Sep 25, 2024

# Pan-Cancer Analysis of the Prognostic and Immunological Roles of SHP-1/*ptpn6*

DOI
**dx.doi.org/10.17504/protocols.io.rm7vzjdpxlx1/v1**

Ping Cui[1,2], Jie Lian[1,2], Yang Liu[2,3], Dongsheng Zhang[1,2], Yao Lin[1,2], Lili Lu[1], Li Ye[1,2], Hui Chen[2,3], Sanqi An[1,2], Jiegang Huang[2,4,5], Hao Liang[1,2]

[1]Life Science Institute, Guangxi Medical University, Nanning, China;
[2]Guangxi Key Laboratory of AIDS Prevention and Treatment, Guangxi Medical University, Nanning, China;
[3]Geriatrics Digestion Department of Internal Medicine, The First Affiliated Hospital of Guangxi Medical University, Nanning, China;
[4]School of Public Health, Guangxi Medical University, Nanning, China;
[5]Guangxi Colleges and Universities Key Laboratory of Prevention and Control of Highly Prevalent Diseases, Guangxi Medical University, Nanning, 530021, China

Ping Cui: First authorship;
Jie Lian: First authorship;
Sanqi An: Correspondence;
Jiegang Huang: Correspondence;
Hao Liang: Correspondence

PTPN6

Ping Cui
GXMU

OPEN ACCESS

**DOI: dx.doi.org/10.17504/protocols.io.rm7vzjdpxlx1/v1**

**Protocol Citation:** Ping Cui, Jie Lian, Yang Liu, Dongsheng Zhang, Yao Lin, Lili Lu, Li Ye, Hui Chen, Sanqi An, Jiegang Huang, Hao Liang 2024. Pan-Cancer Analysis of the Prognostic and Immunological Roles of SHP-1/ptpn6. **protocols.io** **https://dx.doi.org/10.17504/protocols.io.rm7vzjdpxlx1/v1**

**Manuscript citation:**

**Protocol status:** Working
**We use this protocol and it's working**

**Created:** September 23, 2024

**Last Modified:** September 25, 2024

**Protocol Integer ID:** 108198

**Keywords:** ptpn6, pan-cancer, immune infiltration, prognosis

**Funders Acknowledgement:**
**the National Natural Science Foundation of China**
**Grant ID: NSFC 82002134**
**One Thousand of Young and Middle-aged Key Teachers Training Program in Guangxi Colleges and Universities (To Cui Ping)**
**Grant ID: DC2300017000**
**National Natural Science Foundation of Guangxi**
**Grant ID: 2023GXNSFDA026036**
**2022 Innovation and Entrepreneurship Training Program of Guangxi Medical University**
**Grant ID: 202210598034**

## Abstract

In this paper, we present a comprehensive analysis of *ptpn6* across various cancers using multiple online databases, such as TIMER, GEPIA2, and cBioPortal, for differential expression, survival prognosis, immune infiltration, genetic alterations, epigenetic alterations, and functional state evaluations. We expect to reveal significant correlations between *ptpn6* expression and clinical outcomes, as well as its association with immune cell infiltration and biological pathways, to provide insight into presenting a potential prognosis biomarker and immunotherapy target. Additionally, this analysis aims to highlight the heterogeneity of *ptpn6* across different cancer types to help understand its role in tumorigenesis and development.

## Protocal

1 **Differential expression analysis**

The Tumor Immune Estimation Resource (TIMER) 2.0 (http://timer.cistrome.org/) is an online website used to investigate the pan-cancer analysis of gene expression or correlation, and immune infiltration [1]. Difference of *ptpn6* expression between tumors and adjacent normal tissues can be obtained through the “Gene_DE” module of TIMER2. The results were validated using Gene Expression Profiling Interactive Analysis 2 (GEPIA2) database (http://gepia2.cancer-pku.cn/). The expression of p*tpn6* in different pathological stages of cancers was also obtained by GEPIA2 [2].

2 **Survival prognosis analysis**

The heatmaps of overall survival (OS) and disease-free survival (DFS) of *ptpn6* in all TCGA tumors were acquired through GEPIA2. The corresponding survival plots with their 95% confidence interval, *p* value and hazard ratio (HR) can be obtained by the Kaplan-Meier plotter database (https://kmplot.com/analysis/)[3]. To evaluate the expression of *ptpn6* in predicting the prognosis of cancer patients, ROC analysis was conducted using the pROC package in R language (version 4.2.2).

3 **Immune infiltration analysis**

The correlation between *ptpn6* expression and immune infiltration in pan-cancer was investigated using TIMER (https://cistrome.shinyapps.io/timer/) [4]. The tumor purity, B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils and dendritic cells were selected. The results were visualized as scatter plots. Heatmaps and scatter plots of the correlation between *ptpn6* expression and cancer associated fibroblasts (CAFs) were generated through TIMER2 [5].

4 **Enrichment analysis**

Enrichment analysis helps to discover novel biological functions, genotype-phenotype relationships and disease mechanisms. Experimentally determined SHP-1-binding proteins can be obtained through the STRING database (https://string-db.org/), by the following parameters: minimum required interaction score[low confidence (0.150)], meaning of network edges (evidence), maximum number of interactors to show (no more than 50 interactors in 1st shell), and active interaction sources (experiments)[6]. The top 100 *ptpn6*-related genes, were obtained by GEPIA2 and the top five genes were selected to draw the correlation scatter plot with *ptpn6*. The heat map between the selected genes and different types of tumors can be acquired through TIMER2. In addition, the intersection analysis of SHP-1-binding proteins and *ptpn6*-related genes was conducted using Jvenn (https://bioinformatics.psb.ugent.be/webtools/Venn/)[7]. These two sets of data were also combined for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis[8-10] and Gene Ontology (GO) analysis. The functional annotation data were obtained through the Database for Annotation, Visualization, and Integrated Discovery (DAVID,

https://david.ncifcrf.gov/) and enriched pathways were visualized via bioinformatic (https://www.bioinformatics.com.cn/).

5 **Relevance of *ptpn6* across 14 functional states in distinct cancers**

Single-cell RNA sequencing (scRNA-seq) can help researchers understand the functional specificity of cancer cells. CancerSEA (http://biocc.hrbmu.edu.cn/CancerSEA/) is a database for functional states of cancer cells at single-cell level, including angiogenesis, apoptosis, cell cycle, differentiation, DNA damage, DNA repair, EMT, hypoxia, inflammation, invasion, metastasis, proliferation, quiescence, and stemness [11]. The functional state of *ptpn6* in multiple cancers was explored using CancerSEA. Correlations between *ptpn6* expression and functional states in different single-cell datasets were filtered by a correlation strength >0.3 and the *p* value < 0.05.

6 **Genetic alteration analysis**

The cBioPortal (http://www.cbioportal.org), a comprehensive database of cancer genomics datasets [12], is applied to the analysis of *ptpn6* genetic alteration. We explored the copy number alteration (CNA) and mutation status of *ptpn6* across all TCGA tumors using cBioPortal. The results of the alteration frequency, mutation type and CNA in various cancers were derived from the 'Cancer Types Summary' module. The OS, DFS, progression free survival (PFS), and disease free survival (DSS) of patients with *ptpn6* genetic altered were also obtained from cBioPortal.

7 **Analysis of the methylation and phosphorylation of *ptpn6***

UALCAN performed protein expression analysis from the clinical proteomic tumor analysis consortium (CPTAC) dataset and the International Cancer Proteogenome Consortium (ICPC) datasets[13]. The methylation and phosphorylation levels of *ptpn6* between different cancers and normal tissues was investigated by UALCAN database (http://ualcan.path.uab.edu/analysis.html).

8 **Immunohistochemistry (IHC) Staining**

Human Protein Atlas (HPA) (https://www.proteinatlas.org/) is a database of proteins in human organs, tissues and cells based on multiple omics approaches[14,15]. To analyze the differential expression of *ptpn6* at the protein level, the expression of *ptpn6* proteins (SHP-1) in tumor tissues and their corresponding normal tissues was downloaded from HPA and analyzed. Furthermore, the IHC images of some typical immune markers were also acquired from HPA.

9 **Statistical analysis**

Alterations in *ptpn6* expression levels in cancer and normal tissues were estimated using two sets of t-tests. The Kaplan-Meier curve and Cox regression model were used for survival analyses in this study. The Hazard Ratio was calculated by the Cox regression model. The correlation expression analysis between the two variables was analyzed using Spearman's or Pearson's test. *P*-value< 0.05 was considered statistically significant.

## Protocol references

1 Viljević, N., Scibior-Bentkowska, D., Brentnall, A. R., Cuzick, J. & Lorincz, A. T. Credentialing of DNA methylation assays for human genes as diagnostic biomarkers of cervical intraepithelial neoplasia in high-risk HPV positive women. *Gynecologic oncology* **132**, 709-714, doi:10.1016/j.ygyno.2014.02.001 (2014).
2 Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic acids research* **47**, W556-w560, doi:10.1093/nar/gkz430 (2019).
3 Hou, G. X., Liu, P., Yang, J. & Wen, S. Mining expression and prognosis of topoisomerase isoforms in non-small-cell lung cancer by using Oncomine and Kaplan-Meier plotter. *PloS one* **12**, e0174515, doi:10.1371/journal.pone.0174515 (2017).
4 Peng, L. *et al.* A Pan-Cancer Analysis of SMARCA4 Alterations in Human Cancers. *Frontiers in immunology* **12**, 762598, doi:10.3389/fimmu.2021.762598 (2021).
5 Li, T. *et al.* TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic acids research* **48**, W509-w514, doi:10.1093/nar/gkaa407 (2020).
6 Szklarczyk, D. *et al.* The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research* **49**, D605-d612, doi:10.1093/nar/gkaa1074 (2021).
7 Bardou, P., Mariette, J., Escudié, F., Djemiel, C. & Klopp, C. jvenn: an interactive Venn diagram viewer. *BMC bioinformatics* **15**, 293, doi:10.1186/1471-2105-15-293 (2014).
8 Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein science : a publication of the Protein Society* **28**, 1947-1951, doi:10.1002/pro.3715 (2019).
9 Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic acids research* **51**, D587-d592, doi:10.1093/nar/gkac963 (2023).
10 Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**, 27-30, doi:10.1093/nar/28.1.27 (2000).
11 Yuan, H. *et al.* CancerSEA: a cancer single-cell state atlas. *Nucleic acids research* **47**, D900-d908, doi:10.1093/nar/gky939 (2019).
12 Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery* **2**, 401-404, doi:10.1158/2159-8290.Cd-12-0095 (2012).
13 Chandrashekar, D. S. *et al.* UALCAN: An update to the integrated cancer data analysis platform. *Neoplasia (New York, N.Y.)* **25**, 18-27, doi:10.1016/j.neo.2022.01.001 (2022).
14 Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419, doi:10.1126/science.1260419 (2015).
15 Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science* **357**, doi:10