PTPN2 Expression in Cystadenocarcinoma Ovary and It's Clinical Value Based on TCGA Database

Qian Li
Zhengzhou University Third Hospital and Henan Province Women and Children's Hospital
Yannan Chen
Zhengzhou University Third Hospital and Henan Province Women and Children's Hospital
Li Yang
Zhengzhou University Third Hospital and Henan Province Women and Children's Hospital
Yaguang Fan
Zhengzhou University
Feiyan Li
Zhengzhou University Third Hospital and Henan Province Women and Children's Hospital
Feng Zhang
Zhengzhou University Third Hospital and Henan Province Women and Children's Hospital
Chenchen Ren (1241149737@qq.com)
Zhengzhou University Third Hospital and Henan Province Women and Children's Hospital
https://orcid.org/0000-0001-8530-7703
Jiaxi Liu
Zhengzhou University Third Hospital and Henan Province Women and Children's Hospital

Research article

Keywords: PTPN2, Ovarian serous cystadenocarcinoma, The Cancer Genome Atlas, Gene set enrichment analysis

Posted Date: January 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-136135/v1

Abstract

Ovarian serous cystadenocarcinoma (OV) is a malignant tumor that often has a poor prognosis because of its late detection. The expression of PTPN2 is associated with a variety of tumors, but its effect on OV is not well understood. Therefore, we analyzed the relationship between PTPN2 and the prognosis of OV. Analysis of patients with OV using The Cancer Genome Atlas revealed an association between PTPN2 expression and the prognosis of OV. We established a model of the relationship between these factors by logistic regression, which showed a significant correlation between the tumor grade and decreased expression of PTPN2. Kaplan-Meier survival analysis showed that low PTPN2 expression was associated with poor overall survival. Further analysis of the expression of immune cells in OV using the ssGSEA package revealed a significant correlation between the expression level of PTPN2 in OV and the numbers of mast, gamma delta, helper, and central memory T cells. We also found differences between the phenotypic pathways associated with low PTPN2 expression and pathways of genes and proteins that determine epithelial-mesenchymal transformation. Finally, a network diagram of protein molecular interactions was drawn using the STRING database, which showed that PTPN2 was closely related to the signal converter and transcriptional activator family and Janus kinase family. Thus, PTPN2 shows potential for use as a prognostic biomarker in OV and is associated with immune infiltration.

Introduction

OV is a common gynecological malignancy[1]. Because it is asymptomatic in its early stages, the detection of OV is difficult; as a result, many patients are initially diagnosed with advanced disease (stage III, IV), and thus the prognosis is poor[2; 3]. The unsatisfactory prognosis of OV also reflects the limitations of current treatment strategies. Therefore, studies aimed at identifying an effective prognostic marker are urgently needed, which is a current research hotspot. Although various biomarkers, such as carbohydrate antigen 125 and carbohydrate antigen 199, are currently considered as relevant in clinical practice, these markers are have become controversial[4]. Therefore, new prognostic molecular markers for early diagnosis and new treatment regimens are needed to improve patient survival.

Phosphorylation of tyrosine kinase is an important cell signaling mechanism in tumorigenesis. Protein tyrosine phosphatases regulate phosphorylation by removing phosphorylating groups and returning tyrosine kinases to their original state[5]. PTPN2, also known as TCPTP, is a widely expressed intracellular nontransmembrane phosphatase[6].

In this study, we assessed the prognostic value of PTPN2 expression in human OV by analyzing data available from The Cancer Genome Atlas (TCGA). Additionally, biological pathways related to PTPN2 in OV were investigated by gene set enrichment analysis (GSEA).

Materials And Methods

2.1. Data acquisition

We downloaded the dataset of patients with OV from TCGA7. This dataset included 71 normal tissues and 263 tumor tissues. We then excluded pathologies with missing data such as age and survival time. RNA sequencing data were transformed into transcripts per million reads (TPM) for subsequent analyses. Tumor tissues were divided into two groups according to the expression level of PTPN2.

2.2. Expression analysis by USCS XENA

We downloaded TCGA GTEx and TPM RNASeq data format from UCSC XENA (https://xena.ucsc.edu/). RNAseq data of TCGA and GTEx were processed using Toil software[8]. The Wilcoxon rank sum test was performed to compare the expression of PTPN2 in GTEx and TCGA in tumor samples. PTPN2 expression data from normal samples from GTEx and TCGA were combined, and OV samples from TCGA were included in the comparison. The stage diagram with pathological stage as the variable was analyzed to compare PTPN2 expression in different pathological stages. A boxplot using the disease state (tumor or normal state) as a variable was drawn to calculate the differential expression of PTPN2.

2.3. Survival analysis

A Kaplan-Meier plot using the SurvMiner package was drawn to evaluate the prognostic value of PTPN2 in terms of overall survival (OS) of patients with OV[9]. Gene expression values were divided into high and low expression groups according to the median value. The risk ratios (HR) and logarithmic rank P values of 95% confidence intervals were also calculated.

2.4. Statistical analysis

Data obtained from TCGA were statistically analyzed using R-3.6.3 software. The correlation between PTPN2 expression and clinical data was analyzed by logistic regression, and the influence of other clinical factors on the survival rate was evaluated by multivariate Cox analysis. We also analyzed the expression of PTPN2 in various tumors and analyzed the correlations between PTPN2 expression and levels of 24 immune cells.

2.5. Gene set enrichment analysis

GSEA was conducted to assess the distribution trends of genes of pre-defined gene sets in the gene table to determine their influence on phenotype[10; 11]. GSEA was conducted using the gseGO and gseKEGG functions to identify potential functional associations with differential expression of PTPN2 based on RNA sequences from TCGA database, and htSEQ-counts data were analyzed with the DESeq2 packet. There were 52 differential molecules of |log fold change (FC)|>2 and padj < 0.05. The expression level of PTPN2 was considered as a phenotype, and the enrichment pathways of each phenotype were determined by the regulator P-value < 0.05, false discovery rate (FDR) q-value < 0.25, and standardized enrichment score |NES| >1[12].

2.6. Analysis of immune infiltration of PTPN2 by ssGSEA

Gene set enrichment analysis methods using the GSEA package URL or GSEA package literature in R (3. 6. 3) was conducted to analyze 24 types of immune cells in tumor samples, including neutrophils, mast

cells, eosinophils, macrophages, natural killer (NK) cells, CD56dim NK cells, CD56bright NK cells, dendritic cells (DCs), immature DCs, activated DCs, plasmacytoid DCs, plasma cells, T cells to CD8 + T cells, T helper cells and Th1/Th2 cells, Th17 cells, T follicular helper cells, Tregs, effector memory T cells, central memory CD4 + T cells (TCM), gamma delta T cells, cytotoxic cells, and B cells. Based on the characteristic genes of 24 immune cells in the literature[13], the relative enrichment scores of each cell type were quantified from the gene expression profiles of each tumor sample. The correlations between PTPN2 and levels of immune cells were analyzed by Spearman correlation, and infiltration of immune cells between the high and low-expression groups was analyzed by Wilcoxon rank sum test.

2.7. Protein-protein interaction

STRING (https://string-db.org) is a comprehensive network of protein interactions that enables searches for known protein interactions and prediction of protein interactions. These interactions include both direct physical interactions between proteins and indirect functional correlations between proteins. To further explore the interactions of PTPN2, a PPI network of PTPN2 was determined the STRING (http://string-db.org/) database. For each PPI relationship pair distributed between 1 and 0, the database generates a composite score; a higher total score indicates a more reliable PPI relationship. The commonly used comprehensive scoring threshold is 0.4. In this study, we used an interaction score > 0.4 as the cut-off criterion.

Results

3.1. Patient characteristics and multivariate analysis

In June 2020, we obtained clinical and gene expression data from 376 primary tumors from TCGA (Table). Patients were stratified by age group (55.1% were under 60 years old and 44.9% were over 60 years old) and by tumor status [71 (21.3%) had no tumors and 263 (78.7%) had tumors]. In terms of disease status, one case (0.3%) was stage I, 22 cases (5.9%) were stage II, 293 cases (78.6%) were stage III, and 57 cases (15.3%) were stage IV. Tumor tissues were divided into two groups according to PTPN2 expression levels to examine the effect of PTPN2 expression on the immune microenvironment.

According to the Kaplan-Meier plot (Fig. 1a), the group with high PTPN2 expression was associated with better OS of patients with OV. In addition, PTPN2 expression in tumor tissues was significantly lower than that in normal tissues. In the Cox regression model (Table), variables with $P < 0.1$ in univariate Cox regression were included in multivariate Cox regression. Variables satisfying this threshold included primary therapy outcome ($P < 0.001$), tumor residual ($P < 0.001$), age ($P = 0.017$), race ($P = 0.047$), tumor status ($P < 0.001$), and PTPN2 ($P = 0.017$). Multivariate Cox regression showed that primary therapeutic outcome ($P < 0.001$), age ($P = 0.017$), tumor status ($P < 0.001$), and PTPN2 ($P = 0.036$) were independent prognostic factors of OS ($P < 0.05$). PTPN2 expression was significantly correlated with tumor grade. The expression of PTPN2 decreased with increasing tumor grades.

3.2. Expression difference analysis

TCGA OV samples were divided into high and low PTPN2 expression groups. There were 52 differential molecules of |logFC| >2 and padj < 0.05, among which eight were high-expression genes and 44 were low-expression genes (Fig. 1b). We also constructed a heat map(Fig. 2).

3.3. Relationship between PTPN2 expression and clinicopathological variables

A total of 376 OV samples were analyzed from TCGA (Table), including PTPN2 expression data for all patient characteristics. Univariate analysis using logistic regression showed that PTPN2 was significantly correlated with the FIGO stage (P = 0.001), tumor residual (P = 0.038), and tumor status (P = 0.041). These results suggest that endothelial cells with low PTPN2 expression are more likely to progress to a higher late and distant metastasis than those with high PTPN2 expression. Kaplan-Meier survival analysis also showed that low PTPN2 expression was associated with poor OS (HR = 0.73 (0.56–0. 94); P = 0.017).

We also conducted logistic regression to analyze the relationship between the clinicopathological features of OV and value of PTPN2 TPM. PTPN2 was significantly correlated with the FIGO stage (P = 0.001), tumor residual (P = 0.038), and tumor status (P = 0.041).

We also drew a receiver operating characteristic curve with the false-positive rate on the horizontal axis and true positive rate on the vertical axis (Fig. 3c). We evaluated the diagnostic efficacy of PTPN2 for OV by receiver operating characteristic curve analysis. The area under the curve of PTPN2 in the figure is 0.710, suggesting that PTPN2 is a diagnostic molecule.

A nomogram was also used to draw the prognostic model (Fig. 3b). Tumor status, primary therapeutic outcome, age, histologic grade, tumor residual, and PTPN2 were included in the model shown in Fig. 3b. The C-index of the model was 0.740 (0.720–0.760).

3.4. Relationship between PTPN2 expression and tumor-infiltrating immune cells

To determine the difference in PTPN2 expression between tumor and normal tissues, the expression levels of PTPN2 mRNA in normal tissues and multiple tumor types were analyzed using TGCA and GTEx databases (Fig. 3a). The results showed that PTPN2 was significantly differentially expressed (P < 0.05) in adrenal cortical carcinoma, bladder urothelial carcinoma, cervical squamous carcinoma and adenocarcinoma, cholangiocarcinoma, colon cancer, diffuse large B cell lymphoma, esophageal, squamous cell carcinoma of the head and neck, renal clear cell carcinoma, renal papillary cell carcinoma, acute myeloid leukemia, low-grade glioma of the brain, lung adenocarcinoma, lung squamous carcinoma, OV, pancreatic cancer, prostate cancer, rectal adenocarcinoma, sarcoma, cutaneous melanoma, gastric cancer, testicular cancer, thyroid cancer, thymic carcinoma, endometrial cancer, and uterine sarcoma. Moreover, the expression levels of PTPN2 in normal samples from GTEx and TCGA combined and OV samples of TCGA of ovary serous cystadenocarcinoma were significant (P < 0.05) (Fig. 3d).

Our results showed that PTPN2 expression was significantly low in OV ($P < 0.001$) (pan carcinoma differentially expressed figure) and mast cell ($R = -0.195$, $P < 0.01$) and T gamma delta ($R = -0. 153$, $P = 0.03$) invasion levels and negatively correlated with T helper cells ($R = 0.291$, $P < 0.01$). The TCM ($R = 0.327$, $P < 0.01$) level was significantly positively related to PTPN2 expression (Fig. 4a and b).

3.5. Differential expression of PTPN2

GSEA of differences between low and high PTPN2 expression data sets was performed to identify key signaling pathways associated with PTPN2. The results showed significant enrichment of several Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes terms (FDR < 0.05, NOM $P < 0.05$)[14]. GO analysis showed that genes related to PTPN2 mostly act on mRNA, including in mRNA 5'-splice site recognition and mRNA splice site selection. A histogram of the GO enrichment analysis results is shown in Fig. 5a.

Based on the OV expression matrix from TCGA data created using the cluster Profiler package [PMID:22455463][12], GSEA was carried out on both low-expression and high-expression PTPN2 groups, and H. lati. v7. symbol. GMT [Hallmarks] and C5. all. v7[14; 15]. GMT (GO) was selected as a reference gene collection in MSigDBCollections. An FDR < 0.25 and Padjust < 0.05 were used to define significant enrichment. PTPN2 and genes involved in election-mesenchymal transition and epithelial-mesenchymal transition were significantly enriched (Fig. 5b).

3.6. Protein-protein interaction

By using the STRING search tool (Fig. 5c), we found that non-receptor type tyrosine-specific phosphatase dephosphorylates receptor protein tyrosine kinases including CSF1R, INSR, PDGFR, and EGFR. PTPN2 also dephosphorylates non-receptor protein tyrosine kinases including Src, JAK1, JAK2, and JAK3 family kinases, and STAT1, STAT3, and STAT6 in either the cytoplasm or the nucleus. It also negatively regulates numerous signaling pathways and biological processes such as inflammatory response, hematopoiesis, glucose homeostasis, and cell proliferation and differentiation[16].

Discussion

Here, we showed that the prognosis of OV was correlated with the expression level of PTPN2, with high PTPN2 expression related to a positive prognosis. In addition, our study showed that PTPN2 expression was associated with different sets of immune markers and immune infiltration. Therefore, PTPN2 may affect tumor immunity. We also found that PTPN2 expression differed between OV tumor tissue and normal tissue and was correlated with tumor grade. Multivariate analysis showed that PTPN2 expression was an independent prognostic factor affecting the prognosis of patients with OV. Therefore, PTPN2 is a promising cancer tumor biomarker.

PTPN2 is an enzyme that removes phosphate groups from its substrate. Phosphatase plays an important role in maintaining cell homeostasis. The subcellular localization of PTPN2 is variable, and under certain

conditions, the protein shuttles between the nucleus and cytoplasm. PTPN2 is widely expressed in adult cells and plays an important role in regulating cell life activities. Therefore, abnormal PTPN2 protein can cause specific diseases. PTPN2 is closely related to the immune system and affects most cells in the immune system. PTPN2 is closely related to JAK1, JAK3, STAT1, and STAT3, as observed in the PPI graph. Expression of PTPN2 was decreased in some breast cancers, and re-transfer of PTPN2 inhibited the growth of breast cancer cells. This regulation may occur via inhibition of the transforming growth factor (TGF)-β-transforming signal pathway. Therefore, it is important to study the relationship between the TGF-β signal pathway and PTPN2 from a physiological and pathological perspective.

The TGF-β signal pathway is a signal transmission process mediated by transforming growth factor.[17] This pathway plays a key role in the growth, development, and differentiation of cells and tissues, as well as an important regulatory role in cell proliferation, interstitial generation, differentiation, apoptosis, embryonic development, organ formation, immune function, inflammatory response, wound repair, etc. TGF-β expression and signal transduction disorders are associated with the development of many diseases, such as cancer, fibrosis, and hereditary hemorrhagic telangiectasia, familial primary pulmonary hypertension, and many other genetic diseases. TGF-β is associated with the occurrence, progression, and metastasis of tumors.

The classic TGF-β signal pathway contains TGF-β receptors I and II and its downstream proteins Smad, Co-Smad, and Smad4, with Smad4 forming complexes with phosphorylated R-Smad in the nucleus to induce transcription. Thus, Smad4, although not directly regulated by TGF-β activity, is a key factor in the entire TGF-β signaling pathway[18].

In some cancers, the TGF-β signaling pathway is inhibited by continuously activated kinase NPM-ALK[19], which continuously phosphorylates Smad4 and in turn prevents Smad4 from forming a complex with phosphorylated R-Smad in the nucleus for transcription.

PTPN2 has an obvious dephosphorylation effect on Smad4 phosphorylated by NPM-ALK, and the expression of NPM-ALK can dimerize and autophosphorylation is activated, thus showing sustained kinase activity[20]. Studies have shown that PTPN2 also dephosphorylates NPM-ALK, thus reducing the activity of NPM-ALK. Therefore, PTPN2 plays a role by reducing the activity of NPM-ALK and thus affecting Smad4.

The TGF-β signaling pathway also crosstalks with many other pathways, including the JAK/STAT pathway[21]. PTPN2 can mediate the dephosphorylation of pSTAT3(705) and negatively regulate the transcriptional activity of STAT3, whereas STAT3 can bind to Smad3 to inhibit the TGF-β signaling pathway. Therefore, high expression of PTPN2 inhibits STAT3 activity, thus preventing the binding of STAT3 to Smad3 and then restores the TGF-β signaling pathway. In summary, the effect of PTPN2 on the TGF-β signaling pathway through other pathways should be further studied.

Here, we showed that the prognosis of OV was correlated with the expression level of PTPN2, with high PTPN2 expression associated with a positive prognosis. In addition, our study showed that PTPN2 expression was associated with different sets of immune markers and immune infiltration. Thus, PTPN2 may affect tumor immunity. Further, the expression of PTPN2 differed between OV tumor tissue and normal tissue and was correlated with tumor grade. Multivariate analysis showed that PTPN2 expression was an independent prognostic factor affecting the prognosis of patients with OV. Therefore, PTPN2 is a promising cancer tumor biomarker.

Smad4, a tumor suppressor gene, contains mutations or deletions in many cancer cells[22; 23]. Our study showed that PTPN2 can restore the activity of the TGF-β signaling pathway, which is of high clinical value, and a safe, reliable, and effective PTPN2 agonist may be applicable in the clinical treatment of OV

Declarations

Acknowledgements

Not applicable.

Funding

This work was supported by the Medical science and Technology research project of Henan Province (NO. LHGJ20190358).

Authors' contributions

QL and CCR conceived and designed the study. YNC and LY performed the statistical analysis. FYL,FZ,YGF and JXL were involved in the writing of the manuscript and in the interpretation of the results. YGF was involved in the drawing.All authors read and approved the fial manuscript.

Ethics approval and consent to participate

The present study was approved by The Medical Ethics Committees of The Third Affiliated Hospital of Zhengzhou University

Patient consent for publication

Not applicable

Availability of data and materials

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

References

[1] Dhillon, P.K., Yeole, B.B., Dikshit, R., Kurkure, A.P., and Bray, F. (2011). Trends in breast, ovarian and cervical cancer incidence in Mumbai, India over a 30-year period, 1976-2005: an age-period-cohort analysis. Br J Cancer 105, 723-730.

[2] Dong, X., Men, X., Zhang, W., and Lei, P. (2014). Advances in tumor markers of ovarian cancer for early diagnosis. Indian J Cancer 51 Suppl 3, e72-76.

[3] Colombo, N., Peiretti, M., Garbi, A., Carinelli, S., Marini, C., Sessa, C., and Group, E.G.W. (2012). Non-epithelial ovarian cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 23 Suppl 7, vii20-26.

[4] Xie, Y.J., Chen, M., and Chen, S.J. (2014). Prediction of pregnancy outcomes with combined ultrasound scanning of yolk sacs and serum CA125 determinations in early threatened abortion. Clin Exp Obstet Gynecol 41, 186-189.

[5] Bussieres-Marmen, S., Hutchins, A.P., Schirbel, A., Rebert, N., Tiganis, T., Fiocchi, C., Miranda-Saavedra, D., and Tremblay, M.L. (2014). Characterization of PTPN2 and its use as a biomarker. Methods 65, 239-246.

[6] Veenstra, C., Karlsson, E., Mirwani, S.M., Nordenskjold, B., Fornander, T., Perez-Tenorio, G., and Stal, O. (2019). The effects of PTPN2 loss on cell signalling and clinical outcome in relation to breast cancer subtype. J Cancer Res Clin Oncol 145, 1845-1856.

[7] Blum, A., Wang, P., and Zenklusen, J.C. (2018). SnapShot: TCGA-Analyzed Tumors. Cell 173, 530.

[8] Vivian, J., Rao, A.A., Nothaft, F.A., Ketchum, C., Armstrong, J., Novak, A., Pfeil, J., Narkizian, J., Deran, A.D., Musselman-Brown, A., et al. (2017). Toil enables reproducible, open source, big biomedical data analyses. Nat Biotechnol 35, 314-316.

[9] Liu, J., Lichtenberg, T., Hoadley, K.A., Poisson, L.M., Lazar, A.J., Cherniack, A.D., Kovatich, A.J., Benz, C.C., Levine, D.A., Lee, A.V., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell 173, 400-416 e411.

[10] Subramanian, A., Kuehn, H., Gould, J., Tamayo, P., and Mesirov, J.P. (2007). GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics 23, 3251-3253.

[11] Tu, Z., Xiong, J., Xiao, R., Shao, L., Yang, X., Zhou, L., Yuan, W., Wang, M., Yin, Q., Wu, Y., et al. (2019). Loss of miR-146b-5p promotes T cell acute lymphoblastic leukemia migration and invasion via the IL-17A pathway. J Cell Biochem 120, 5936-5948.

[12] Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545-15550.

[13] Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A.C., Angell, H., Fredriksen, T., Lafontaine, L., Berger, A., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity 39, 782-795.

[14] Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16, 284-287.

[15] Kim, S.Y., and Volsky, D.J. (2005). PAGE: parametric analysis of gene set enrichment. BMC Bioinformatics 6, 144.

[16] Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47, D607-D613.

[17] Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell 113, 685-700.

[18] Xu, L. (2006). Regulation of Smad activities. Biochim Biophys Acta 1759, 503-513.

[19] Hallberg, B., and Palmer, R.H. (2016). The role of the ALK receptor in cancer biology. Ann Oncol 27 Suppl 3, iii4-iii15.

[20] Drexler, H.G., Gignac, S.M., von Wasielewski, R., Werner, M., and Dirks, W.G. (2000). Pathobiology of NPM-ALK and variant fusion genes in anaplastic large cell lymphoma and other lymphomas. Leukemia 14, 1533-1559.

[21] Su, F., Ren, F., Rong, Y., Wang, Y., Geng, Y., Wang, Y., Feng, M., Ju, Y., Li, Y., Zhao, Z.J., et al. (2012). Protein tyrosine phosphatase Meg2 dephosphorylates signal transducer and activator of transcription 3 and suppresses tumor growth in breast cancer. Breast Cancer Res 14, R38.

[22] Hahn, S.A., Schutte, M., Hoque, A.T., Moskaluk, C.A., da Costa, L.T., Rozenblum, E., Weinstein, C.L., Fischer, A., Yeo, C.J., Hruban, R.H., et al. (1996). DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 271, 350-353.

[23] Miyaki, M., and Kuroki, T. (2003). Role of Smad4 (DPC4) inactivation in human cancer. Biochem Biophys Res Commun 306, 799-804.

Tables

Table ⍰⍰Composition of clinical variables and corresponding proportions of OV in TCGA data.

Characteristic	Level	Overall
n		376
FIGO stage (%)	Stage I	1 (0.3%)
	Stage II	22 (5.9%)
	Stage III	293 (78.6%)
	Stage IV	57 (15.3%)
Histologic grade (%)	G1	1 (0.3%)
	G2	42 (11.5%)
	G3	322 (88.0%)
	G4	1 (0.3%)
Primary therapy outcome (%)	CR	213 (69.8%)
	PD	27 (8.9%)
	PR	43 (14.1%)
	SD	22 (7.2%)
Race (%)	Asian	11 (3.0%)
	Black or African American	25 (6.9%)
	White	326 (90.1%)
Anatomic neoplasm subdivision (%)	Bilateral	253 (71.5%)
	Unilateral	101 (28.5%)
Venous invasion (%)	No	40 (38.8%)
	Yes	63 (61.2%)
Tumor residual (%)	NRD	66 (19.8%)
	RD	267 (80.2%)
Tumor status (%)	Tumor free	71 (21.3%)
	With tumor	263 (78.7%)
Lymphatic invasion (%)	No	48 (32.4%)
	Yes	100 (67.6%)
TP53 status (%)	Mut	248 (90.5%)
	WT	26 (9.5%)

Characteristic	Level	Overall
Age (median [IQR])		59. 00[51. 00,68. 00]

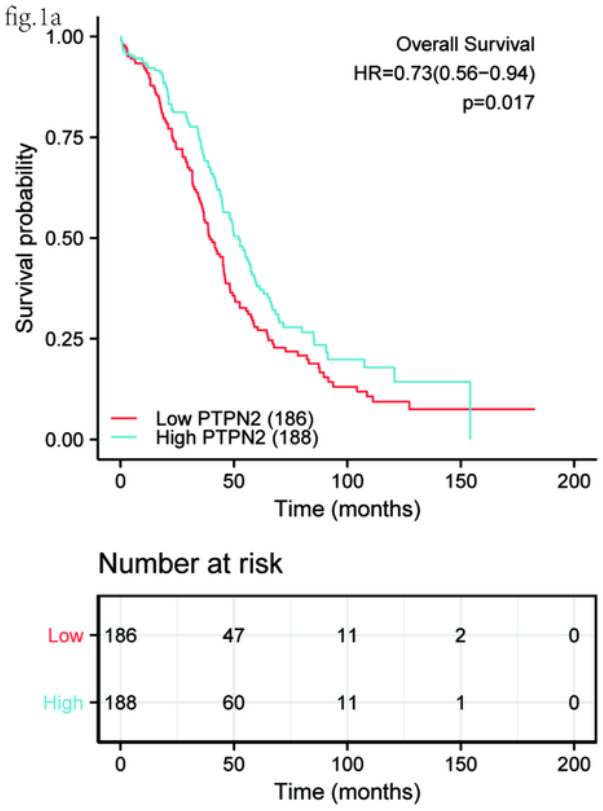
Table Univariate/multivariate Cox regression analysis results predicting OS. In the Cox regression model, variables with p < 0.1 in univariate Cox regression were included in multivariate Cox regression. Variables satisfying this threshold included primary therapy outcome (P < 0. 001), tumor residual (P < 0.001), age (P = 0.017), race (P = 0.047), tumor status (P < 0.001), and PTPN2 (P = 0.017). Multivariate Cox regression showed that primary therapeutic outcome (P < 0.001), age (P = 0.017), tumor status (P < 0.001), and PTPN2 (P = 0.036) were independent prognostic factors in OS (P <0. 05).

Characteristic	Total (N)	Univariate analysis		Multivariate analysis	
		HR (95% CI)	P-value	HR (95% CI)	P-value
FIGO stage (Stage III & Stage IV vs. Stage I & Stage II)	371	2.085 (0. 925–4. 699)	0. 076	1. 820 (0.431–7.687)	0.415
Histologic grade (G3&G4 vs. G1&G2)	364	1.194 (0.797–1. 789)	0. 389		
Primary therapy outcome (CR vs.PD&SD&PR)	304	0.234 (0.169–0.324)	<0.001	0. 340 (0.237–0.488)	<0.001
Venous invasion (Yes vs. No)	103	0.905 (0.487–1.683)	0.753		
Tumor residual (RD vs. NRD)	332	2.302 (1.479-3.583)	<0.001	1. 166 (0.678–2.005)	0.579
Age (>60 vs. <=60)	374	1.373 (1.059–1.780)	0.017	1. 496 (1.075–2.081)	0.017
Tumor status (Tumor free vs. with tumor)	333	0.092 (0.041–0.209)	<0.001	0. 057 (0. 014–0.233)	<0. 001
Lymphatic invasion (Yes vs. No)	147	1.422 (0.839–2.411)	0.191		
TP53 status (Mut vs. WT)	273	0.692 (0.423–1.132)	0.143		
PTPN2 (High vs. Low)	374	0.727 (0.560–0.944)	0.017	0. 700 (0.501–0.977)	0. 036

Table □. The relationship between the clinicopathological features of serous cystadenocarcinoma and the value of PTPN2 TPM was analyzed using the logistic regression method. PTPN2 was significantly correlated with FIGO stage (P = 0.001), tumor residual (P = 0.038), and tumor status (P = 0. 041)

Characteristic	N	Odds Ratio (OR)	P value
FIGO stage (Stage III & Stage IV vs. Stage I & Stage II)	373	0. 93(0.88–0. 97)	0. 001
Histologic grade (G3&G4 vs. G1&G2)	366	1. 02(0. 97–1. 08)	0. 423
Primary therapy outcome (CR vs. PD&SD&PR)	305	1. 02(0. 99–1. 06)	0. 293
Anatomic neoplasm subdivision (Bilateral vs. Unilateral)	354	0. 97(0. 94–1. 00)	0. 052
Venous invasion (Yes vs. No)	103	1. 01(0. 96–1. 06)	0. 714
Tumor residual (RD vs. NRD)	333	0. 96(0. 92–1. 00)	0. 038
Tumor status (Tumor free vs. With tumor)	334	1.04(1.00–1. 08)	0. 041
Lymphatic invasion (Yes vs. No)	148	0. 99(0.95–1.03)	0. 488
TP53 status (Mut vs. WT)	274	1. 01(0.96–1.08)	0. 629

Figures



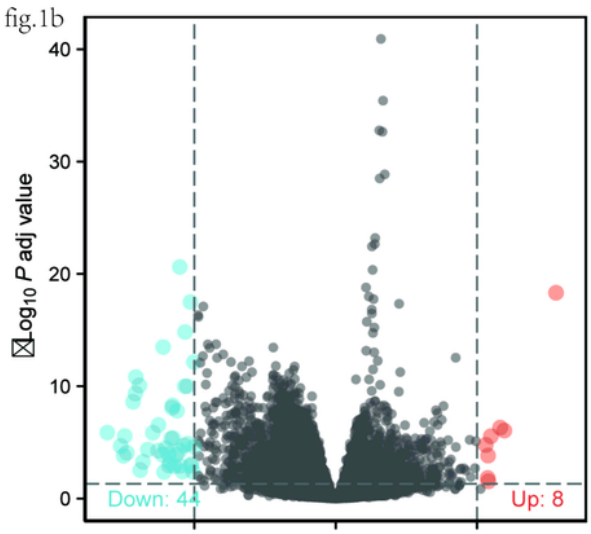


Figure 1

a. Kaplan-Meier plot using SurvMiner package was used to evaluate the prognostic value of PTPN2 in predicting the overall survival of patients with OV. b. In difference analysis, there were 52 differentially expressd molecules with |logFC| > 2 and Padj < 0. 05.

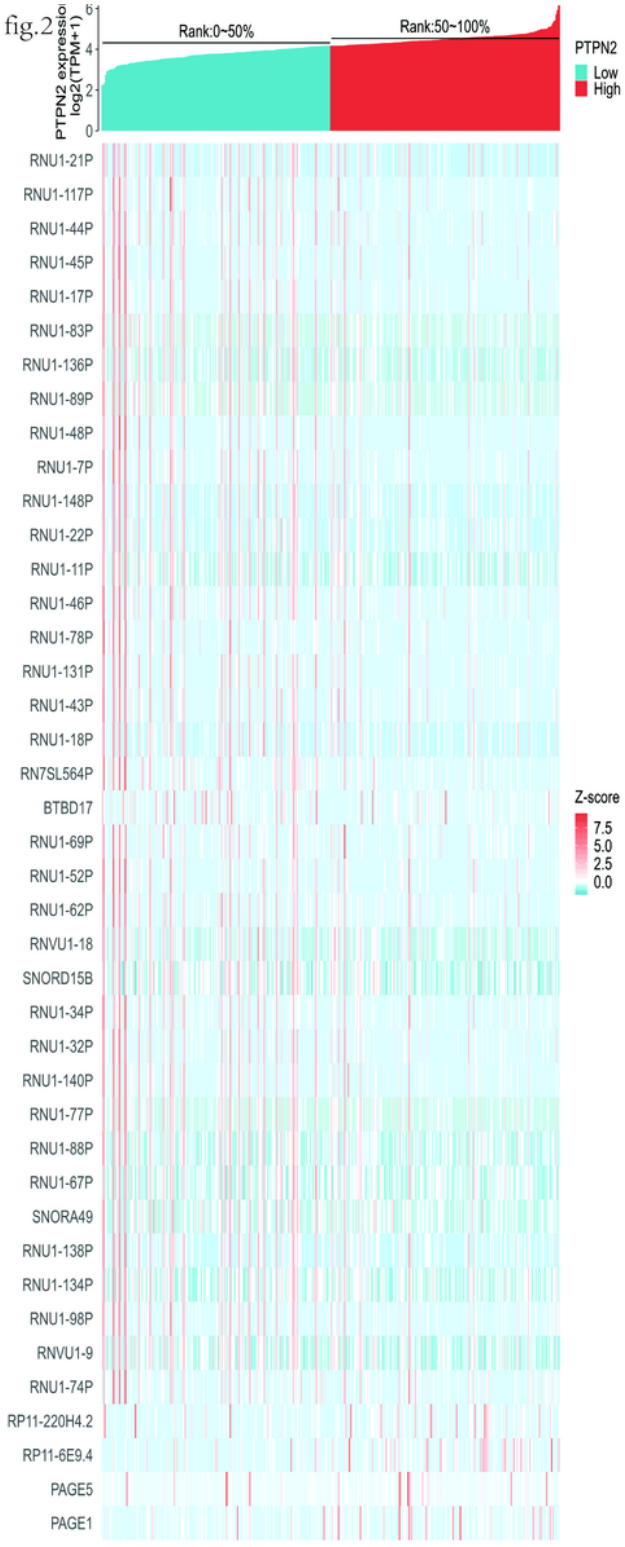


Figure 2

In difference analysis, in TCGA patients with OV, PTPN2 was divided into high and low expression groups, showing the co-expression difference of genes.

fig.3a

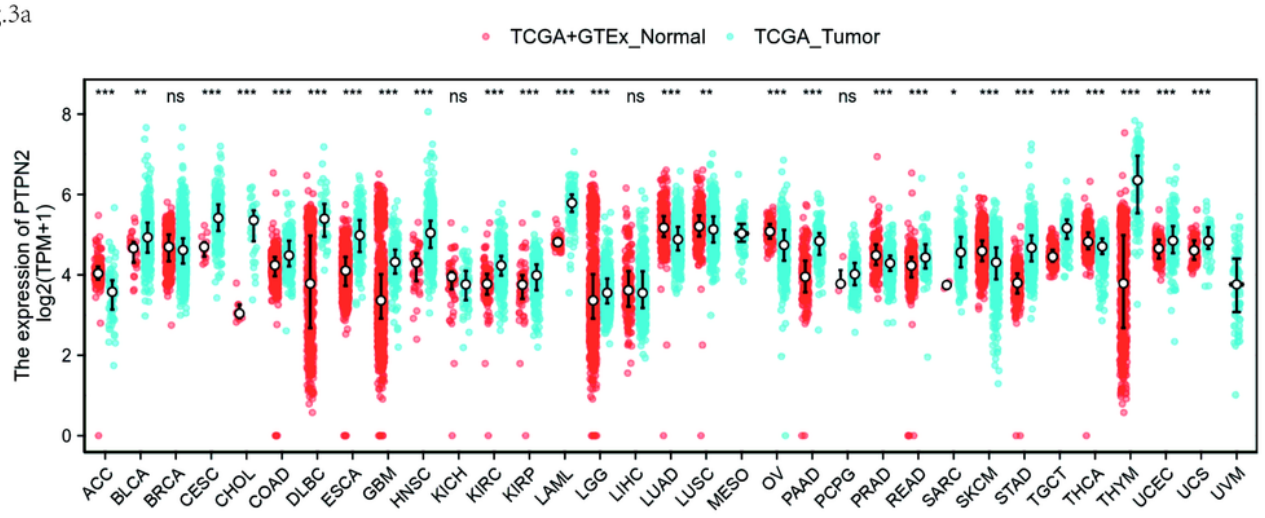


fig.3b

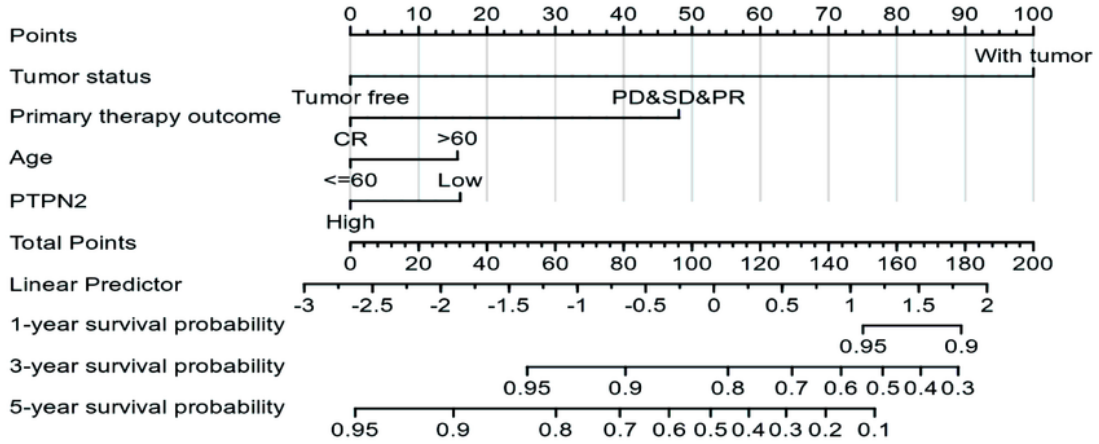


fig.3c

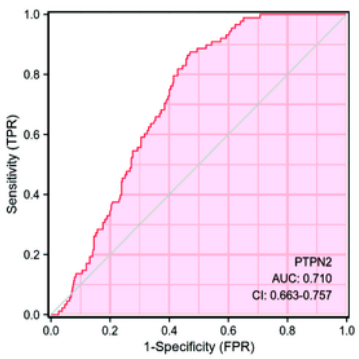


fig.3d

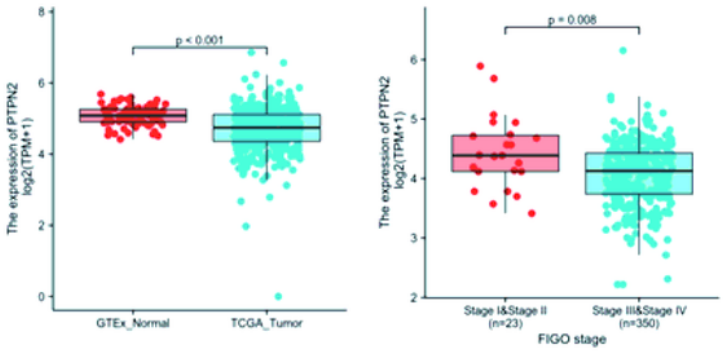


Figure 3

a. Wilcoxon Rank SUM test was used to compare the expression of PTPN2 in GTEx combined with TCGA in tumor samples. b. In the nomogram, the c-index value is generally 0.5–1, with 0.50–0.70 indicating low accuracy, 0.71–0.90 indicating moderate accuracy, >0.90 indicating high accuracy. c. Receiver operating characteristic curve, with an area under the curve of 0.5–1. An area under the curve close to 1 indicates a better diagnostic effect. The area under the curve of 0.5–0.7 has low accuracy, 0.7–0.9 has some

accuracy, and >0.9 is high accuracy. d. Wilcoxon Rank sum test was used to compare PTPN2 expression with the tumor status of patients with OV in TCGA serous cystadenocarcinoma of the ovary.

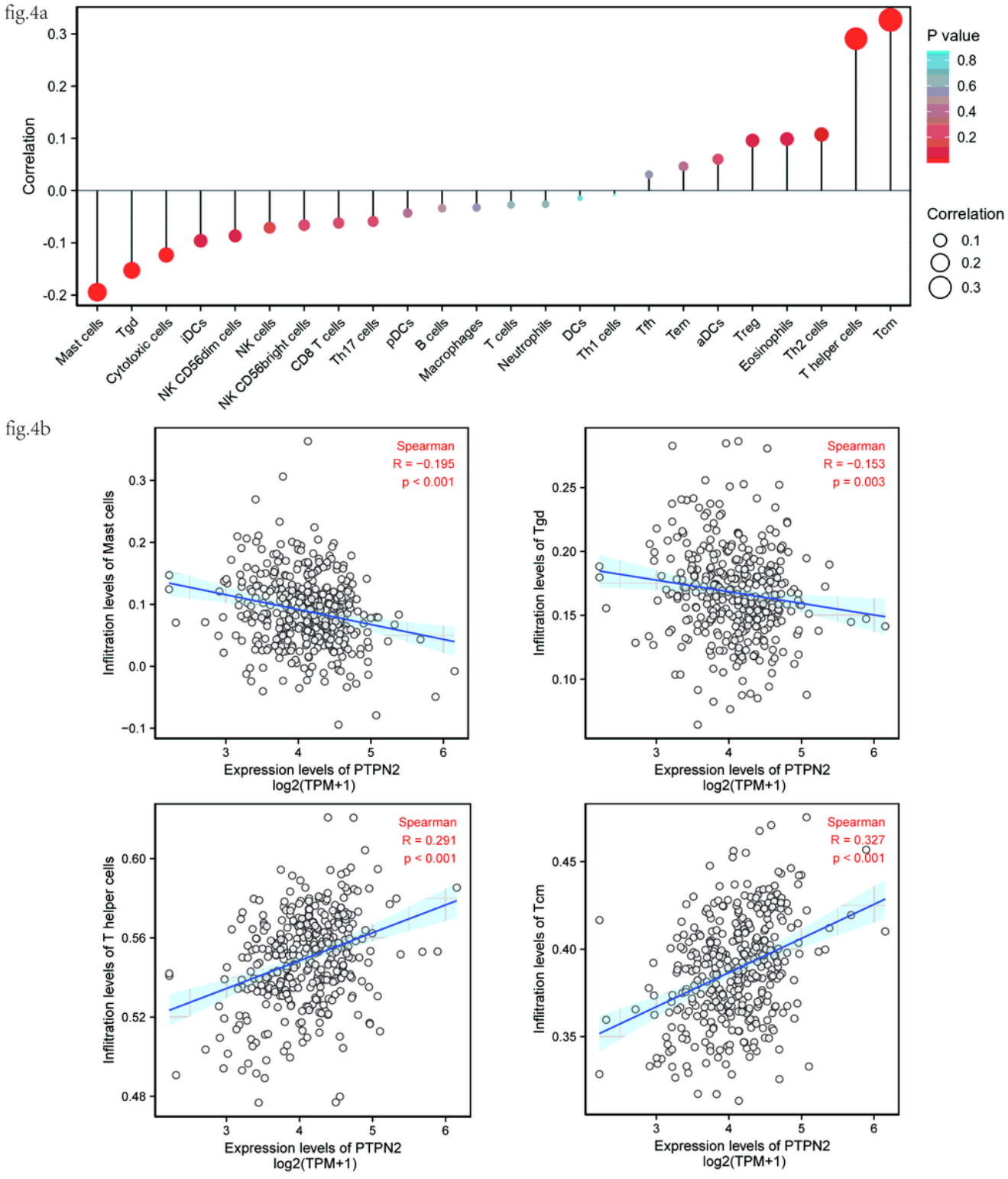


Figure 4

a. Correlation between PTPN2 expression in 24 kind of immune cells. b. Spearman correlation method was used to analyze the correlation between PTPN2 expression and Tcm, Th2 cells, Tgd, and mast cells.

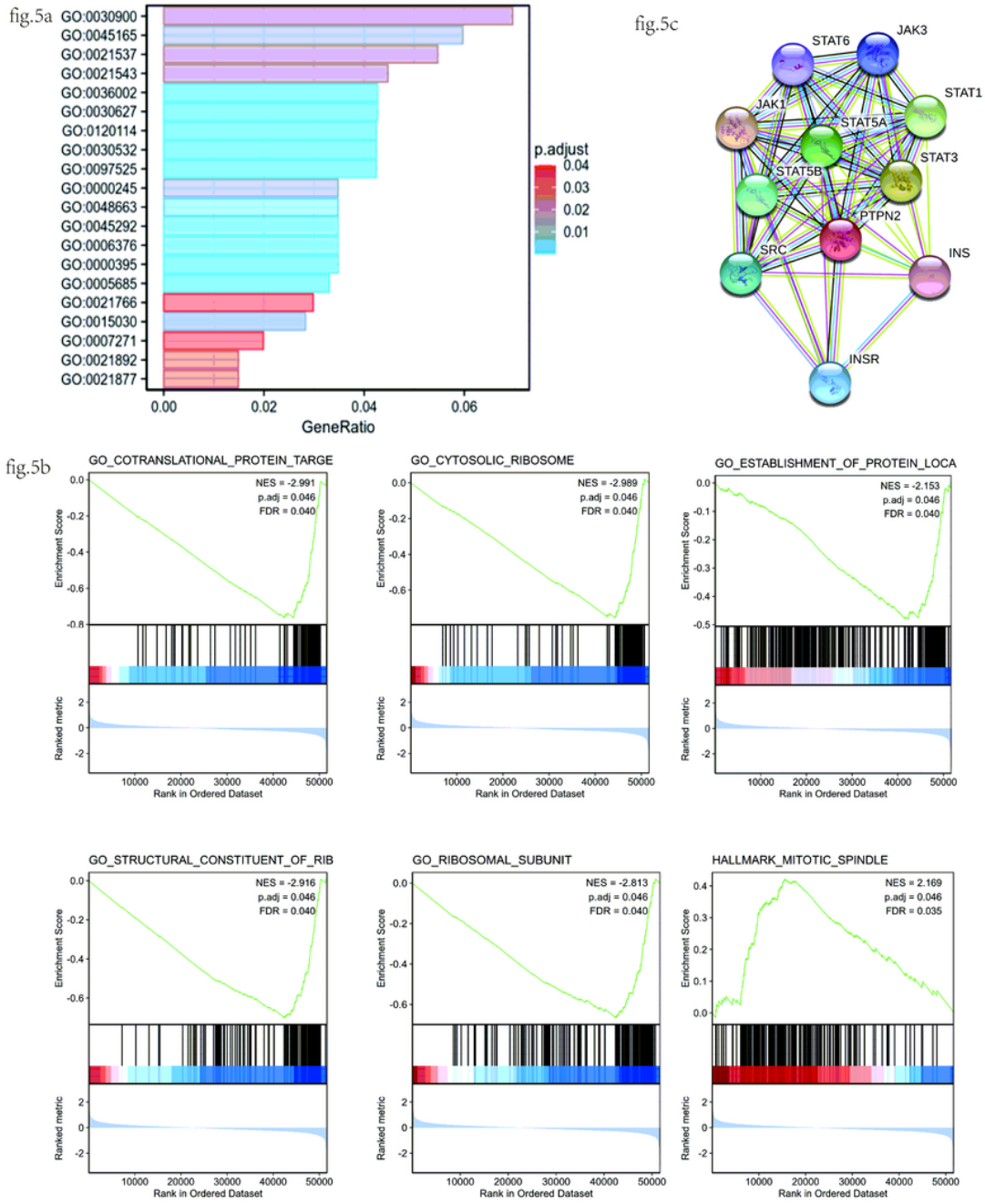


Figure 5

a. Gene ontology (GO) enrichment analysis of the input gene list, including biological process (BP), cellular components (CC), and molecular function (MF). b. Typical results of single gene set GSEA visualization. c. Protein-protein interaction from the STRING database.