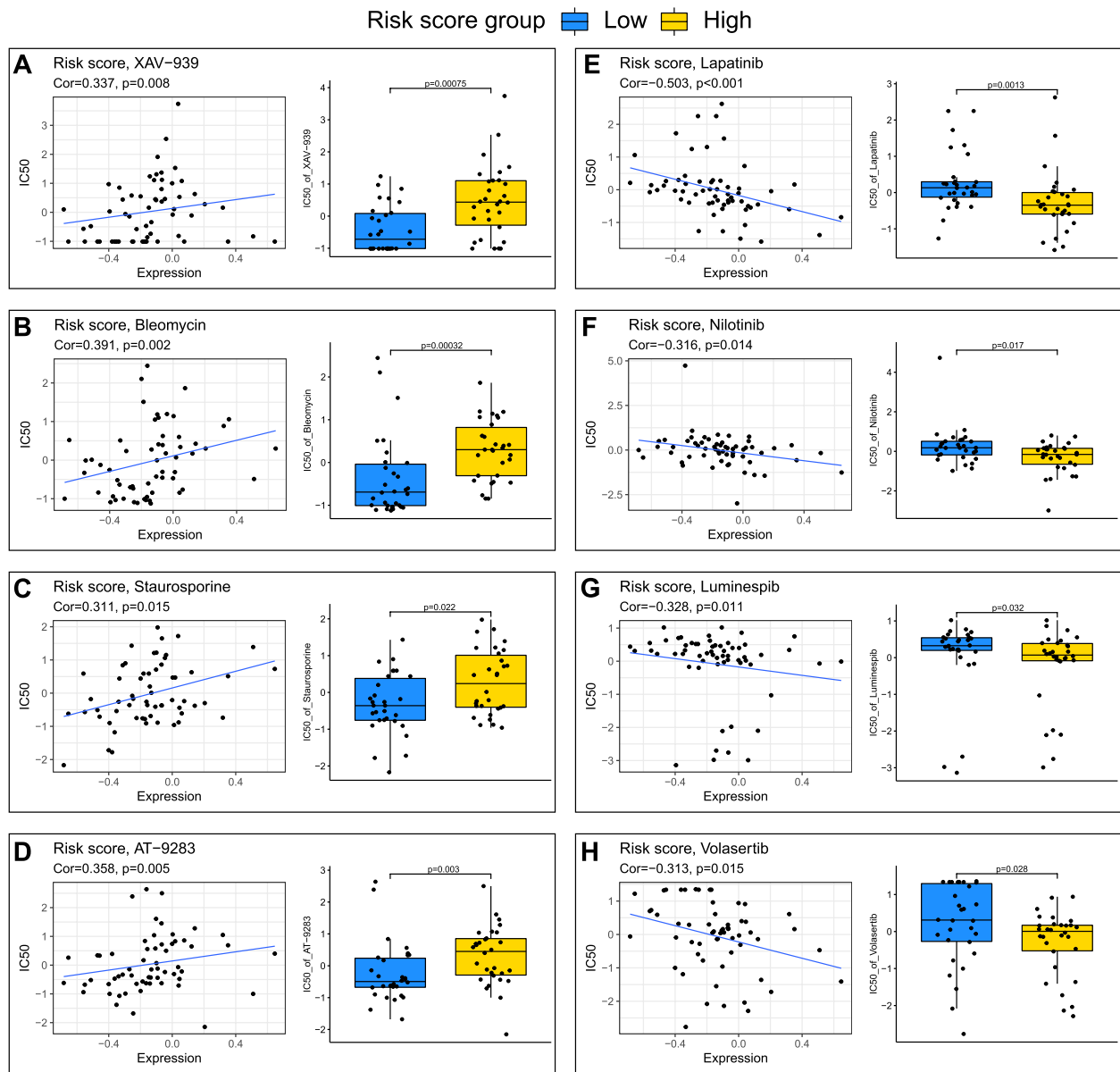
Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 10 of 13

Fig. 7 Prediction of drug sensitivity in distinct EMTG risk score groups of OC. **A**-**H** Scatter and box plots showing the correlations of EMTG risk score with IC50 of XAV-939 (**A**), Bleomycin (**B**), Staurosporine (**C**), AT-9283 (**D**), Lapatinib (**E**), Nilotinib (**F**), Luminespib (**G**), and Volasertib (**H**)

TCGA and constructed an EMT gene-based risk score model to improve OC prognosis. Our risk score model was successfully validated by independent dataset.

During the model construction, we noticed that the set of EMT genes had low frequency of genomic mutations (Fig. 2). Instead, EMT gene expressions were dynamically regulated at the transcriptional level, supported by the fact that these genes were differentially expressed in the tumor tissues and normal controls. This point was further enhanced by previous findings showing that distinct EMT gene expression existed in morphologically indistinguishable samples from primary and primary and metastatic ovarian tumors [47]. The dynamic expressions of EMT genes are actually controlled by upstream regulations, including epigenetic modulation [48–50]. Therefore, understanding the molecular alteration of EMT genes will not only provide prognostic values, but also afford druggable targets in OC interventions.

In addition to provide a clinical tool for prognosis prediction, our EMT-related model also clarifies why the high-risk score group had rather poor survival. At the immune level, the immune infiltration analysis

Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 11 of 13

revealed that immunosuppressive cells were increased in the high-risk group (Fig. 6). The interconnective network of EMT signature genes with oncogenic signaling (e.g., Wnt) [36] exacerbates the malignancy in the high-risk group (Fig. 5). At the molecular level, the high-risk score group had elevated expressions of tumor enhancing factors, including *MMP7* [51], and *SFRP2* [52] (Fig. 3), which may synergistically promote tumor growth, proliferation, invasion and chemoresistance. For example, *MMP7* encodes matrix metalloproteinase-7 and facilitates the invasion and migration of ovarian tumor cells by remodeling the extracellular matrix [53, 54]. On the contrary, the low-risk score group had elevated expressions of tumor suppressing factors, including *ESR2* [55, 56]. For example, *ESR2* encode the estrogen receptor beta, and appeared to inhibit the EMT induced by E2 (17β-estradiol) in OC [55]. Moreover, elevated *ESR2* increases ERβ expression to block EMT and OC metastasis [56]. These experimental results are consistent with our bioinformatic analysis, showing that ESR2 expression is decreased as the risk score increasing. All these findings collaboratively address the reason for poor survival in high-risk score group of OC.

Nevertheless, our 11 EMT-gene based risk score model could improve the current prediction of survival status of OC patients. In fact, the predictive power of the clinical factors including age, WHO grade and histological stage are not satisfied, and our EMT-related risk score performed best compared with all these factors (Fig. 4). To the best of our knowledge, our study is the unique model using EMT genes in OC prognosis. Thus, our work will improve current understandings on gene set-based prognostic modeling, in parallel with facilitating the optimization of chemotherapeutic strategies in clinical practice (Fig. 7).

Conclusion

Collectively, our study established a novel EMT-based gene signature for OC, and revealed that the model was a reliable tool for prognostic prediction, with mechanistic indications of immune infiltration features and chemotherapeutic sensitivities.

Abbreviations

OC	Ovarian cancer
EMT	Epithelial-mesenchymal transition
EMTG	Epithelial-mesenchymal transition related gene
TCGA	The Cancer Genome Atlas
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
ROC	Receiver operating characteristic
AUC	Area under the ROC curve
PPI	Protein-protein interaction

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13048-023-01132-2.

Additional file 1: Figure S1. The genomic mutation of EMT genes in OC cohort. The landscape of genomic mutations in OC cohort of TCGA, showing the variant classification (A), variant type (B), SNV class (C), variants per sample (D), variant classification (E), and top 10 mutated genes (F). **Figure S2.** Construction and exploration of the EMTG risk score model in OC; related to Fig. 3. (A) The procedure of LASSO Cox regression analysis for construction of the EMTG model in OC cohort of TCGA. (B-G) Kaplan–Meier analysis showing the distinct survival tendency in low- and high-risk score groups in sub-set of the TCGA cohort, including the young set (age ≤ 65, B), old set (age > 65, C), grade 1/2 set (D), grade 3/4 set (E), stage I/II set (F), and stage III/IV set (G).

Availability of data and materials
The datasets used and/or analyzed during the current study are available on the open repositories, including TCGA and GEO. 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
All gene expression datasets were collected from public database. All data can be downloaded and analyzed freely, and not require additional ethics statements.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References

1. Matulonis UA, Sood AK, Fallowfield L, Howitt BE, Sehouli J, Karlan BY. Ovarian cancer. Nat Rev Dis Primers. 2016;2:16061.
2. Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. Lancet. 2014;384(9951):1376–88.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.

Li *et al. Journal of Ovarian Research*   (2023) 16:51

Page 12 of 13

4. Lheureux S, Gourley C, Vergote I, Oza AM. Epithelial ovarian cancer. Lancet. 2019;393(10177):1240–53.
5. Kurnit KC, Fleming GF, Lengyel E. Updates and new options in advanced epithelial ovarian Cancer treatment. Obstet Gynecol. 2021;137(1):108–21.
6. Shih IM, Wang Y, Wang TL. The origin of ovarian Cancer species and precancerous landscape. Am J Pathol. 2021;191(1):26–39.
7. Doherty JA, Peres LC, Wang C, Way GP, Greene CS, Schildkraut JM. Challenges and opportunities in studying the epidemiology of ovarian Cancer subtypes. Curr Epidemiol Rep. 2017;4(3):211–20.
8. Alsop K, Fereday S, Meldrum C, deFazio A, Emmanuel C, George J, et al. BRCA mutation frequency and patterns of treatment response in BRCA mutation-positive women with ovarian cancer: a report from the Australian ovarian Cancer study group. J Clin Oncol. 2012;30(21):2654–63.
9. Walker JL, Powell CB, Chen LM, Carter J, Bae Jump VL, Parker LP, et al. Society of Gynecologic Oncology recommendations for the prevention of ovarian cancer. Cancer. 2015;121(13):2108–20.
10. Morch LS, Lokkegaard E, Andreasen AH, Kruger-Kjaer S, Lidegaard O. Hormone therapy and ovarian cancer. JAMA. 2009;302(3):298–305.
11. Rice MS, Hankinson SE, Tworoger SS. Tubal ligation, hysterectomy, unilateral oophorectomy, and risk of ovarian cancer in the Nurses' health studies. Fertil Steril. 2014;102(1):192–8 e3.
12. Melamed A, Rauh-Hain JA, Gockley AA, Nitecki R, Ramirez PT, Hershman DL, et al. Association between overall survival and the tendency for Cancer programs to administer neoadjuvant chemotherapy for patients with advanced ovarian Cancer. JAMA Oncol. 2021;7(12):1782–90.
13. Stewart C, Ralyea C, Lockwood S. Ovarian Cancer: an integrated review. Semin Oncol Nurs. 2019;35(2):151–6.
14. McMullen M, Karakasis K, Rottapel R, Oza AM. Advances in ovarian cancer, from biology to treatment. Nat Can. 2021;2(1):6–8.
15. Vergara D, Merlot B, Lucot JP, Collinet P, Vinatier D, Fournier I, et al. Epithelial-mesenchymal transition in ovarian cancer. Cancer Lett. 2010;291(1):59–66.
16. Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M, et al. Guidelines and definitions for research on epithelial-mesenchymal transition. Nat Rev Mol Cell Biol. 2020;21(6):341–52.
17. Greenburg G, Hay ED. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. J Cell Biol. 1982;95(1):333–9.
18. Derynck R, Weinberg RA. EMT and Cancer: more than meets the eye. Dev Cell. 2019;49(3):313–6.
19. Hanahan D. Hallmarks of Cancer: new dimensions. Cancer Discov. 2022;12(1):31–46.
20. Matsumoto T, Yokoi A, Hashimura M, Oguri Y, Akiya M, Saegusa M. TGF-beta-mediated LEFTY/Akt/GSK-3beta/snail axis modulates epithelial-mesenchymal transition and cancer stem cell properties in ovarian clear cell carcinomas. Mol Carcinog. 2018;57(8):957–67.
21. Amoroso MR, Matassa DS, Agliarulo I, Avolio R, Lu H, Sisinni L, et al. TRAP1 downregulation in human ovarian cancer enhances invasion and epithelial-mesenchymal transition. Cell Death Dis. 2016;7(12):e2522.
22. Chen Y, Wang DD, Wu YP, Su D, Zhou TY, Gai RH, et al. MDM2 promotes epithelial-mesenchymal transition and metastasis of ovarian cancer SKOV3 cells. Br J Cancer. 2017;117(8):1192–201.
23. Fang D, Chen H, Zhu JY, Wang W, Teng Y, Ding HF, et al. Epithelial-mesenchymal transition of ovarian cancer cells is sustained by Rac1 through simultaneous activation of MEK1/2 and Src signaling pathways. Oncogene. 2017;36(11):1546–58.
24. Zhao M, Liu Y, Zheng C, Qu H. dbEMT 2.0: an updated database for epithelial-mesenchymal transition genes with experimentally verified information and precalculated regulation information for cancer metastasis. J Genet Genomics. 2019;46(12):595–7.
25. Zhao M, Kong L, Liu Y, Qu H. dbEMT: an epithelial-mesenchymal transition associated gene resource. Sci Rep. 2015;5:11459.
26. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res. 2010;38(Web Server issue):W214–20.
27. Zhang Z, Reinikainen J, Adeleke KA, Pieterse ME, Groothuis-Oudshoorn CGM. Time-varying covariates and coefficients in cox regression models. Ann Transl Med. 2018;6(7):121.
28. Tibshirani R. The lasso method for variable selection in the cox model. Stat Med. 1997;16(4):385–95.
29. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773–82.
30. Reinhold WC, Sunshine M, Liu H, Varma S, Kohn KW, Morris J, et al. CellMiner: a web-based suite of genomic and pharmacologic tools to explore transcript and drug patterns in the NCI-60 cell line set. Cancer Res. 2012;72(14):3499–511.
31. Li QJ, Wang J, Jiang J, Lin B. Identification and validation of a gene-based signature reveals SLC25A10 as a novel prognostic indicator for patients with ovarian cancer. J Ovarian Res. 2022;15(1):106.
32. Wu X, Li S, Chen D, Zheng G, Zhang Z, Li Z, et al. An inflammatory response-related gene signature associated with immune status and prognosis of acute myeloid leukemia. Am J Transl Res. 2022;14(7):4898–917.
33. Wang W, Pan F, Lin X, Yuan J, Tao C, Wang R. Ferroptosis-related hub genes in hepatocellular carcinoma: prognostic signature, immune-related, and drug resistance analysis. Front Genet. 2022;13:907331.
34. Zhang Y, Cao L, Nguyen D, Lu H. TP53 mutations in epithelial ovarian cancer. Transl Cancer Res. 2016;5(6):650–63.
35. Shih Ie M, Kurman RJ. Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. Am J Pathol. 2004;164(5):1511–8.
36. Shin SY, Rath O, Zebisch A, Choo SM, Kolch W, Cho KH. Functional roles of multiple feedback loops in extracellular signal-regulated kinase and Wnt signaling pathways that regulate epithelial-mesenchymal transition. Cancer Res. 2010;70(17):6715–24.
37. Jimenez-Sanchez A, Memon D, Pourpe S, Veeraraghavan H, Li Y, Vargas HA, et al. Heterogeneous tumor-immune microenvironments among differentially growing metastases in an ovarian Cancer patient. Cell. 2017;170(5):927–38 e20.
38. Kim S, Han Y, Kim SI, Kim HS, Kim SJ, Song YS. Tumor evolution and chemoresistance in ovarian cancer. NPJ Precis Oncol. 2018;2:20.
39. Trelstad RL, Revel JP, Hay ED. Tight junctions between cells in the early chick embryo as visualized with the electron microscopy. J Cell Biol. 1966;31(1):C6–10.
40. Hay ED, Fischman DA. Origin of the blastema in regenerating limbs of the newt Triturus viridescens. An autoradiographic study using tritiated thymidine to follow cell proliferation and migration. Dev Biol. 1961;3:26–59.
41. Kan T, Wang W, Ip PP, Zhou S, Wong AS, Wang X, et al. Single-cell EMT-related transcriptional analysis revealed intra-cluster heterogeneity of tumor cell clusters in epithelial ovarian cancer ascites. Oncogene. 2020;39(21):4227–40.
42. Latha NR, Rajan A, Nadhan R, Achyutuni S, Sengodan SK, Hemalatha SK, et al. Gene expression signatures: a tool for analysis of breast cancer prognosis and therapy. Crit Rev Oncol Hematol. 2020;151:102964.
43. Xin W, Zhao C, Jiang L, Pei D, Zhao L, Zhang C. Identification of a novel epithelial-mesenchymal transition gene signature predicting survival in patients with HNSCC. Pathol Oncol Res. 2021;27:585192.
44. Shi Y, Wang J, Huang G, Zhu J, Jian H, Xia G, et al. A novel epithelial-mesenchymal transition gene signature for the immune status and prognosis of hepatocellular carcinoma. Hepatol Int. 2022;16(4):906–17.
45. Ruan T, Wan J, Song Q, Chen P, Li X. Identification of a novel epithelial-mesenchymal transition-related gene signature for endometrial carcinoma prognosis. Genes (Basel). 2022;13(2). https://www.mdpi.com/2073-4425/13/2/216.
46. Busuioc C, Ciocan-Cartita CA, Braicu C, Zanoaga O, Raduly L, Trif M, et al. Epithelial-mesenchymal transition gene signature related to prognostic in Colon adenocarcinoma. J Pers Med. 2021;11(6). https://www.mdpi.com/2073-4425/13/2/216.
47. Lili LN, Matyunina LV, Walker LD, Wells SL, Benigno BB, McDonald JF. Molecular profiling supports the role of epithelial-to-mesenchymal transition (EMT) in ovarian cancer metastasis. J Ovarian Res. 2013;6(1):49.
48. Al-Radhawi MA, Tripathi S, Zhang Y, Sontag ED, Levine H. Epigenetic factor competition reshapes the EMT landscape. Proc Natl Acad Sci U S A. 2022;119(42):e2210844119.
49. Nowak E, Bednarek I. Aspects of the epigenetic regulation of EMT related to Cancer metastasis. Cells. 2021;10(12). https://www.mdpi.com/2073-4409/10/12/3435.
50. Chung VY, Tan TZ, Ye J, Huang RL, Lai HC, Kappei D, et al. The role of GRHL2 and epigenetic remodeling in epithelial-mesenchymal plasticity in ovarian cancer cells. Commun Biol. 2019;2:272.

51. Brun JL, Cortez A, Commo F, Uzan S, Rouzier R, Darai E. Serous and mucinous ovarian tumors express different profiles of MMP-2, −7, −9, MT1-MMP, and TIMP-1 and -2. Int J Oncol. 2008;33(6):1239–46.
52. Duan H, Yan Z, Chen W, Wu Y, Han J, Guo H, et al. TET1 inhibits EMT of ovarian cancer cells through activating Wnt/beta-catenin signaling inhibitors DKK1 and SFRP2. Gynecol Oncol. 2017;147(2):408–17.
53. Wang FQ, So J, Reierstad S, Fishman DA. Matrilysin (MMP-7) promotes invasion of ovarian cancer cells by activation of progelatinase. Int J Cancer. 2005;114(1):19–31.
54. Zhang H, Wang Y, Chen T, Zhang Y, Xu R, Wang W, et al. Aberrant activation of hedgehog Signalling promotes cell migration and invasion via matrix Metalloproteinase-7 in ovarian Cancer cells. J Cancer. 2019;10(4):990–1003.
55. Park SH, Cheung LW, Wong AS, Leung PC. Estrogen regulates snail and slug in the down-regulation of E-cadherin and induces metastatic potential of ovarian cancer cells through estrogen receptor alpha. Mol Endocrinol. 2008;22(9):2085–98.
56. Banerjee A, Cai S, Li N, Lavudi K, Pal T, Bai X, et al. Abstract 3094: estrogen receptor β agonist depletes ovarian cancer stem cells via repressing epithelial to mesenchymal transition. Cancer Res. 2021;81(13_Supplement):3094.

Publisher's Note