

SPOP基因与前列腺癌相关性的生物信息学分析

于佳¹ 张景楠² 马均柏¹ 刘欢¹

1. 宁夏医科大学基础医学院 宁夏银川 750004

2. 宁夏医科大学临床医学院 宁夏银川 750004

摘要: **目的:** 本研究立足于TCGA的基因数据集,对前列腺癌与基因SPOP的相关数据进行分析,以期前列腺癌的基因治疗及分子标志物的最终判定奠定基础。**方法:** 正常组织及前列腺癌组织相关数据取自TCGA公共组学数据库。利用SPSS 23.0软件进行数据处理,行t检验、 χ^2 检验和非参数检验对检测指标进行分析比较,以 $P < 0.05$ 为差异具有统计学意义。**结果:** 1.SPOP在前列腺癌中的表达基于不同的指标,包括患者不同的样本类型、年龄、种族、格里森分数、淋巴结转移状态和TP53的转移状态等,表达结果各有不同:SPOP在正常组织中的表达明显高于原发性肿瘤,比较具有极显著的统计学意义($P = 1.05E-06 < 0.01$);SPOP的表达在正常组对高加索人组和正常组对非裔美国人组中, P 值分别为 $P = 5.437E-04 < 0.01$ 和 $P = 5.382E-03 < 0.01$,差异极显著且有统计学意义($P < 0.01$);SPOP的表达在正常组与格里森评分等于6、7、8、9的组间分别比较, P 值分别为 $4.24E-05$ 、 $2.27E-06$ 、 $1.95E-06$ 和 $2.04E-07$,差异均极显著,有统计学意义($P < 0.01$);SPOP表达在正常组与无区域淋巴结转移组和有区域淋巴结转移组间的比较, P 值分别为 $1.390E-06$ 和 $4.244E-07$,差异极显著,均具统计学意义($P < 0.01$);SPOP表达在正常组与TP53突变组及TP53非突变状态组间的比较, P 值分别为 $3.620E-06$ 和 $9.703E-07$,差异极显著,均具统计学意义($P < 0.01$)。2.经Log-Rank检验比较得到,患者SPOP的表达水平基于格里森评分是其预后的不良因素,差异极显著,有统计学意义($P < 0.01$)。**结论:** 1.SPOP在前列腺癌中是作为抑癌基因存在的,此基因突变会促进前列腺癌的发生发展。2.前列腺癌患者的发病和死亡与其所在地域和种族密切相关。3.临床诊断应重视SPOP突变检测。

关键词: 前列腺癌; SPOP; 基因表达; 预后生存

Bioinformatics analysis of the correlation between spop gene and prostate cancer

JiaYu¹, Jingnan Zhang², Junbai Ma¹, Huan Liu¹

1. School of basic medicine, Ningxia Medical University, Yinchuan, Ningxia, 750004

2. School of clinical medicine, Ningxia Medical University, Yinchuan, Ningxia, 750004

Abstract: **Objective:** Based on the gene data set of TCGA, this study analyzed the relevant data of prostate cancer and gene spop, in order to lay a foundation for gene therapy and final determination of molecular markers of prostate cancer. **Methods:** the data of normal tissues and prostate cancer tissues were obtained from TCGA public omics database. SPSS 23.0 software was used for data processing, t-test χ^2 test and nonparametric test were used to analyze and compare the detection indexes, and the difference was statistically significant ($P < 0.05$). **Results:** 1. The expression of spop in prostate cancer was based on different indicators, including different sample types, age, race, Gleason score, lymph node metastasis and TP53 metastasis. The expression results were different: the expression of spop in normal tissues was significantly higher than that in primary tumors ($P = 1.05e-06 < 0.01$); The expression of spop in normal group vs Caucasian group and normal group vs African American group, $P = 5.437e-04 < 0.01$ and $P = 5.382e-03 < 0.01$, respectively, the difference was very significant and statistically significant ($P < 0.01$); The expression of spop was significantly different between the normal group and the groups

基金项目: 宁夏自然科学基金 (2020AAC03156)

with Gleason score equal to 6, 7, 8 and 9 ($P < 0.01$); The comparison of spop expression between normal group, non regional lymph node metastasis group and regional lymph node metastasis group was $1.390e-06$ and $4.244e-07$, respectively, with significant difference ($P < 0.01$); The P values of spop expression between normal group, TP53 mutation group and TP53 non mutation group were $3.620e-06$ and $9.703e-07$, respectively, with significant difference ($P < 0.01$). 2. Compared with log rank test, the expression level of spop in patients based on Gleason score is an adverse factor of prognosis, and the difference is very significant ($P < 0.01$). Conclusion: 1. Spop exists as a tumor suppressor gene in prostate cancer. This gene mutation will promote the occurrence and development of prostate cancer. 2. The morbidity and mortality of patients with prostate cancer are closely related to their region and race. 3. Clinical diagnosis should pay attention to spop mutation detection.

Keywords: prostate cancer; SPOP; Gene expression; Prognostic survival

前列腺癌 (Prostate Cancer, PCa) 是男性最常见的泌尿系统恶性肿瘤之一, 其排名继肺癌之后位居全球男性癌症死因的第二位^[1], 年龄和家族史为其主要危险因素。因其恶性程度较高, 预后相对较差, 严重影响男性的生活质量。前列腺癌发病率呈明显上升的趋势, 我国前列腺癌发病率虽远低于欧美国家, 但近年来, 随着居民预期寿命的延长, 饮食结构、生活方式的改变以及前列腺癌检出率和诊断水平的不断提高^[2], 已成为男性最常见的十大肿瘤之一^[3]。前列腺癌起病隐匿, 临床发现多为晚期且死亡率也在随着年龄的增长而上升, 目前尚不能有效地治疗这种病情, 因此急需更直观、更有效、能够直接或辅助预测前列腺癌高危人群以及指导治疗的标志物, 借此提升患者的生存率。

TCGA (The Cancer Genome Atlas) 数据库由美国国家癌症研究所 (NCI) 和国家人类基因组研究所 (NHGRI) 合作创建, 收集了各种癌症相关的组学数据^[4], 为人们提供了一个大型且免费的癌症研究参考数据库^[5-7]。SPOP是一种已被鉴定结构的E3泛素连接酶, 研究发现, SPOP参与调控多种细胞活动, 影响细胞的基因组修饰和信号转导, 在肿瘤的形成和发展中起重要作用。SPOP在肿瘤中所发挥的作用依赖于特定的肿瘤类型, 有研究表明, 在前列腺癌中, SPOP发挥抑癌基因作用, 其基因突变会诱发肿瘤^[8]。本研究旨在阐明SPOP的表达与前列腺癌的相关性及其对患者预后生存率的影响, 以期前列腺癌分子标志物的判定及选择相应靶序列进行基因治疗奠定基础。

1 一般资料与方法

1.1 一般资料

笔记本电脑、校园网以及图书馆资源。以“prostate cancer; SPOP”为关键词, 在数据库TCGA (网址: <https://portal.gdc.cancer.gov/>) 中搜索, 对得到的数据进行收集和整理, 筛选SPOP基因的有效信息。

1.2 评价指标

从3个角度评价SPOP与前列腺癌的关系: 基因表达、甲基化状态和总体生存率。

1.2.1 基因表达及其相关指标: 患者的样本类型、年龄、种族、格里森评分、淋巴结转移状态以及TP53突变状态。

1.2.2 总体生存率通过Kaplan-Meier法分析得到。

1.3 统计学方法

利用SPSS 23.0软件对收集到的数据进行相关分析, 行t检验、 χ^2 检验和非参数检验对不同指标进行统计比较。数据均以平均值 \pm 标准差 (SD) 或者患者人数和百分比表示, 以 $P < 0.05$ 为差异具有统计学意义^[8]。

2 结果

2.1 SPOP在前列腺癌中表达的结果

根据不同检测指标, 得出下图1 A-F:

1. 基于样本类型: SPOP在正常组织中的表达明显高于原发肿瘤, 比较具有极显著的统计学意义 ($P=1.05E-06 < 0.01$)。

2. 基于患者年龄: SPOP的表达在正常人群与前列腺癌患者年龄介于41-60岁和61-80岁比较, 均不适用统计分析; 而患者年龄介于41-60岁和61-80岁比较, 差异无统计学意义 ($P > 0.05$)。

3. 基于患者种族: 不论是正常组对高加索人组还是正常组对非裔美国人组, SPOP的表达均有极显著的差异, 分别为 $P=5.437E-04 < 0.01$ 和 $P=5.382E-03 < 0.01$, 差异有统计学意义。高加索人和非裔美国人的比较中, 差异无统计学意义 ($P > 0.05$)。

4. 基于患者的格里森得分: SPOP的表达在正常组与格里森评分等于6、7、8、9的组间分别比较, P值分别为 $4.24E-05$ 、 $2.27E-06$ 、 $1.95E-06$ 和 $2.04E-07$, 差异均极显著, 有统计学意义 ($P < 0.01$)。正常组与格里森评分为10的组间, 差异不明显 ($P > 0.05$)。

5. 基于淋巴结转移状态: SPOP表达在正常组与无区域淋巴结转移组和有区域淋巴结转移组间的比较, P值

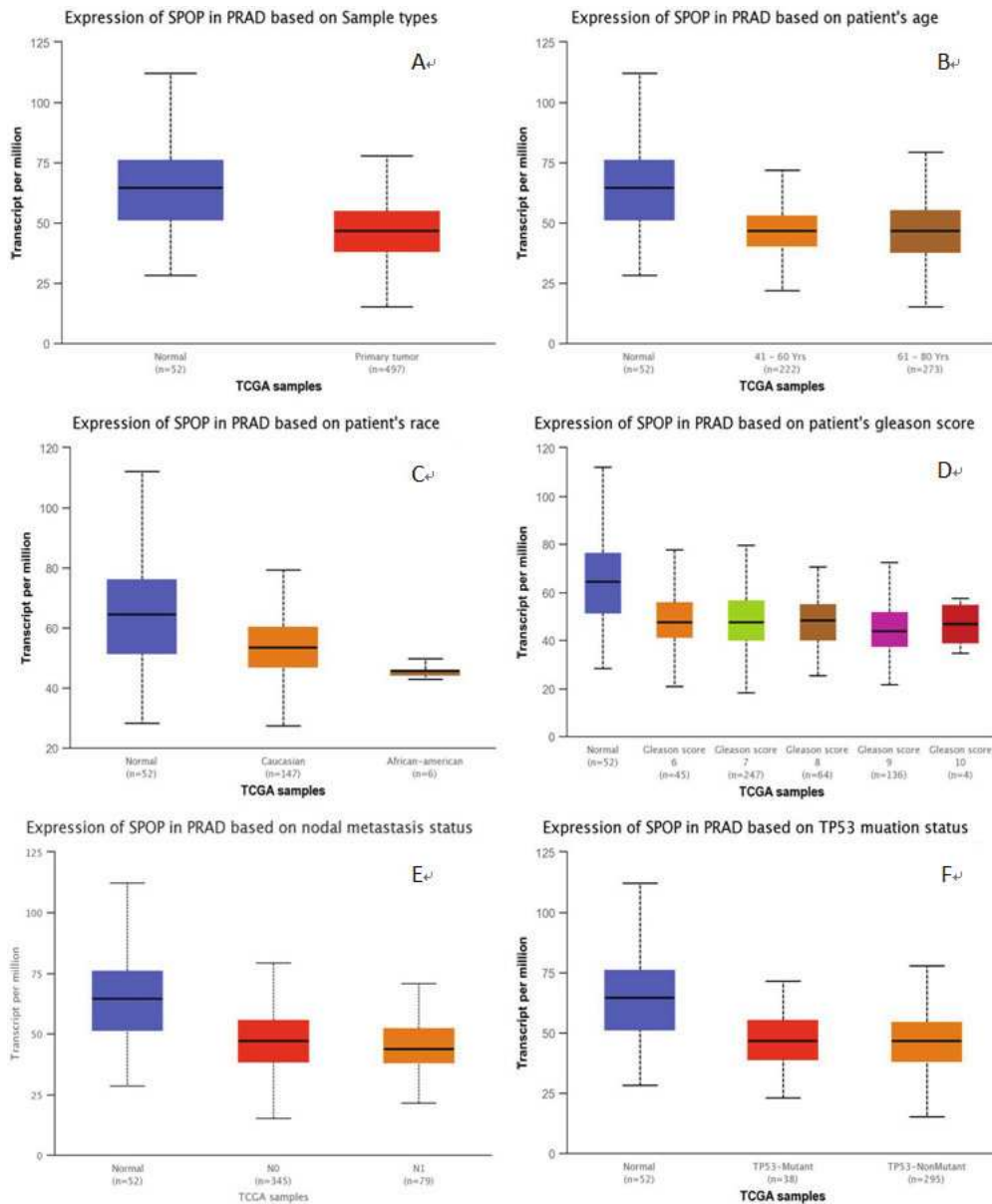


图1 SPOP在前列腺癌中的表达结果

A. 基于样本类型, B. 基于患者年龄, C. 基于患者种族, D. 基于患者的格里森评分, E. 基于淋巴结转移状态, F. 基于TP53突变状态

分别为 $1.390E-06$ 和 $4.244E-07$, 差异极显著, 均具统计学意义 ($P < 0.01$)。

6. 基于TP53的突变状态: SPOP表达在正常组与TP53突变组及TP53非突变状态组间的比较, P值分别为 $3.620E-06$ 和 $9.703E-07$, 差异极显著, 均具统计学意义 ($P < 0.01$)。

2.2 Kaplan-Meier法分析前列腺癌的总生存率结果
患者样本类型、种族、分子特征和格里森评分这些危险因素, 影响着患者的预后生存。利用Kaplan-Meier法分析上述因素^[9], 经Log-Rank检验比较, SPOP的格里

森评分为患者预后不良因素, $P = 0.0001 < 0.05$, 差异极显著, 有统计学意义, 如图2 B。

3 讨论

通过数据分析, 不难看出, 基于患者的种族、样本类型、格里森分数、淋巴结转移状态以及TP53突变状态等考量, SPOP的表达在正常人群及前列腺肿瘤患者中的表达均有极大的不同, 正常人群表达量高, 癌症患者表达量低, 差异极显著且具有统计学意义 ($P < 0.01$)。我们的结果和前文的研究成果不谋而合——SPOP在前列腺癌中是作为抑癌基因存在的。若此基因突变, 失去抑癌

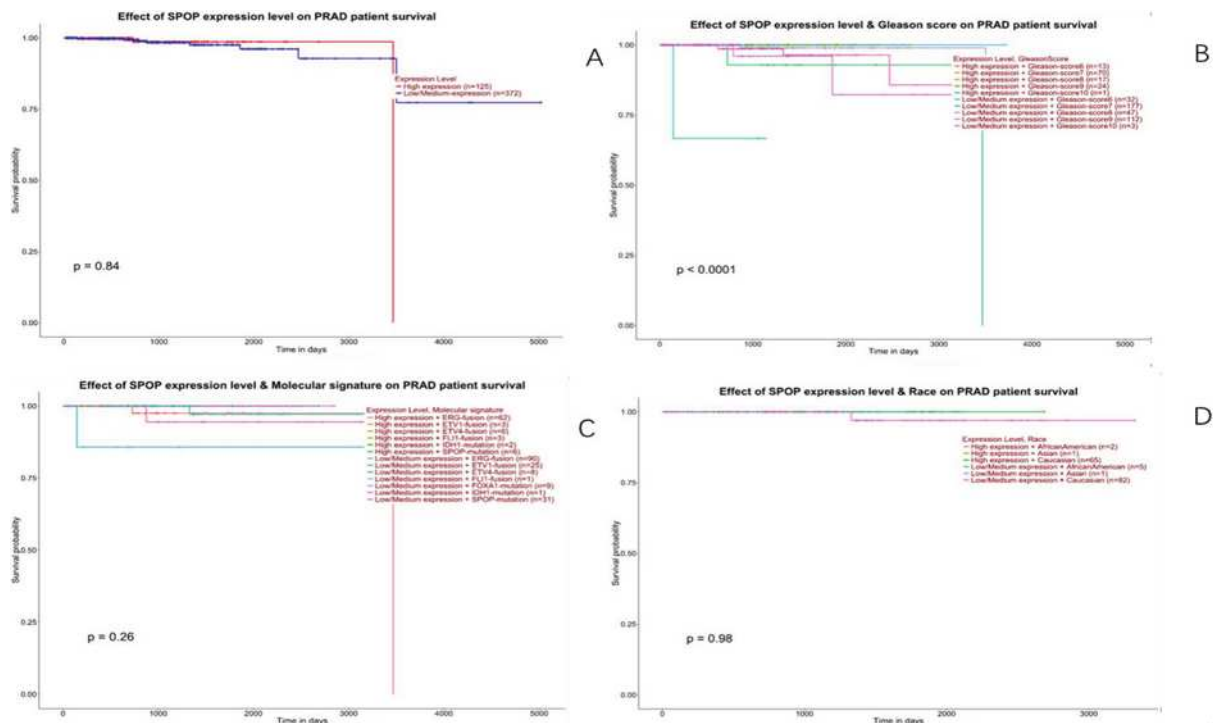


图2 Kaplan-Meier 曲线分析患者总体生存率与样本类型 (A)、格里森评分 (B)、分子特征 (C) 和种族 (D) 的关系

作用,明显会促进前列腺癌的发生发展。因此,建议将SPOP基因的突变检测纳入临床常规检查中。

格里森评分是反应前列腺癌组织分化程度的金指标,评分高低与肿瘤的恶性程度呈正比。在本研究中,格里森评分为患者预后不良因素,且SPOP的表达在正常组与格里森评分6-9的患者组间比较,差异均极显著且有统计学意义 ($P < 0.01$),对应的P值几乎是越来越小,说明差异越来越大。侧面反映了SPOP的抑癌作用会随着肿瘤恶性程度的增加而减弱。这或许是因为前列腺癌的早期,肿瘤恶性程度低,SPOP能正常发挥抑癌作用;而发展到后期,肿瘤恶性程度高,SPOP的抑癌作用已经不能遏制肿瘤的疯狂扩增,也因此,比较到格里森评分为10的患者组,数据上反倒没有了统计学意义上的差异。

而前列腺癌患者的发病和死亡往往也与其所在地域和种族密切相关。因此,建议男性前列腺癌高发的种族和地区,重视SPOP基因突变的筛查。

参考文献:

[1]熊家伟,王秀云,汪佳旭等.前列腺癌诊断的最新进展[J].中华医学超声杂志(电子版),2019,016(006):470-473.

[2]李铁求.前列腺癌转移相关基因的生物信息学分析及功能预测[D].南方医科大学,2014.

[3]张娟娟,陈英红,林旭红等.超声靶向微泡破坏联合脂质体介导pSilencer 3.1-SATB1基因转染前列腺癌细胞效果观察[J].山东医药,2017,30(v.9):28-31.

[4]徐磊,郭剑明.晚期前列腺癌治疗的规范及最新进展[J].实用肿瘤杂志,2020,v.35(02):12-18.

[5]Xu S, Kong D, Chen Q, et al. Oncogenic long noncoding RNA landscape in breast cancer[J]. Mol Cancer, 2017,16(1):129.

[6]Giulietti M, Occhipinti G, Principato G, et al. Weighted gene co-expression network analysis reveals key genes involved in pancreatic ductal adenocarcinoma development[J]. Cell Oncol, 2016,39(4):379-388.

[7]顾方六.现代前列腺病学[M].北京:人民军医出版社,2002:276.

[8]鲍镇美.晚期前列腺癌的治疗新动向[J].中华泌尿外科杂志,2002,23(2):69-71.

[9]韩苏军,张思维,陈万青等.中国前列腺癌发病现状和流行趋势分析[J].临床肿瘤学杂志,2013,18(4):330-334.