Oncology and Translational Medicine

Volume 7 • Number 6 • December 2021

Oncology and Translational Medicine Volume 7 • Number 6 • December 2021 pp 245-313

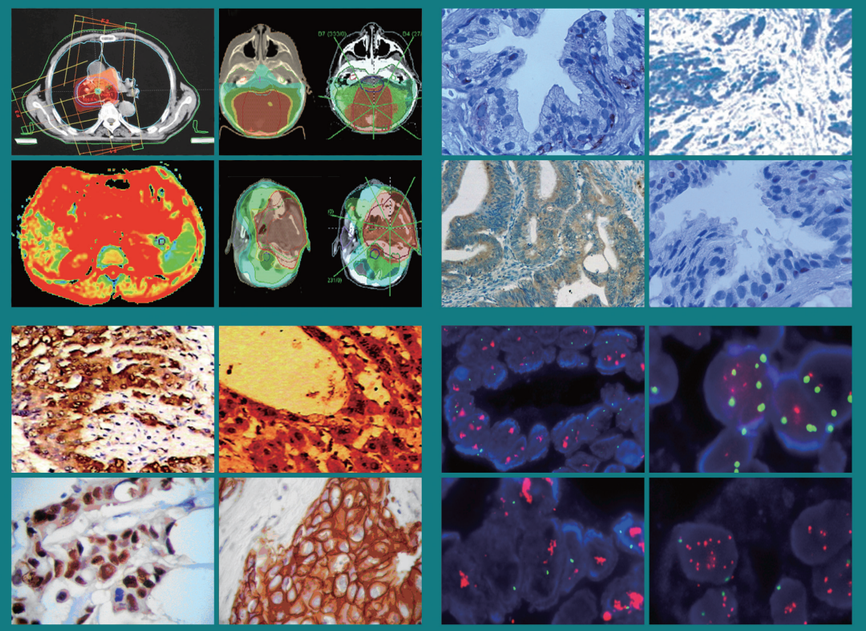
邮发代号：38-121
ISSN 2095-9621

GENERAL INFORMATION
≫otm.tjh.com.cn

ISSN 2095-9621
CN 42-1865/R

Oncology and Translational Medicine

肿瘤学与转化医学（英文）



Volume 7
Number 6
December 2021

Oncology and Translational Medicine

Oncology and Translational Medicine
December 2021 Volume 7 Number 6

Contents

Oncology and Translational Medicine

Volume 7 • Number 6 • December 2021

Aims & Scope

Oncology and Translational Medicine is an international professional academic periodical. The Journal is designed to report progress in research and the latest findings in domestic and international oncology and translational medicine, to facilitate international academic exchanges, and to promote research in oncology and translational medicine as well as levels of service in clinical practice. The entire journal is published in English for a domestic and international readership.

Copyright

Supervised by

Ministry of Education of the People's Republic of China.

Administered by

Tongji Medical College, Huazhong University of Science and Technology.

Submission information

Manuscripts should be submitted to:
http://otm.tjh.com.cn
dmedizin@sina.com

Subscription information

ISSN edition: 2095-9621
CN: 42-1865/R

■ Subscription rates
Subscription may begin at any time. Remittances made by check, draft or express money order should be made payable to this journal. The price for 2021 is as follows: US $ 30 per issue; RMB ￥28.00 per issue.

Database

Oncology and Translational Medicine is abstracted and indexed in EMBASE, Index Copernicus, Chinese Science and Technology Paper Citation Database (CSTPCD), Chinese Core Journals Database, Chinese Journal Full-text Database (CJFD), Wanfang Data; Weipu Data; Chinese Academic Journal Comprehensive Evaluation Database.

Business correspondence

All matters relating to orders, subscriptions, back issues, offprints, advertisement booking and general enquiries should be addressed to the editorial office.

Mailing address
Editorial office of
Oncology and Translational Medicine
Tongji Hospital
Tongji Medical College
Huazhong University of Science and Technology
Jie Fang Da Dao 1095
430030 Wuhan, China
Tel.: +86-27-69378388
Email: dmedizin@sina.com

Printer
Changjiang Spatial Information Technology Engineering Co., Ltd. (Wuhan) Hangce Information Cartorgraphy Printing Filial, Wuhan, China
Printed in People's Republic of China

Editors-in-Chief

Managing director

Executive editors

Oncology and Translational Medicine
DOI 10.1007/s10330-021-0511-1

December 2021, Vol. 7, No. 6, P245–P252

REVIEW ARTICLE

Recent progress in understanding the role of genes in the pathogenesis of cutaneous squamous cell carcinoma*

Yong He[1], Yilin Wu[1] (Co-first author), Yueyue Zhang[2], Qun Lv[1], Liming Li[1] (✉), Mingjun Jiang[1] (✉)

[1] Hospital of Dermatology, Chinese Academy of Medical Sciences and Peking Union Medical College, Nanjing 210042, China
[2] Outpatient Department, Affiliated Jinling Hospital, Medical School of Nanjing University, Nanjing 210002, China

Abstract

Cutaneous squamous cell carcinoma (cSCC) is the second most common skin tumor in humans. Ultraviolet (UV) radiation is an important environmental risk factor for cSCC; other risk factors include human papilloma virus (HPV) infection, chronic inflammation, and chronic wounds. A large proportion of patients present with an aggressive form of cSCC at the time of diagnosis, which is often accompanied by regional lymph node involvement and distant metastases. The long-term prognosis for these highly metastatic diseases is extremely poor, with a 10-year survival rate of less than 10%. Therefore, clarifying the pathogenesis of this tumor is of great significance and may contribute to the identification of novel biomarkers and development of new therapeutic strategies. In this review, we focus on the recent progress in genes related to the development of this tumor, intending to aid future investigations into the genetic alterations related to cSCC.

Received: 2 August 2021
Revised: 5 October 2021
Accepted: 2 November 2021

Key words: cutaneous squamous cell carcinoma (cSCC); genetics; pathogenesis; carcinogenesis

Skin cancers are classified into melanoma and non-melanoma types. Most skin cancers are of the non-melanoma type, which originate from epidermal keratinocytes and are further classified into cutaneous basal cell carcinoma (cBCC) and cutaneous squamous cell carcinoma (cSCC). cSCC is the second most common skin tumor in humans, adversely affecting the quality of life [1]. It most frequently develops in the skin due to long-term exposure to the sun, which results in ultraviolet (UV)-induced DNA damage in the epidermal keratinocytes [2]. cSCC carcinogenesis includes premalignant lesions [actinic keratosis (AK) and *in situ* squamous carcinoma/Bowen's disease] and invasive and metastatic cSCCs, however, a multistep process is not always detected [3]. Although multiple AKs are clearly strong risk factors of developing invasive cSCC, the rate of progression of an AK to invasive cSCC is not precisely known [4]. Most patients with localized cSCC usually have an excellent outcome if the lesion is completely excised by surgery [5]. However, a large number of patients have developed an aggressive form of cSCC with distant metastases by the time of diagnosis, leading to both severe morbidity and mortality rates [6]. Furthermore, although radiotherapy and chemotherapy have been utilized in the treatment of advanced cSCC, their effect is very limited [7]. Besides, at present, the knowledge on the molecular basis of cSCC progression from premalignant lesions to cSCC *in situ* and eventually to invasive cSCC is limited [8]. Therefore, clarifying the pathogenesis of this tumor is of great significance and may contribute to the identification of novel biomarkers and new therapeutic strategies [7]. Increasing evidence has demonstrated that tumorigenesis, progression, invasion and metastasis of cSCC involve several genes such as *TP53*, *NOTCH1/2*, *CDKN2A*, *TGFBR1*, and *RAS* [9]. Notwithstanding these advances, the genetic mechanisms of tumor development are far

✉ Correspondence to: Liming Li. Email: llm0706@163.com; Mingjun Jiang. Email: drmingjunjiang@163.com
* Supported by the Natural Science Foundation of Jiangsu Province (No. BK20191136), the Fundamental Research Funds for the Central Universities (No. 3332019104), and the Open Project of Jiangsu Biobank of Clinical Resources (No. JSSWYB2020-05-003).

from clarified. With this in mind, we sought to review the latest advances regarding the role of mutated genes in carcinogenesis (Table 1). This review intends to aid future investigations into the genetic alterations related to cSCC. To simplify the analysis, the genetic alterations are presented according to different mechanisms associated with the development of cSCC.

Proliferation and apoptosis

In previous studies, many genes have been shown to play crucial roles in regulating the proliferation and apoptosis of tumor cells. For example, *p53*, the gene most commonly and earliest mutated in cSCC, prevents tumor cell apoptosis and allows clonal expansion of tumor cells. The *CDKN2A* gene encodes two alternatively spliced proteins, p16 and p14, which regulate cell cycle progression and proliferation through the retinoblastoma and p53 pathways, respectively [10]. In recent years, more genes such as *NBPF1*, *miR-221*, and *ID4* have been found to be involved in the development of cSCC by regulating the proliferation and apoptosis of tumor cells.

Inhibitor of DNA binding/differentiation 4 (ID4)

ID4 is a downstream mediator of the TGF-β/BMP/SMAD signaling pathway and regulates the growth and differentiation of embryonic tissues [11]. Our research group found that UVB exposure could downregulate ID4 expression via DNA methylation to initiate cutaneous tumorigenesis [12]. Silencing of *DNMT1* and overexpression of TET1 and TET2 can increase ID4 expression, leading to reduced cell proliferation, migration, and invasion, and increased apoptosis in cSCC cell lines [12]. Based on the results presented above, *ID4* acts as a tumor suppressor gene in cSCC carcinogenesis [12].

miR-221

miR-221 is a member of the *miR-221/222* cluster, which is located on the X chromosome [13]. It is significantly upregulated in cSCC tissues and cell lines. It can regulate several hallmarks of cSCC, including cell growth and colony formation. In addition, *miR-221* may act as anoncogene, and its aberrant expression may be linked to the progression of human cSCC. By targeting and repressing the expression of PTEN, miR-221 can regulate the expression of numerous genes related to cell proliferation, apoptosis, and invasion and is implicated in the progression of several tumors. These results suggest that miR-221 may be a potential target for cSCC diagnosis and treatment [14].

Table 1 The role of mutated genes in carcinogenesis

Function	Gene
Proliferation and apoptosis	*ID4*, *ALK*, *miR-221*, *NBPF1*, *miR-506*, *Drp1*
Wnt signaling pathway	*SFRP*, *TCF4*, *HOXB7*
MAPK pathway	*TPL2*, *miR-202*
Terminal differentiation factors	*CYFIP1*, *P63*
Glycolysis	*HOXA9*
Inflammsome	*ASC*
Epidermal growth factor receptor	*EGFR-PPARGC1A*
Migration, invasion and microenvironment	*OVOL1-OVOL2*, *EphB2*, *NQO1*

Anaplastic lymphoma kinase (ALK)

ALK, a receptor tyrosine kinase of the insulin receptor superfamily, plays a pivotal role in the pathogenesis of cSCC. The overexpression of the mutated ALK, ALK^{F1174L}, is able to initiate the development of cSCC. ALK^{F1174L} cooperates with oncogenic $Kras^{G12D}$ and loss of *p53*, resulting in a more aggressive cSCC type associated with a higher histological grade. As mentioned above, inactivation of *p53* induces cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. As a key player in the pathogenesis of cSCC, oncogenic ALK signaling may serve as a target for future clinical trials [15].

Neuroblastoma breakpoint family member 1 (NBPF1)

NBPF1 is located on chromosome 1p36 where many tumor suppressor genes have been identified. The *NBPF1* gene is expressed at low levels in cSCC tissues and shows the lowest expression in the A431 cell line. In the A431 cell line, increased expression of NBPF1 leads to a significant decrease in cell viability and cell cycle arrest in the G1 phase. Meanwhile, overexpression of NBPF1 promotes apoptosis by promoting Bax and inhibiting Bcl-2 and survivin. Bax directly regulates apoptosis-related proteins and promotes apoptosis. Bcl-2 can inhibit apoptosis, and its overexpression is common in cSCC. Survivin is associated with cell viability [16–17]. In addition, *NBPF1* can inhibit the activation of the Akt-p53-Cyclin signaling pathway. Akt regulates a variety of signaling pathways and is involved in tumor proliferation and cell apoptosis [18]. Cyclin D1 and p53 are important Akt downstream factors that directly regulate the cell cycle [19]. By inhibiting the phosphorylation of Akt protein, *NBPF1* can inhibit the activation of p53 and cyclin D1, thereby promoting apoptosis and arresting the cell cycle in the G1 phase [20].

miR-506

MicroRNAs (miRNAs) are non-coding RNAs that have a regulatory effect on protein expression at the post-transcription level [21]. They are involved in the regulation of many biological processes such as proliferation, differentiation, migration, and invasion of cells. miR-506 expression is upregulated in cSCC tissues. Decreased miR-506 levels result in decreased proliferation of cSCC cells. Furthermore, miR-506 inhibition can also induce apoptosis and autophagy in cSCC cells. In addition, miR-506 decreases cSCC cell migration and invasion *in vitro*. *miR-506* functions as an oncogene in cSCC by targeting *p65* and *LAMC1* genes. P65 is a member of the NF-κB family, which can regulate many genes associated with apoptosis, proliferation, and differentiation of cells. The silencing of *miR-506* increases p65 expression, and consequently increases cellular apoptosis and impairs cell viability. *LAMC1*, an oncogene that belongs to the laminin family, is associated with the metastasis, signaling, differentiation, and adhesion of tumor cells. Silencing of *LAMC1*, which can be directly targeted by miR-506 in cSCC cells, restores the migration and invasion properties of cSCC cells. Thus, it plays an important role in the activation and progression of cSCC. In conclusion, reduced miR-506 expression is highly associated with impaired tumor cell growth. Therefore, miR-506 can be further developed as a diagnostic and prognostic biomarker for cSCC [22].

Dynamin-related protein 1 (Drp1)

Drp1, a cytosolic protein, can mediate mitochondrial fission [23] and is involved in the process of cell proliferation or cell remodeling that facilitates the development of malignant neoplasms. Drp1 is more highly expressed in SCC than in the normal epidermis. Drp1 knockdown causes ATM-dependent G2/M cell cycle arrest and apoptosis. Morphologically, the depletion of Drp1 results in an elongated, hyper-fused mitochondrial network [24]. Disrupted mitochondrial networks promote cell cycle arrest and apoptosis. In addition, Drp1 can also be regarded as a prognostic factor in several malignancies, and the expression levels of Drp1 positively correlate with advanced clinical stages. In conclusion, Drp1 plays an important role in cell proliferation, apoptosis, and cell cycle in cSCC and may serve as a novel target for skin tumor therapies [25].

Wnt signaling pathway

The Wnt/β-catenin signaling pathway is involved in cell growth, and its inhibition may lead to abnormal cell growth and differentiation. The abnormal expression of the Wnt pathway activates growth- and mitosis-related genes such as *c-myc* and *cyclin D1*, thus leading to the proliferation of tumor cells.

Secreted frizzled-related protein (SFRP)

SFRPs have been identified as negative regulators of Wnt signaling and play an important role in oncogenic activation of the Wnt pathway and tumor progression.

In hepatocellular carcinoma, SFRP1 can attenuate Wnt signaling, decrease the abnormal accumulation of β-catenin in the nucleus, and suppress cell growth. However, the precise role of the Wnt pathway in cSCC is unclear. Moreover, SFRP promoter hypermethylation has been identified in human cancers. Hypermethylation of these SFRPs, particularly SFRP1, is associated with the development of cSCC. The SFRP1 CpG site can be a possible biomarker of cSCC [26]. Besides, SFRP1 loss results in early tumor initiation and cancer stem cell regulation. In an *in vivo* mouse skin carcinogenesis model of multiple human epithelial cancers, SFRP1 downregulation was found to be associated with poor survival [27]. Nevertheless, further studies are necessary to clarify the roles of SFRPs in Wnt signaling and tumor pathogenesis.

T-cell factor 4 (TCF4)

TCF4 is a high-mobility group (HMG) box-containing transcription factor that activates the Wnt/β-catenin signaling pathway in many cancers. Silencing of TCF4 can effectively inhibit tumor cell growth and invasion, indicating that TCF4 plays an oncogenic role in carcinogenesis and the development of cSCC. It may also be a novel therapeutic target for cancer treatment. In addition, TCF4 knockdown can also regulate the MAPK, insulin, and Rap1 signaling pathways. The MAPK pathway could antagonize the activity of Wnt/β-catenin, whereas insulin and Rap 1 can affect downstream targets of the Wnt/β-catenin pathway. Additionally, in cSCC cells, aberrant activation of TCF4 may result in a Wnt/β-catenin-independent regulation of gene transcription. In conclusion, TCF4 plays an important role in the development of cSCC via activation of different signaling pathways and may be a new therapeutic target for cSCC [28].

Homeobox B7 (HOXB7)

HOXB7 gene, a member of the HOX family, serves as a transcriptional factor that regulates cell viability, growth, morphogenesis, and differentiation. Overexpression of HOXB7 is common in various cancers and is associated with tumorigenesis and tumor proliferation. Cancer patients with a higher expression of HOXB7 are more susceptible to distant metastasis and have a lower survival rate. The knockdown of *HOXB7* can inhibit the expression of Wnt/β-catenin signaling pathway-related downstream genes, including *c-myc*, *cyclin D1*, and *LEF1*. Through inactivation of the Wnt/β-catenin signaling

pathway, silencing of *HOXB7* can promote cell apoptosis and suppress cell migration and invasion in cSCC. Further studies are needed to assess whether HOXB7 can serve as a therapeutic target and prognostic biomarker [6].

MAPK pathway

The MAPK/ERK pathway is the most important cell survival pathway in non-tumorigenic keratinocytes and is triggered by EGF. Negative MAPK regulation and EGFR-induced STAT3 activation can increase the expression of anti-apoptotic molecules and thus lead to malignant progression of keratinocytes towards cSCC [1].

Tumor progression locus 2 (TPL2)

TPL2 is a serine/threonine MAP kinase kinase kinase 8 (MAP3K8) that regulates various signaling pathways associated with inflammation and cell growth. TPL2 overexpression has been found in cutaneous metastatic SCC and plays an important role in the tumorigenesis of cSCC. The overexpression of TPL2 in immortalized human keratinocytes promotes cell proliferation, inhibits apoptotic cell death, and induces cell transformation by activating its downstream signaling pathways, MEK/ERK MAPK, mTOR, NF-κB, and p38 MAPK. Rapamycin, an mTOR inhibitor, is routinely used for the treatment of SCC. In addition, TPL2 overexpression is necessary for maintaining the iTPL2 TG-driven SCC. The data presented above show that TPL2 is an oncogenic driver in cSCC, and further studies are needed to assess its potential as a new therapeutic target for cSCC treatment [29].

miR-204

miR-204 is an intronic miRNA located at the *TRPM3* gene, and its aberrant expression can affect several biological processes, including proliferation, apoptosis, and invasiveness. cSCC shows low expression of miR-204 compared to AK, a type of precancerous lesion. miR-204 may act as a "rheostat" that controls signaling towards the MAPK pathway or the STAT3 pathway in the progression from AK to cSCC. DNA methylation of the miR-204 promoter can lead to miR-204 silencing, which results in STAT3 activation and translocation to the nucleus. miR-204 activates the MAPK pathway, which is the most important cell survival pathway in non-tumorigenic keratinocytes. Both the MAPK and STAT3 pathways can drive the expression of multiple anti-apoptotic molecules and transform AK to cSCC [30].

Terminal differentiation factors

Previous studies have shown that NOTCH is involved in terminal differentiation of cSCC. Notch signaling has been associated with cellular development, progression, and differentiation [10]. Besides, CYFIP1 and p63 can also regulate differentiation through different mechanisms.

CYFIP1

CYFIP1 functions as a novel invasion inhibitor in a variety of epithelial cancers. It is downregulated in cSCC at both mRNA and protein levels and is associated with differentiation and metastatic properties of tumors. CYFIP1 is a direct Notch1 target in keratinocytes. Notch signaling plays an important role in cell fate determination, and it induces differentiation and suppresses development of cSCC [27]. Moreover, Notch activation is involved in the control of the cell cycle of keratinocytes via p21WAF1/Cip1. NOTCH 1 can also function as a promoter of differentiation and an inhibitor of invasion by inducing CYFIP1 expression. CYFIP1 may be a link between the loss of differentiation and invasive properties of cSCC [31].

P63

P63 gene is a member of the p53/p63/p73 family of transcription factors and plays a critical role in the development and homeostasis of squamous epithelium, such as the epidermis. Dysregulated expression of p63 has been found in many squamous cancers and may contribute to cancer development through disruption of many cellular processes. ΔNP63α, the predominant p63 isoform in stratified squamous epithelium, is a key regulator of epidermal morphogenesis and epithelial tissue homeostasis. It influences keratinocyte lineage commitment, proliferation, and survival and blocks terminal differentiation, apoptosis, and senescence; additionally, it modulates the tissue microenvironment through remodeling of the extracellular matrix and vasculature and potentially influences the tumor immune microenvironment [32]. Besides, p63 may be a strong predictor of poor differentiation in non-melanoma skin cancer [33]. The clarification of the molecular mechanism of p63 holds promise for novel interventions in cancer prevention and treatment.

Glycolysis – HOXA9

HOXA9, a direct target of onco-miR-365, functions as a tumor suppressor in cSCC and is significantly downregulated in cSCC. The hypoxia-inducible factor (HIF)-1 pathway is involved in cancer-related biological processes, including hypoxic response, angiogenesis, glycolysis, and proliferation of cSCC stem-like cells. Loss of HOXA9 upregulates HIF-1α and the downstream glycolytic genes of the HIF-1 pathway, which contributes to glycolytic reprogramming, a key pro-survival mechanism of cancer that helps tumor cells to meet their oxygen demand. Besides, HOXA9 interacts with CRIP2 and epigenetically represses HIF-1α expression and

inhibits the expression of glycolytic genes, such as *HK2*, *GLUT1*, and *PDK1*, which is critical for the inhibition of tumor cell growth. Future studies should focus on the newly identified miR-365-HOXA9-HIF-1α axis that may provide novel intervention targets for cSCC therapy [34].

Inflammasome – Apoptosis-associated speck-like protein (ASC)

The inflammasome adaptor ASC is essential for the secretion of pro-tumorigenic innate cytokines. ASC not only regulates caspase-1 activation and IL-1 expression but also controls cell proliferation in cSCC. *ASC* functions as a tumor suppressor gene, and downregulation of ASC expression by aberrant methylation has been found in numerous cancers. In addition, ASC might regulate the epithelial-mesenchymal transition (EMT)-like dedifferentiation of keratinocytes through activation of p53. Moreover, ASC expression does not correlate with metastatic potential but with the degree of dedifferentiation and can serve as an indicator for highly differentiated tumors. ASC is silenced in cSCC by promoter-specific methylation and impairs inflammasome function. This could be of therapeutic relevance as some treatment options for early skin cancers demand immune system activation [35].

Epidermal growth factor receptor (EGFR) – EGFR-PPARGC1A

Wild-type full-length EGFR is a transmembrane glycoprotein that binds EGF. EGFR activation or overexpression leads to upregulation of both MAPK and PI3K signaling pathways and is involved in the proliferation and pathogenesis of SCC, including cSCC. EGFR-PPARGC1A may induce tumor formation via phosphorylation, probably through conformational changes or through interaction with wild-type endogenous EGFR. *EGFR-PPARGC1A* is a fusion gene that is associated with chronic sun exposure. Detection of EGFR-PPARGC1A by RT-PCR may be useful for the early diagnosis of cSCC, because this fusion can be detected *in situ*. EGFR inhibitors (erlotinib and getinib) and EGFR antibodies (cetuximab and panitumumab) are widely used for lung SCC [36], and cetuximab has been reported to have therapeutic effects against cSCC. Further studies are needed to explore how the fusion gene *EGFR-PPARGC1A* regulates tumor formation in cSCC, which may lead to a better understanding of the pathogenesis of cSCC and the development of EGFR-targeted cancer therapies [7].

Migration, invasion, and microenvironment

Metastasis begins with the invasion of tumor cells into the stroma and migration toward the bloodstream. Multiple genes are involved in the regulation of tumor cell migration and invasion through different signaling pathways.

OVOL1-OVOL2

OVOL1 and *OVOL2* are ubiquitous and conserved genes that encode C2H2 zinc-finger transcription factors in mammals. *OVOL1* and *OVOL2* act as guardians against epithelial-to-mesenchymal transition (EMT) [37] and govern the development, maintenance, and proliferation of epithelial cells via the Wnt signaling pathway. OVOL1, an upstream suppressor of c-myc in squamous cell carcinoma cells, is markedly downregulated in cSCC, and the downregulation of OVOL1 may be responsible for the aberrant expression of c-myc and is related to poor tumor prognosis. OVOL2 is typically expressed in the cytoplasm, but only sporadically in the nucleus. OVOL2 negatively affects the EMT process, and the downregulation of OVOL2 activity in SCC might be involved in the invasiveness of this tumor. OVOL1 can suppress OVOL2 expression, and the OVOL1-OVOL2 axis coordinately regulates the EMT transition process and invasiveness of cSCC [38].

EPHB2

Erythropoietin-producing hepatocellular (Eph) receptors and their ligand ephrins are membrane-bound molecules that are highly expressed in cSCC. EPHB2 functions as a biomarker for cSCC and plays an important role in the early stages of tumor progression to invasive cSCC. *EPHB2* knockdown suppresses the expression of genes involved in cell viability, migration, and invasion. Among the genes most downregulated by *EphB2* knockdown are *MMP1* and *MMP13*, two important proteinases that promote cSCC cell invasion [8]. Besides, silencing of *EPHB2* induces EMT-like morphological changes, which reduce cell-cell adhesion and expression of E-cadherin on the cell surface. EPHB2 plays a crucial role in promoting anchorage-independent cell growth through the suppression of EMT [39]. Therefore, EphB2 may serve as an effective therapeutic target in this invasive skin cancer.

NAD(P)H dehydrogenase 1 (NQO1)

NQO1 is a ubiquitous flavoenzyme that plays a role in the mechanism of cellular defense against oxidative stress. Knockdown of *NQO1* promotes colony forming activity and the proliferation, invasion, and migration

of SCC cells, which may promote cancer development. By contrast, the overexpression of NQO1 can suppress the cell proliferation and colony forming activity. Besides, the expression of NQO1 can regulate the levels of phosphorylated AKT, JNK, and p38 MAPK. This may be one possible mechanism underlying the suppressive function of *NQO1*. Further studies are needed to clarify the precise link between NQO1 and intracellular signaling pathways [40].

Conclusion and future perspectives

In this review, we summarized the latest advances in genes involved in the pathogenesis of cSCC and analyze their role in the development of this cancer. As mentioned before, these genes can regulate many biological processes, such as proliferation, apoptosis, terminal differentiation, glycolysis, and many signaling pathways. Specifically, *HOXB7*, *TCF4*, and *SFRP* can target the Wnt pathway. *TPL2* and many other genes participate in the MAPK and PI3K/Akt/mTOR pathways (Fig. 1). Among these pathways, we identified three pathways that deserve further investigation. The first is the Wnt pathway. It has been long investigated but the mechanism was still not fully elucidated. Recently, our research group (unpublished data) specifically focused on the the Wnt/calcineurin/NFAT pathway, which functions in keratinocyte differentiation, migration, and DNA repair. Furthermore, dysregulation of this signaling pathway contributes to squamous cell carcinoma formation, abnormal growth, and tumorigenic microenvironment. Our research group found that Wnt5a, FZD4, PLC, and NFATc4 are downregulated in cancer tissue. Wnt5a/Ca^{2+} suppresses the development of cSCC, and FZD4 and NFATc4 interact with each other. However, further research is needed to clarify the specific mechanisms. The second pathway is the p63 pathway. P63 can directly target gene transcription and function as a key driver of critical networks linked to cellular identity and cell fate determination. Besides influencing keratinocyte lineage commitment, proliferation, and survival, p63 can modulate the tissue microenvironment and regulate the immune system. As p63 is involved in these coordinated pathways and plays an important role in cSCC development; it may serve as a promising target for cancer treatment. The last pathway involves Drp1, which regulates mitochondrial fission and plays an important role in cell proliferation, apoptosis, and cell cycle in cSCC. Therefore, this pathway may serve as a novel target for skin tumor therapies.

Many new therapeutics targeting these specific pathways are available. For example, cetuximab, an EGFR inhibitor, is administered to patients with cSCC. Patients with locally advanced SCC show good responses to cetuximab. However, it is ineffective in treating distant metastatic diseases. In addition, EGFR inhibitors are used in advanced cSCCs as a second-line treatment after chemotherapy failure and disease progression [10]. The

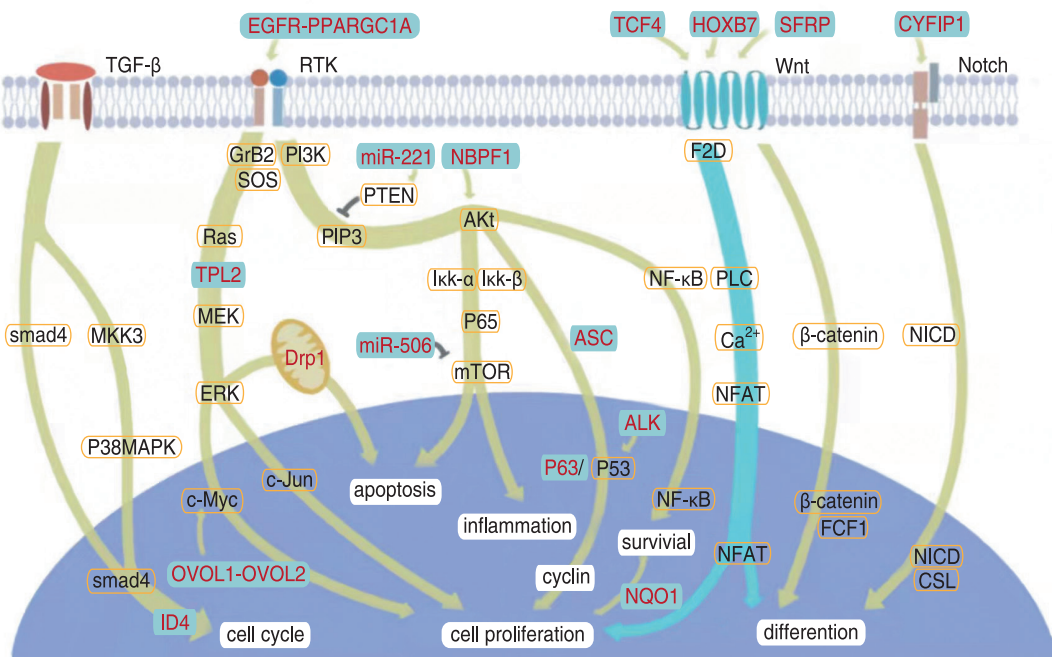


Fig. 1 The figure depicts all the genes that are reviewed in this review. The arrows indicate facilitation of the pathway, and the black T-shaped line indicates inhibition of the pathway. Our research group mainly studied the pathway marked in blue

MEK inhibitor, PD325901, can inhibit cell proliferation, as well as the phosphorylation of ERK1/2 and Drp1 [41]. Good clinical results have been achieved with PD1-inhibitors in the treatment of cSCC, and Cemiplimab is currently the only immune checkpoint inhibitor approved by the FDA and EMA to treat patients with locally advanced or metastatic cSCC. Although many new drugs with various molecular targets have been developed and significant improvements in surgery, chemotherapy, and radiotherapy have been achieved, overall survival of patients with advanced cSCC has not markedly improved [36]. Thus, further studies for a comprehensive understanding of the molecular basis of cSCC are of outstanding importance, especially for patients with metastatic disease in which prognosis is poor and effective therapies are lacking. Considering the complex molecular network, combined therapies targeting different genetic alterations and signaling pathways might provide more effective and personalized therapies for patients with cSCC. More accurate prognostic biomarkers make early intervention possible. In the next few years, scientists will be able to develop effective drugs or prognostic biomarkers that target these genetic alterations and improve the survival rate of patients with cSCC.

Conflicts of interest

The authors indicated no potential conflicts of interest.

References

1. Zhang L, Shan X, Chen Q, *et al*. Downregulation of HDAC3 by ginsenoside Rg3 inhibits epithelial-mesenchymal transition of cutaneous squamous cell carcinoma through c-Jun acetylation. J Cell Physiol, 2019, 234: 22207–22219.
2. Zhang Y, Gao L, Ma SD, *et al*. MALAT1-KTN1-EGFR regulatory axis promotes the development of cutaneous squamous cell carcinoma. Cell Death Differ, 2019, 26: 2061–2073.
3. Campos MA, Macedo S, Fernandes MS, *et al*. Prognostic significance of RAS mutations and p53 expression in cutaneous squamous cell carcinomas. Genes (Basel), 2020, 11: 751.
4. Chen IP, Bender M, Spassova I, *et al*. UV-type specific alteration of miRNA expression and its association with tumor progression and metastasis in SCC cell lines. J Cancer Res Clin Oncol, 2020, 146: 3215–3231.
5. Ci C, Wu C, Lyu D, *et al*. Downregulation of kynureninase restrains cutaneous squamous cell carcinoma proliferation and represses the PI3K/AKT pathway. Clin Exp Dermatol, 2020, 45: 194–201.
6. Gao D, Chen HQ. Specific knockdown of HOXB7 inhibits cutaneous squamous cell carcinoma cell migration and invasion while inducing apoptosis via the Wnt/β-catenin signaling pathway. Am J Physiol Cell Physiol, 2018, 315: C675–C686.
7. Egashira S, Jinnin M, Ajino M, *et al*. Chronic sun exposure-related fusion oncogenes EGFR-PPARGC1A in cutaneous squamous cell carcinoma. Sci Rep, 2017, 7: 12654.
8. Farshchian M, Nissinen L, Siljamäki E, *et al*. EphB2 promotes progression of cutaneous squamous cell carcinoma. J Invest Dermatol, 2015, 135: 1882–1892.
9. Nobeyama Y, Watanabe Y, Nakagawa H. Silencing of G0/G1 switch gene 2 in cutaneous squamous cell carcinoma. PLoS One, 2017, 12: e0187047.
10. Nardo LD, Pellegrini C, Stefani AD, *et al*. Molecular genetics of cutaneous squamous cell carcinoma: perspective for treatment strategies. J Eur Acad Dermatol Venereol, 2020, 34: 932–941.
11. De Candia P, Benera R, Solit DB. A role for Id proteins in mammary gland physiology and tumorigenesis. Adv Cancer Res, 2004, 92: 81–94.
12. Li LM, Li FJ, Xia YD, *et al*. UVB induces cutaneous squamous cell carcinoma progression by de novo ID4 methylation via methylation regulating enzymes. EBioMedicine, 2020, 57: 102835.
13. Mari E, Zicari A, Fico F, *et al*. Action of HMGB1 on miR-221/222 cluster in neuroblastoma cell lines. Oncol Lett, 2016, 12: 2133–2138.
14. Gong ZH, Zhou F, Shi C, *et al*. miRNA-221 promotes cutaneous squamous cell carcinoma progression by targeting PTEN. Cell Mol Biol Lett, 2019, 24: 9.
15. Gualandi M, Iorio M, Engeler O, *et al*. Oncogenic *ALK*^{F1174L} drives tumorigenesis in cutaneous squamous cell carcinoma. Life Sci Alliance, 2020, 3: e201900601.
16. Nomura M, Shimizu S, Sugiyama T, *et al*. 14-3-3 interacts directly with and negatively regulates pro-apoptotic Bax. J Biol Chem, 2015, 290: 6753.
17. Salvador-Gallego R, Mund M, Cosentino K, *et al*. Bax assembly into rings and arcs in apoptotic mitochondria is linked to membrane pores. EMBO J, 2016, 35: 389–401.
18. Bouchard C, Marquardt J, Brás A, *et al*. Myc-induced proliferation and transformation require Akt-mediated phosphorylation of FoxO proteins. EMBO J, 2004, 23: 2830–2840.
19. Sun HF, Ding CM, Zhang HY, *et al*. Let-7 miRNAs sensitize breast cancer stem cells to radiation-induced repression through inhibition of the cyclin D1/Akt1/Wnt1 signaling pathway. Mol Med Rep, 2016, 14: 3285–3292.
20. Gao Y, Zhu H, Mao Q. Effects of neuroblastoma breakpoint family member 1 (NBPF1) gene on growth and Akt-p53-Cyclin D pathway in cutaneous squamous carcinoma cells. Neoplasma, 2019, 66: 584–592.
21. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell, 2004, 116: 281–297.
22. Zhou J, Zhang Y, Han ZF, *et al*. miR-506 contributes to malignancy of cutaneous squamous cell carcinoma via targeting of P65 and LAMC1. Cell Cycle, 2019, 18: 333–345.
23. Frank S, Gaume B, Bergmann-Leitner ES, *et al*. The role of dynamin-related protein 1, a mediator of mitochondrial fission, in apoptosis. Dev Cell, 2001, 1: 515–525.
24. Qian W, Choi S, Gibson GA, *et al*. Mitochondrial hyperfusion induced by loss of the fission protein Drp1 causes ATM-dependent G2/M arrest and aneuploidy through DNA replication stress. J Cell Sci, 2012, 125: 5745–5757.
25. Kitamura S, Yanagi T, Imafuku K, *et al*. Drp1 regulates mitochondrial morphology and cell proliferation in cutaneous squamous cell carcinoma. J Dermatol Sci, 2017, 88: 298–307.
26. Liang JQ, Kang XJ, Halifu Y, *et al*. Secreted frizzled-related protein promotors are hypermethylated in cutaneous squamous carcinoma compared with normal epidermis. BMC Cancer, 2015, 15: 641.
27. Li Y, Huang CL, Yang XC. Characterization of TCF4-mediated oncogenic role in cutaneous squamous cell carcinoma. Int J Clin Exp Pathol, 2019, 12: 3583–3594.
28. Bray SJ. Notch signalling: a simple pathway becomes complex. Nat Rev Mol Cell Biol, 2006, 7: 678–689.

29. Lee JH, Lee JH, Lee SH, *et al*. TPL2 is an oncogenic driver in keratocanthoma and squamous cell carcinoma. Cancer Res, 2016, 76: 6712–6722.
30. Toll A, Salgado R, Espinet B, *et al*. MiR-204 silencing in intraepithelial to invasive cutaneous squamous cell carcinoma progression. Mol Cancer, 2016, 15: 53.
31. Dziunycz PJ, Neu J, Lefort K, *et al*. CYFIP1 is directly controlled by NOTCH1 and down-regulated in cutaneous squamous cell carcinoma. PLoS One, 2017, 12: e0173000.
32. Moses MA, George AL, Sakakibara N, *et al*. Molecular mechanisms of p63-mediated squamous cancer pathogenesis. Int J Mol Sci, 2019, 20: 3590.
33. Koster MI, Kim S, Roop DR. P63 deficiency: a failure of lineage commitment or stem cell maintenance? J Investig Dermatol Symp Proc, 2005, 10: 118–123.
34. Zhou L, Wang YH, Zhou MJ, *et al*. HOXA9 inhibits HIF-1α-mediated glycolysis through interacting with CRIP2 to repress cutaneous squamous cell carcinoma development. Nat Commun, 2018, 9: 1480.
35. Meier K, Drexler SK, Eberle FC, *et al*. Silencing of ASC in cutaneous squamous cell carcinoma. PLoS One, 2016, 11: e0164742.
36. Campos MA, Lopes JM, Soares P. The genetics of cutaneous squamous cell carcinogenesis. Eur J Dermatol, 2018, 28: 597–605.
37. Li S, Yang J. Ovol proteins: guardians against EMT during epithelial differentiation. Dev Cell, 2014, 29: 1–2.
38. Ito T, Tsuji G, Ohno F, *et al*. Potential role of the OVOL1-OVOL2 axis and c-Myc in the progression of cutaneous squamous cell carcinoma. Mod Pathol, 2017, 30: 919–927.
39. Inagaki Y, Tokunaga T, Yanai M, *et al*. Silencing of *EPHB2* promotes the epithelial-mesenchymal transition of skin squamous cell carcinoma-derived A431 cells. Oncol Lett, 2019, 17: 3735–3742.
40. Zhang QL, Li XM, Lian DD, *et al*. Tumor suppressive function of NQO1 in cutaneous squamous cell carcinoma (SCC) cells. Biomed Res Int, 2019, 2019: 2076579.
41. Yanagi T, Kitamura S, Hata H. Novel therapeutic targets in cutaneous squamous cell carcinoma. Front Oncol, 2018, 8: 79.

DOI 10.1007/s10330-021-0511-1
Cite this article as: He Y, Wu YL, Zhang YY, *et al*. Recent progress in understanding the role of genes in the pathogenesis of cutaneous squamous cell carcinoma. Oncol Transl Med, 2021, 7: 245–252.

Oncology and Translational Medicine
DOI 10.1007/s10330-021-0545-5
December 2021, Vol. 7, No. 6, P253–P268

ORIGINAL ARTICLE

Development and validation of a tumor microenvironment-related prognostic signature in lung adenocarcinoma and immune infiltration analysis*

Zhou Li[1], Yanqi Feng[1], Piao Li[1], Shennan Wang[1], Ruichao Li[2], Shu Xia[1] (✉)

[1] Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China
[2] Department of Geriatric, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

Abstract

Objective Tumor-infiltrating immune cells and stromal cells in the tumor microenvironment (TME) significantly affect the prognosis of and immune response to lung adenocarcinoma (LUAD). In this study, we aimed to develop a novel TME-related prognostic model based on immune and stromal genes in LUAD.

Methods LUAD data from the TCGA database were used as the training cohort, and three Gene Expression Omnibus (GEO) datasets were used as the testing cohort. The Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data algorithm was used to analyze the immune and stromal genes involved in the TME. Kaplan-Meier and Cox regression analyses were used to identify prognostic genes and construct a TME-related prognostic model. Gene set enrichment analysis and TIMER were used to analyze the immune features and signaling pathways of the model.

Results A TME-related prognostic model based on six hub genes was generated that significantly stratified patients into the high- and low-risk groups in terms of overall survival. The model had strong predictive ability in both the training (TCGA) and testing (GEO) datasets and could serve as an independent prognostic factor for LUAD. Moreover, the low-risk group was characterized by greater immune cell infiltration and antitumor immune activity than the high-risk group. Importantly, the signature was closely associated with immune checkpoint molecules, which may serve as a predictor of patient response to immunotherapy. Finally, the hub genes BTK, CD28, INHA, PIK3CG, TLR4, and VEGFD were considered novel prognostic biomarkers for LUAD and were significantly correlated with immune cells.

Conclusion The TME-related prognostic model could effectively predict the prognosis and reflect the TME status of LUAD. These six hub genes provided novel insights into the development of new therapeutic strategies.

Received: 19 December 2021
Revised: 25 December 2021
Accepted: 29 December 2021

Key words: lung adenocarcinoma; tumor microenvironment; immunotherapy; immune checkpoint molecules; prognostic biomarkers

Lung cancer is the leading cause of cancer-related deaths worldwide, with a predicted 5-year survival rate of 16% [1]. More than 85% of cases are classified as non-small cell lung cancer (NSCLC), with lung adenocarcinoma (LUAD) being the most common pathological subtype [2]. In recent decades, the discovery of driver gene mutations in tumors has allowed for the introduction of personalized molecular-targeted therapy for NSCLC [3]. However, this approach is not feasible for treating tumors that do not carry gene alterations, and the inevitable resistance to tyrosine kinase inhibitors further suggests the need for alternative therapeutic options in lung cancer patients [4]. In recent decades, immunotherapy targeting immune checkpoints has made great progress in the treatment of

✉ Correspondence to: Shu Xia. Email: xiashutj@hotmail.com
*Supported by grants from the National Natural Science Foundation of China (No.81772471 and 82172716).

NSCLC [5]. Immune checkpoint inhibitors (ICIs) enhance the antitumor effects of the immune system to obtain a potent and durable cure [6]. However, the overall response rate to ICIs is relatively low, and only one-fifth of cancer patients benefit from these agents [7]. Therefore, it is necessary to identify novel biomarkers for predicting LUAD patient survival and response to ICI therapy.

A growing body of evidence has demonstrated the importance of the tumor microenvironment (TME) in oncogenesis and tumors [8]. The TME is a complex network composed of tumor cells, immune cells, mesenchymal stem cells, fibroblast cells, endothelial cells, inflammatory mediators, and extracellular matrix [9]. The interactions between tumor cells and their surrounding supporting cells contribute to the malignant biological behaviors of cancer, such as unlimited proliferation, resistance to apoptosis, and evasion of immune surveillance [10]. Therefore, the TME significantly affects the therapeutic response to and clinical outcomes of patients with cancer. The major non-tumor components of the TME, tumor-infiltrating immune cells and stromal cells have been proposed to be valuable for the diagnostic and prognostic assessment of patients with tumors [11, 12]. The development of a comprehensive model of the TME based on immune and stromal signature genes may contribute to the prognostic evaluation of LUAD patients and predict the efficacy of immunotherapy. With advancements in sequencing techniques, bioinformatics tools such as Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data (ESTIMATE) and CIBERSORT, make it feasible to estimate the distribution of immune and stromal cells in the TME by analyzing specific gene expression signatures of immune and stromal cells [13]. This algorithm has been successfully applied to quantitative analysis of the TME in various tumors and the identification of immune and stromal genes involved in the TME, and its effectiveness has been proven [14].

To date, several predictive models have been constructed for LUAD prognosis stratification, mainly focusing on immune-related genes or immune cells [15]. However, few studies have investigated the influence of TME on LUAD patient survival outcomes and response to immunotherapy, specifically based on immune and stromal components. To fill this knowledge gap, we aimed to develop a TME prognostic model based on immune and stromal genes to predict the survival outcomes and immune responses. In the present study, we systematically investigated the expression details and clinical significance of immune and stromal genes in the TME of LUAD and developed a novel TME-related prognostic model. In addition, we validated this model using independent datasets and analyzed its potential prognostic mechanism and association with immunotherapy responses. Our findings provide promising biomarkers for the prognostic stratification and selection of patients responsive to ICIs, which would facilitate accurate management and appropriate personalized therapies for patients with LUAD.

Materials and methods

Data source and preprocessing

The gene expression profiles of 594 LUAD case were downloaded from the TCGA database (https://portal.gdc.cancer.gov/), along with their corresponding clinical and survival data. Datasets GSE26939, GSE37745, and GSE29016, which contained microarray expression data and clinical information of LUAD patients, were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). In our study, data from TCGA were used as the training cohort, whereas the three GEO datasets were used as validation cohorts.

Generation of the immune score and stromal score

ESTIMATE is an algorithm for estimating the infiltration of immune and stromal cells in tumor samples by analyzing the specific gene expression signatures of immune and stromal cells. Here, we calculated the immune and stromal scores to predict the proportion of immune and stromal components in each sample using the ESTIMATE algorithm with the aid of the R software estimate package.

Identification of differentially expressed genes related to the TME and functional enrichment analysis

All patients in the training cohort were divided into high and low immune/stromal score groups according to the median immune and stromal scores, respectively. Kaplan–Meier analysis was conducted to compare the survival difference between the two groups, and the p-value of the log-rank test was calculated. The limma package was used to identify differentially expressed genes (DEGs) between the high and low immune/stromal score groups with a fold change (FC) =1 and false discovery rate (FDR) < 0.05. DEGs between the high and low immune score groups were defined as immune DEGs, whereas the DEGs between the high and low stromal score groups were defined as stromal DEGs. Finally, the intersecting genes between the immune and stromal DEGs were considered for subsequent analysis. Heatmaps of DEGs were generated to show expression differences using the heatmap package heatmap. Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of intersecting DEGs were performed using the

clusterProfiler, enrichplot, and ggplot2 packages. Only terms with p- and q-values of < 0.05 were considered significantly enriched. Moreover, we downloaded a list of immune-related genes from the Immunology Database and Analysis Portal (Immport) to select immune-related DEGs among these DEGs using the VennDiagram package.

Construction and evaluation of the TME-related prognostic model in the training set

Univariate Cox and Kaplan–Meier analyses were performed in the training cohort to identify significant prognostic genes among the immune-related DEGs. A P value < 0.05 in the log-rank test was considered significant. Multivariate Cox regression analysis was performed to obtain the respective coefficients (βi) of each gene. Finally, the TME prognostic model was constructed on the basis of the key prognostic genes, and the risk score of each patient was calculated on the basis of the expression level of each key prognostic gene and its regression coefficient.

Kaplan-Meier and receiver operating characteristic (ROC) analyses were used to assess the accuracy of the model in predicting clinical outcomes. Univariate and multivariate Cox regression analyses were performed to evaluate the prognostic value of the model and other common clinical factors such as age, sex, stage, and TNM stage.

Validation of the TME-related prognostic model in the testing set

The feasibility and stability of the TME-related prognostic model were confirmed using the GEO database model. Patients in the three testing datasets were divided into the high- and low-risk groups according to the formula of risk score derived from the training dataset using the same methods as above. Kaplan–Meier survival analysis and ROC curve analysis were used to evaluate the performance of the six-gene prognostic model in predicting the outcomes of patients with LUAD.

Evaluation of immune status between the high-risk and low-risk groups stratified by prognostic model

To explore the potential mechanism of the prognostic effects of the model, we analyzed the immune status and pathway enrichment of high-risk and low-risk samples. First, we quantified the enrichment levels of the 29 immune signatures in each LUAD sample using single-sample gene set enrichment analysis (ssGSEA) score. Based on the ssGSEA score, we performed a hierarchical clustering analysis to compare immune activities between the high-risk and low-risk samples. CIBERSORT is an algorithm used for estimating the proportion of immune cell subsets through cell type identification by estimating the relative subsets of RNA transcripts. In this study, we used the CIBERSORT algorithm to construct 21 types of immune cell profiles in LUAD samples and compared the differences in immune cell subtypes between the high- and low-risk groups. KEGG enrichment analysis was performed to analyze the functions or pathways that were upregulated in the two groups. Finally, we compared the mRNA levels of immune checkpoints and their ligands and the expression of HLA genes in the high- and low-risk groups.

Comprehensive analysis of prognostic hub genes in the model

To reveal the regulatory mechanisms of the prognostic hub genes in the TME, we systematically analyzed the genetic alterations and functional enrichment of these six genes. First, RNA expression and gene-encoding protein expression level alterations in LUAD compared with normal tissue were estimated by the Wilcoxon test and immunohistochemistry (IHC), respectively. IHC results for hub genes were obtained from the Human Protein Atlas (HPA) database. The String online database and Cytoscape software were used to construct a protein–protein interaction (PPI) network between the molecules. We then analyzed the pathways of hub genes by gene set enrichment analysis (GSEA), using the gene expression level as the phenotype. The curated KEGG gene set was downloaded from the Molecular Signature Database, and FDR < 0.05 was considered significant. Finally, we evaluated the correlation between hub gene expression and immune cell infiltration in LUAD using TIMER.

Results

Immune scores and stromal scores were correlated with survival outcomes

A total of 510 LUAD cases from TCGA were used as the training cohort, and three GEO datasets were used as the validation cohorts. The clinical information for all cohorts is summarized in Table 1. We calculated the immune and stromal scores of each LUAD patient in TCGA and divided them into high and low immune/stromal score groups on the basis of the median value. Kaplan-Meier survival analysis showed that patients with high immune and stromal scores showed better survival outcomes than those with low scores, with log-rank tests of P = 0.01 and 0.026, respectively (Fig. 1a, 1b).

Identification of DEGs based on immune score and stromal score

The heatmap showed that genes in the high score group had lower expression levels than those in the low score group, both for immune and stromal scores (Figure 1C,

Table 1 Clinical characteristics of LUAD patients included in this study

Features	Number of patients (%)			
	Training cohort TCGA (*n* = 510)	Testing cohort1 GSE26939 (*n* = 116)	Testing cohort2 GSE37745 (*n* = 196)	Testing cohort3 GSE29016 (*n* = 72)
Age (years)				
≤ 65	235 (46.08)	59 (50.86)	102 (52.04)	33 (45.83)
> 65	275 (53.92)	57 (49.14)	94 (47.96)	39 (54.17)
Gender				
Female	271 (53.14)	63 (54.31)	89 (45.41)	41 (56.94)
Male	239 (46.86)	53 (45.69)	107 (54.59)	31 (43.06)
AJCC stage				
Stage I	272 (53.33)	62 (53.45)	130 (66.33)	46 (63.89)
Stage II	124 (24.31)	19 (16.38)	35 (17.85)	15 (20.83)
Stage III	85 (16.67)	19 (16.38)	27 (13.78)	5 (6.95)
Stage IV	22 (4.31)	2 (1.72)	4 (2.04)	0 (0)
Unknown	7 (1.38)	14 (12.07)	0 (0)	6 (8.33)
T stage				
T1	168 (32.94)	–	–	25 (34.72)
T2	276 (54.12)	–	–	36 (0.50)
T3	47 (9.21)	–	–	7 (9.72)
T4	19 (3.73)	–	–	4 (5.56)
N stage				
N0	335 (65.69)	–	–	65 (90.28)
N1–3	175(34.31)	–	–	7 (9.72)
M stage				
M0	349 (68.43)	–	–	68 (94.44)
M1	22 (4.31)	–	–	0 (0)
Unknown	139 (27.26)	–	–	4 (5.56)
Survival status				
Alive 0	184 (36.08)	50 (43.10)	51 (26.02)	22 (30.56)
Dead 1	326 (63.92)	66 (56.90)	145 (73.98)	50 (69.44)

AJCC, American Joint Committee on Cancer

D). A total of 776 immune DEGs were obtained from the comparison of immune scores (samples with high scores vs. low scores), of which 613 genes were upregulated and 163 genes were downregulated. Similarly, 792 stromal DEGs were obtained from a comparison of the stromal scores, consisting of 678 upregulated genes and 114 downregulated genes. Moreover, Venn diagrams showed that 297 DEGs were commonly upregulated in the high-score groups, and 66 DEGs were commonly downregulated (Fig. 1e, 1f). These notable DEGs were potentially determinant factors of TME status.

GO and KEGG enrichment analysis

Results of the GO enrichment analysis showed that these DEGs were mainly involved in immune-related functions, such as T-cell activation and lymphocyte proliferation (Fig. 2a, 2c). KEGG analysis also revealed enrichment of the T cell receptor signaling pathway, chemokine signaling pathway, and hematopoietic cell lineage (Fig. 2b, 2d). Since the DEGs were correlated with immune functions or pathways in LUAD, we further identified the top 89 immune-related DEGs from the Immport database for subsequent analysis (Fig. 3a).

Construction of the TME-related prognostic model

Univariate Cox and Kaplan–Meier analyses were conducted to determine the significant prognostic genes among the 89 immune-related DEGs. A total of 24 genes were identified as significant in the Cox regression analysis (Fig. 3b), of which 6 genes were also significant in the Kaplan-Meier analysis. Among them, higher expression levels of BTK, CD28, PIK3CG, TLR4, and VEGFD correlated positively with poor survival outcomes, whereas INHA expression correlated negatively with prognosis (Fig. 3c). Then, the six prognostic genes were subjected to multivariate Cox regression analysis, and the risk coefficient of each gene was calculated (Table 2). The TME-related prognostic model was constructed as follows: risk score = (–0.111958) × EXP_{BTK} + 0.279096 ×

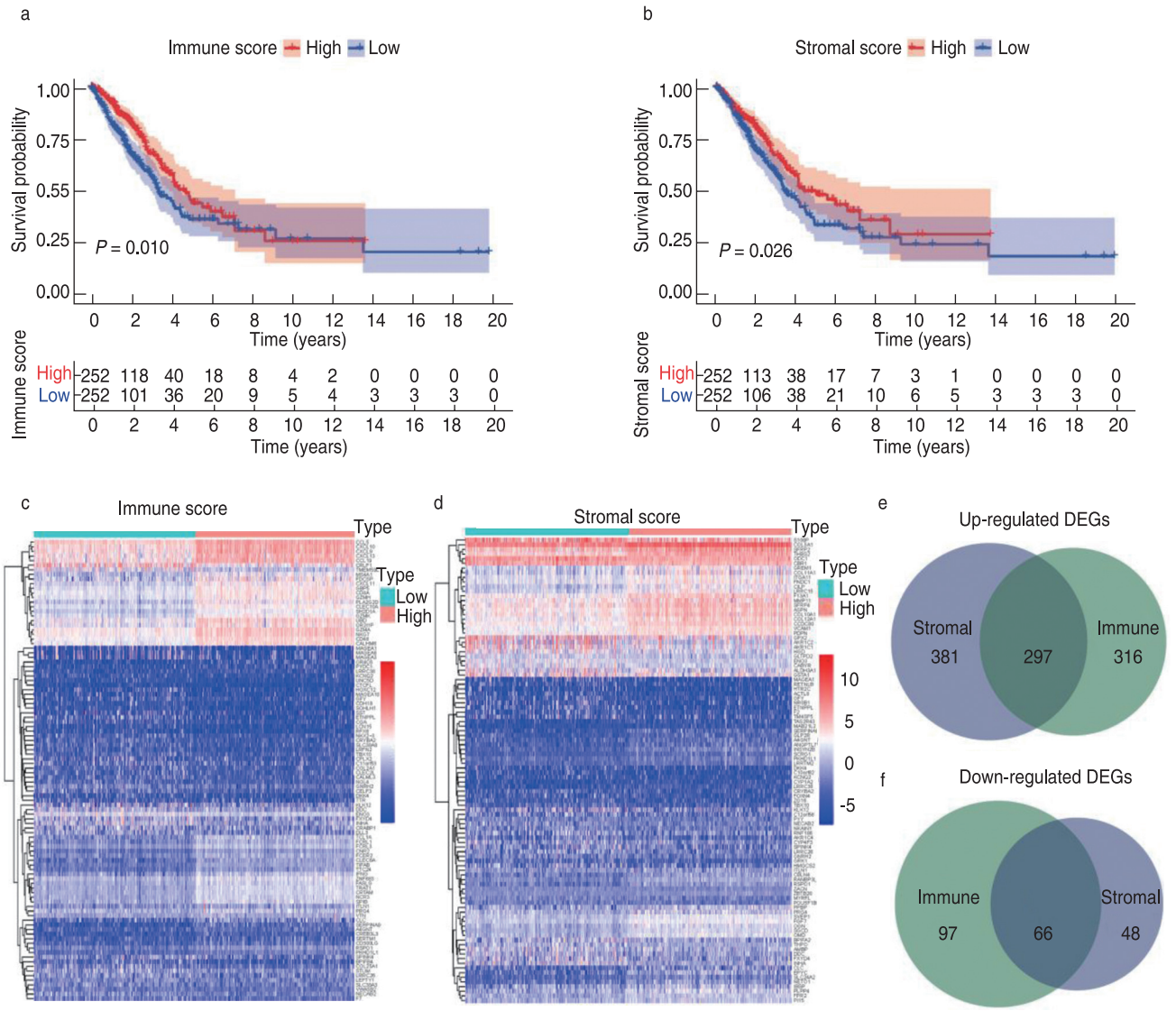


Fig. 1 Identification of differential expressed genes(DEGs). (a) Kaplan-Meier survival curve of high and low immune score groups; (b) Kaplan-Meier survival analysis for high and low stromal score groups; (c) Heatmap for DEGs generated by comparison of gene expression profiles in high and low immune score groups; (d) Heatmap for DEGs in high and low stromal score groups; (e, f) Venn diagrams showed the common up-regulated and down-regulated DEGs shared by immune and stromal score groups

EXP_{CD28} + 0.008079 × EXP_{INHA} + (–0.357674) × EXP_{PIK3CG} +0.099561 × EXP_{TLR4} + (–0.102261) × EXP_{VEGFD}.

Prognostic value of the TME-related model in the training and validation cohorts

We calculated the risk score for each patient in the training cohort (n = 510) and divided them into the high- and low-risk groups according to the median cutoff value (cutoff value: –0.261). The Kaplan-Meier plot showed that patients in the high-risk group had worse survival outcomes than those in the low-risk group (Fig. 4a). The ROC curve of the 5-year survival prediction was drawn to assess the predictive accuracy, with an area under the curve of 0.688 (Fig. 4b). Additionally, the risk curve indicated that the high-risk group had a higher mortality and worse prognosis than the low-risk group (Fig. 4c, 4d). The prognostic value of our model in patients with LUAD was further evaluated using other common prognostic factors. Although univariate Cox analysis indicated that the pathological stage and risk score had prognostic effects, only the risk score could be used as an independent prognostic factor (P < 0.001; Fig. 4e, 4f).

Consistent with the results in the training dataset, the six-gene model stratified the samples of the three GEO testing datasets into high-risk and low-risk groups. Patients with low-risk scores had better survival

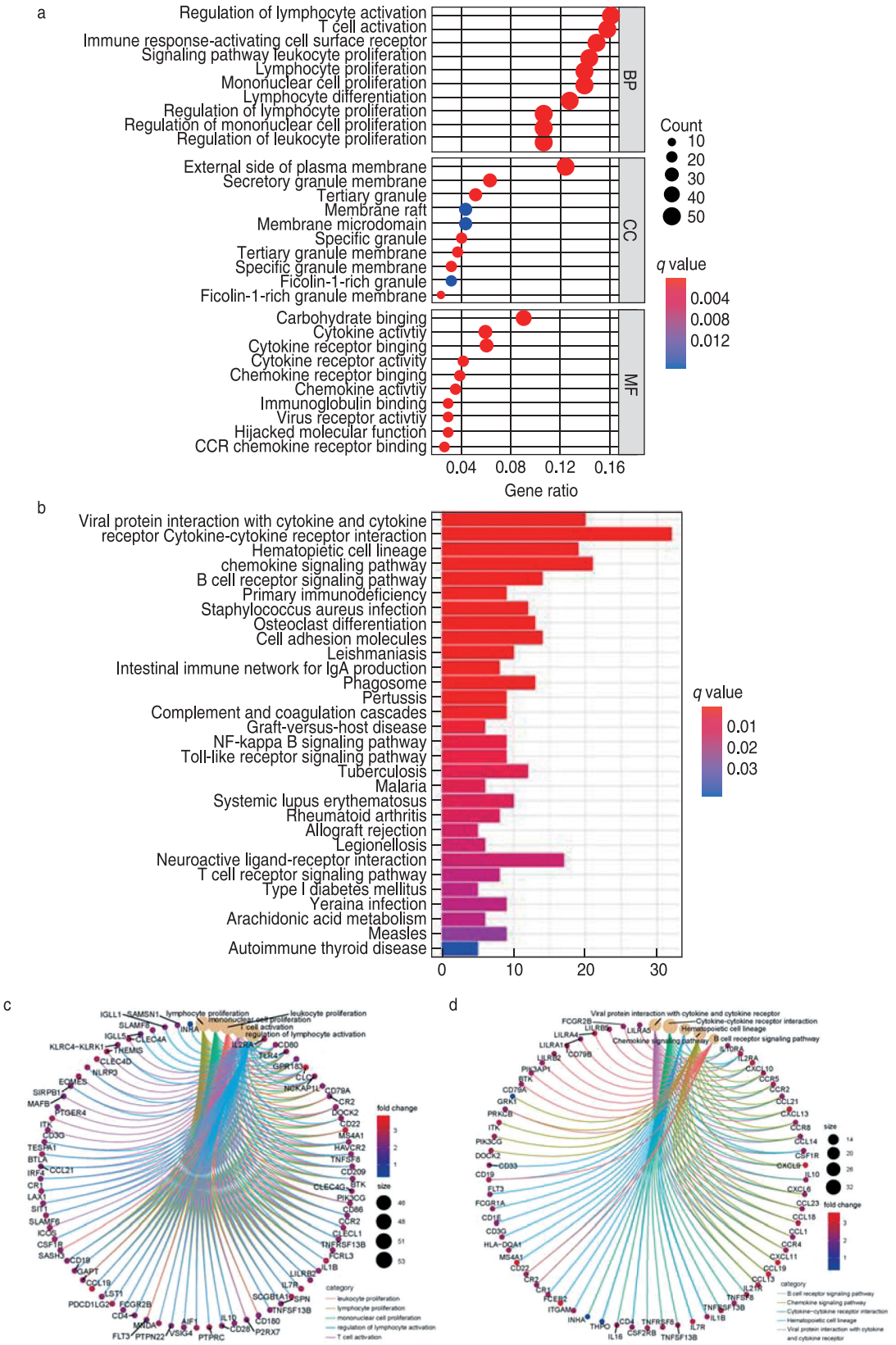


Fig. 2 Functional enrichment analysis. (a) GO enrichment analysis of 363 DEGs; (b) KEGG enrichment analysis of DEGs; (c)Circle plot show the DEGs involved in top 5 enriched terms of GO analysis; (d)Circle plot show the DEGs involved in top 5 enriched terms of KEGG analysis

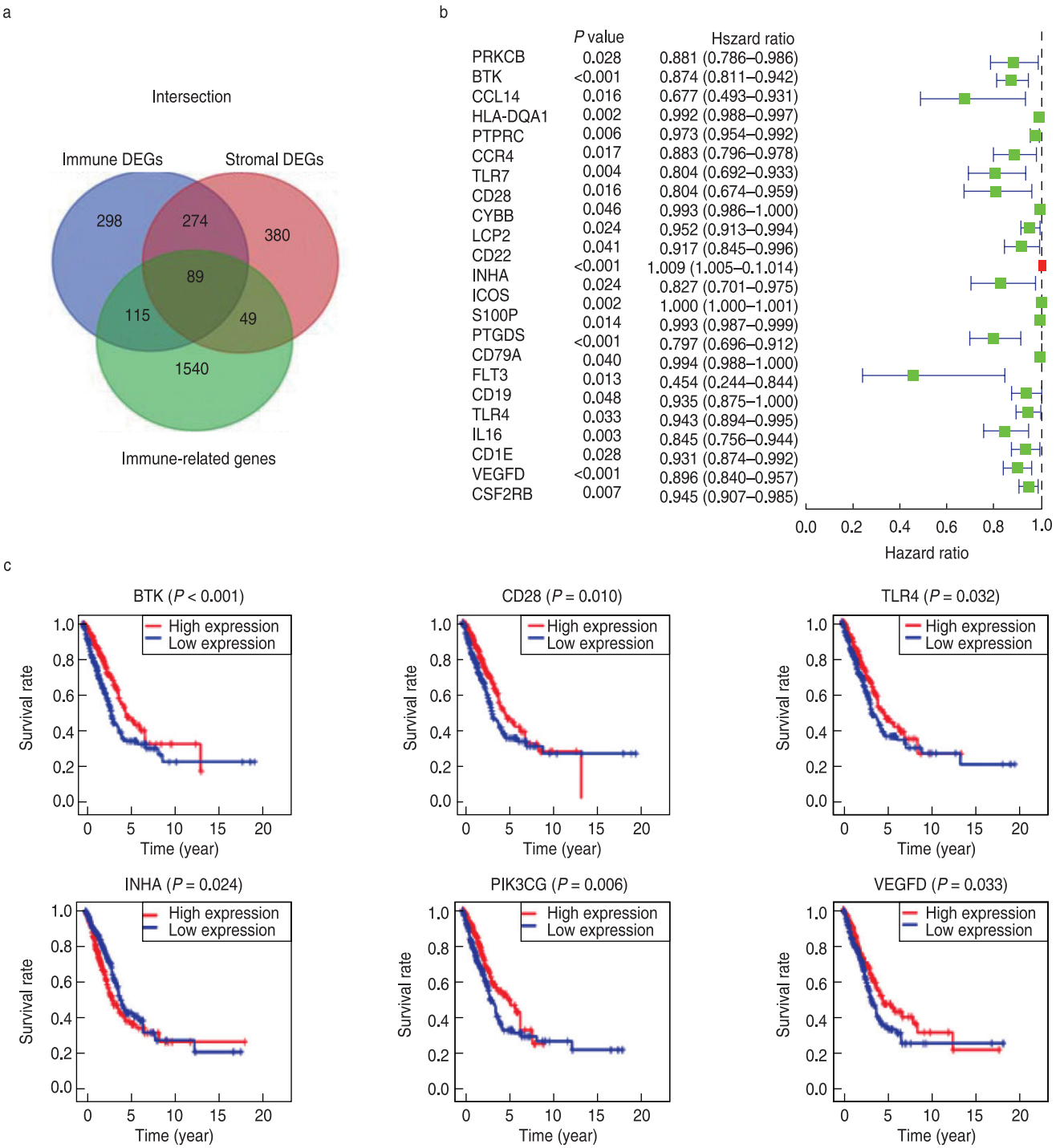


Fig. 3 Univariate Cox and Kaplan Meier analysis for prognostic genes screening. (a) Identification of immune-related DEGs; (b) The forest plot of 24 prognostic immune-related DEGs screened out by Univariate Cox regression analysis with $P < 0.005$; (c) Survival curves of the 6 prognostic genes extracted by the Kaplan-Meier analysis. Patients were labeled with high expression or low expression according to the median expression level of the 6 genes

outcomes than those in the high-risk group ($P < 0.05$; Fig. 5a–5c). The areas under the ROC curves for predicting 5-year survival in the three testing datasets were 0.679, 0.666, and 0.732 (Fig. 5d–5f). These results suggest that

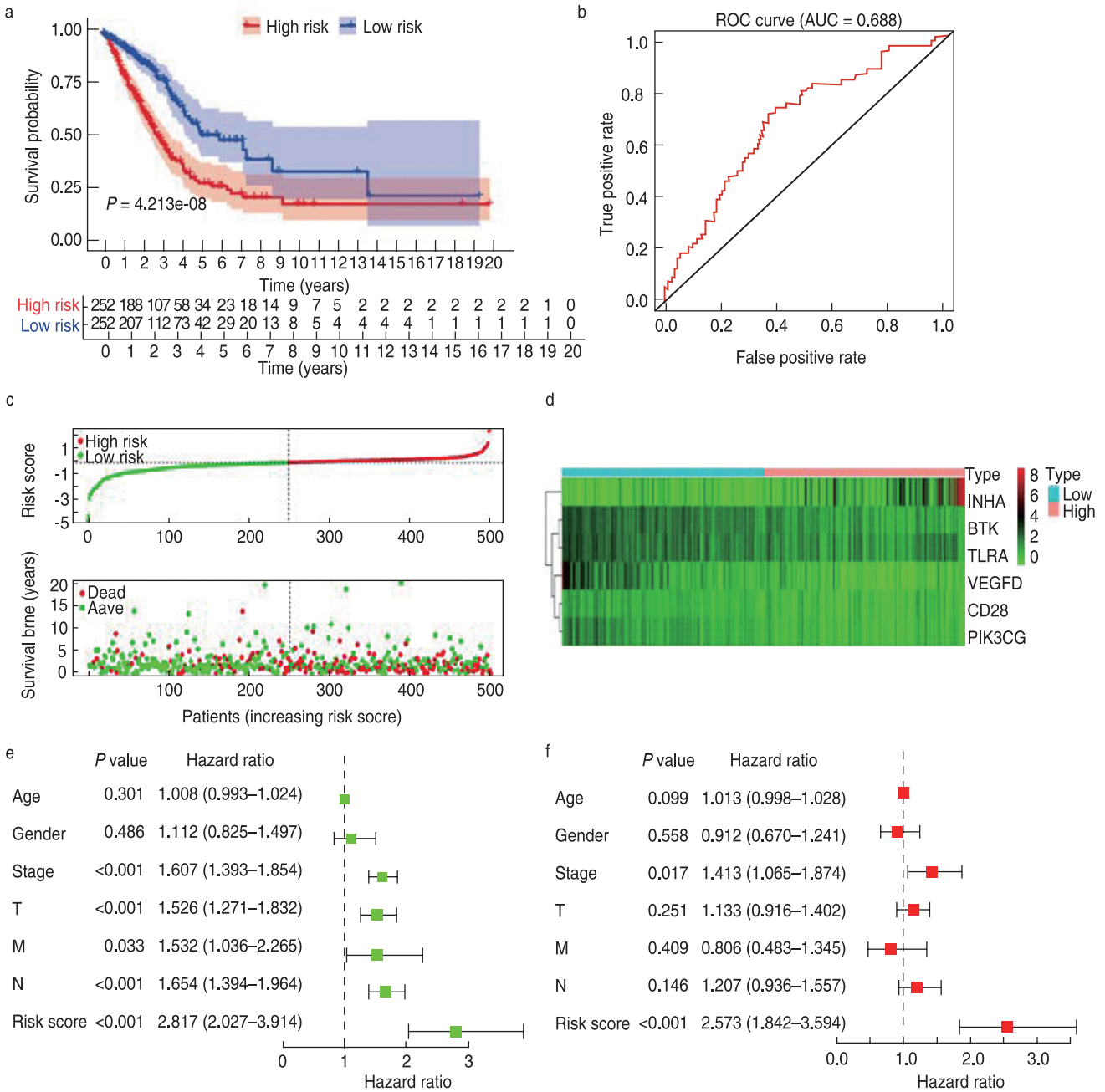


Fig. 4 Construction and validation of the TME-related prognostic model in the training cohort. (a) Kaplan-Meier survival curve of low- and high-risk groups stratified by the TME-related prognostic model; (b) The ROC analysis of the TCGA dataset for survival prediction by the TME prognostic model; (c) The distribution of risk score and survival time in high- and low-risk groups; (d) Heatmap of the six prognostic genes; (e)The Univariate Cox analysis evaluating the prognostic effect of the model and common clinical factors; (f) Multivariate Cox analysis evaluating independent predictive ability of our model for overall survival

the TME-related prognostic model is robust in predicting the survival outcomes of patients with LUAD.

Evaluation of the immune status between low-risk and high-risk groups

The strong stratification power of the TME-related model in predicting the survival of patients with LUAD led us to explore the difference in functional characteristics between the two risk groups. The ssGSEA score of the

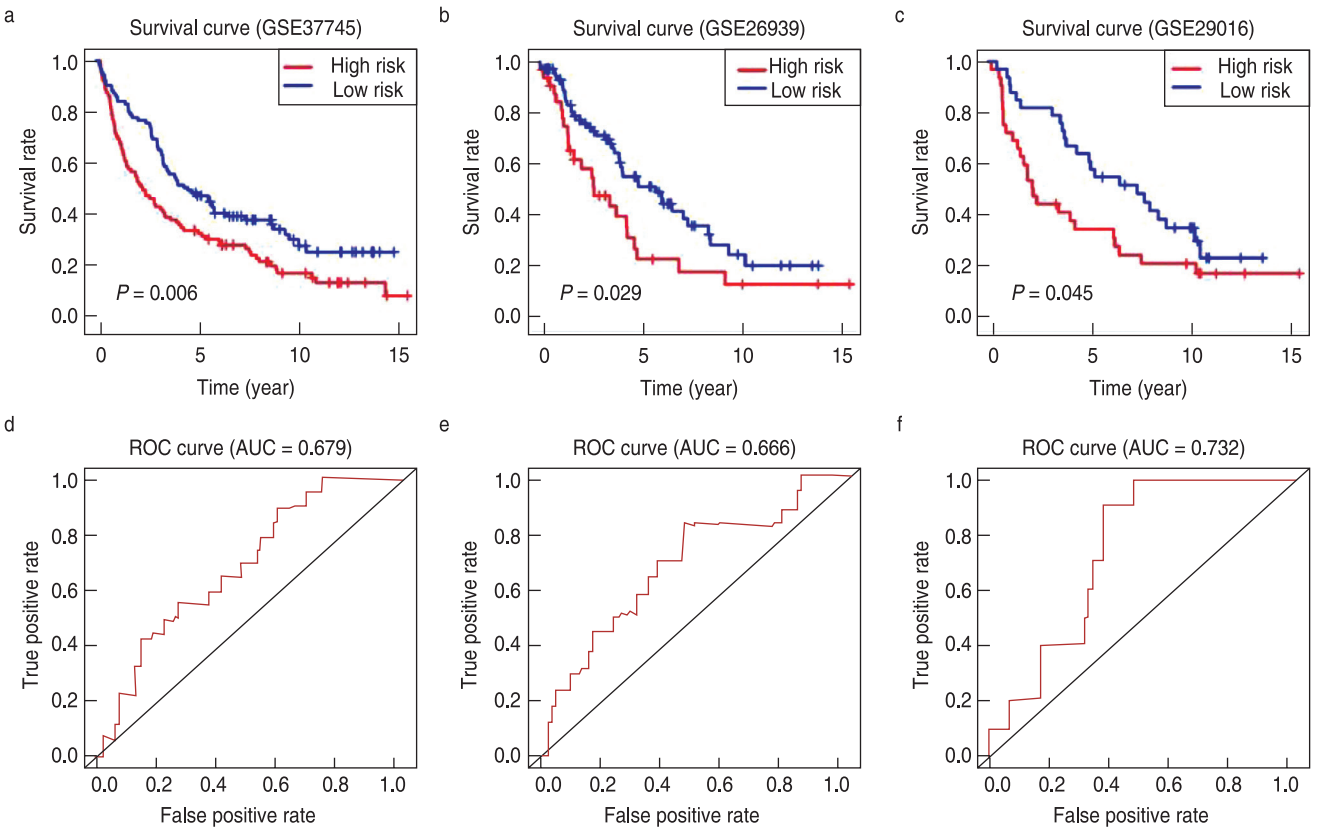


Fig. 5 Validation of the TME-related prognostic model in the testing cohort. Kaplan-Meier survival curves showing overall survival outcomes of high- and low-risk groups in GSE37745 (a), GSE26939 (b) and GSE29016 (c). The ROC curves for judging the predictive accuracy of the model in GSE37745 (d), GSE26939 (e) and GSE29016 (f)

Table 2 Genes in the TME-related prognostic model

Gene symbol	Gene description	Coefficient
BTK	Bruton tyrosine kinase	–0.111958
CD28	CD28 molecule	0.279096
INHA	Inhibition alpha subunit	0.008079
PIK3CG	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma	–0.357674
TLR4	Toll like receptor 4	0.099561
VEGFD	Vascular endothelial growth factor D	–0.102261

29 immune signatures was used to evaluate the immune status of the two groups. The heatmap showed that the low-risk group had a higher immune activity than the high-risk group (Fig. 6a). Consistent with the ssGSEA results, immune and stromal scores in the low-risk group were significantly higher, and the tumor purity of the low-risk group was significantly lower (Fig. 6b–6d). This finding indicated that more immune and stromal cells infiltrated the TME of low-risk samples, whereas more tumor cells were present in high-risk samples. Moreover, we identified the immune cell subtypes in the two groups. The low-risk group had a significantly higher proportion of memory B cells, memory CD4+ T cells, monocytes, and dendritic cells than the high-risk group, whereas the high-risk group had a markedly higher proportion of M0 macrophages (Fig. 6e). Taken together, these results suggest that patients with low-risk scores show elevated antitumor immune activity, leading to more favorable clinical outcomes.

Potential mechanisms of the prognostic effects of the TME-related model

GSEA was conducted to elucidate the specific regulatory mechanisms resulting in the differences in prognosis and immune status between the two risk cohorts. The results showed that the low-risk group was

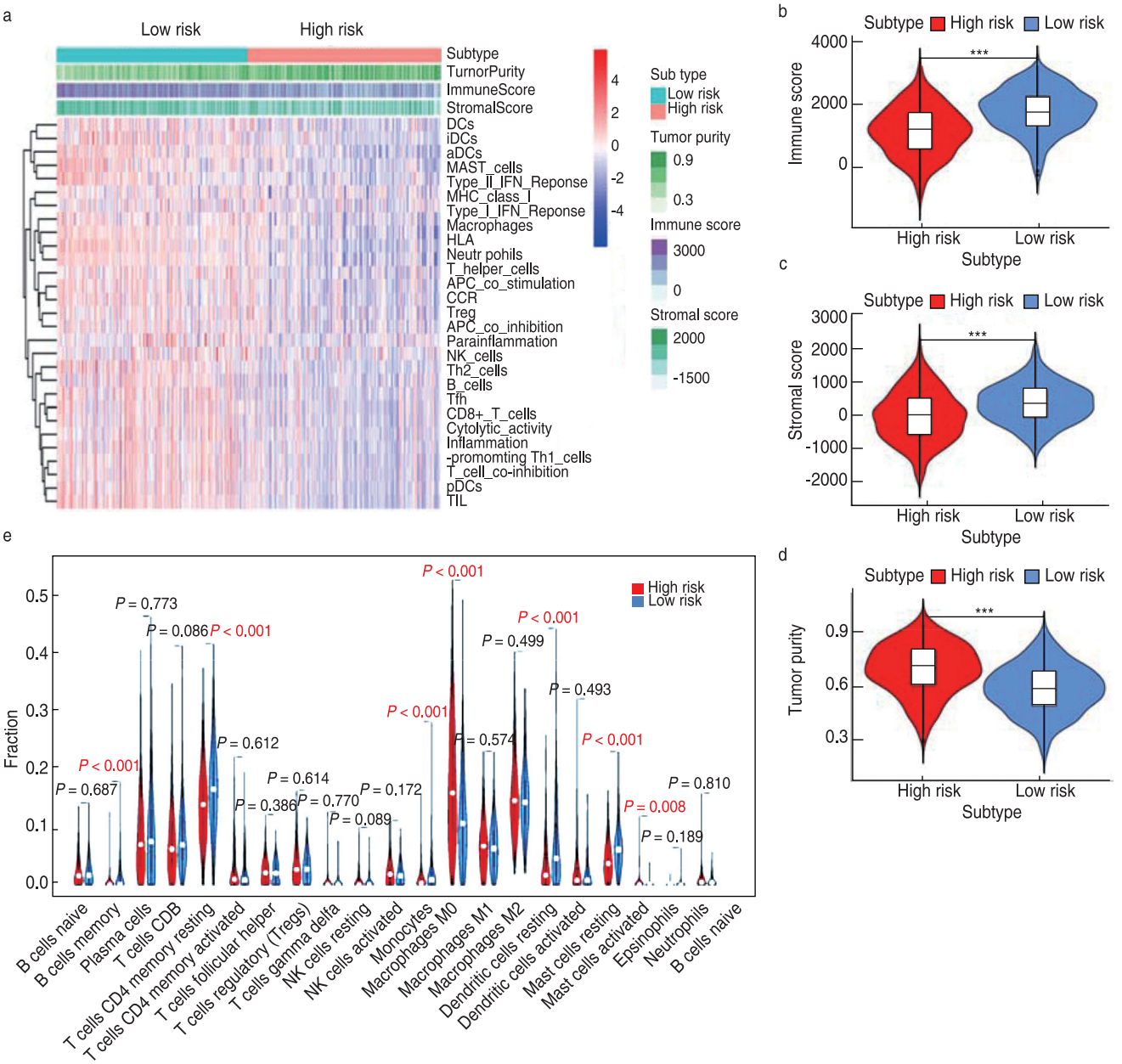


Fig. 6 Evaluation of immune status and immune cell infiltration levels between high- and low- risk groups. (a) The heatmap of the overall immune status of high- and low-risk groups in TCGA database, showing greater heterogeneity between the two groups; (b–d) The violin plots showed the difference in immune score, stromal score and tumor purity between low- and high-risk groups. ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$; (e) The violin plot shows the difference in the proportion of 21 kinds of immune cells between high- and low-risk groups, and the Wilcox rank-sum was used for the significance test

enriched not only in immunoregulation and immune cell activation, but also in many cancer-associated pathways, such as JAK-STAT signaling, cell adhesion molecules, and transendothelial migration (Fig. 7a). In contrast, the high-risk group was impoverished in immune signatures but enriched in metabolic signaling. Notably, most HLA genes showed significantly higher expression in the low-risk group than in the high-risk group, indicating that local immune regulation and immunogenicity were more active in the low-risk group (Fig. 7b).

Relationship between the TME-related model and immunotherapy response

In recent years, immune checkpoint proteins such as cytotoxic T lymphocyte antigen 4 (CTLA-4) or the programmed cell death ligand 1/protein 1 pathway

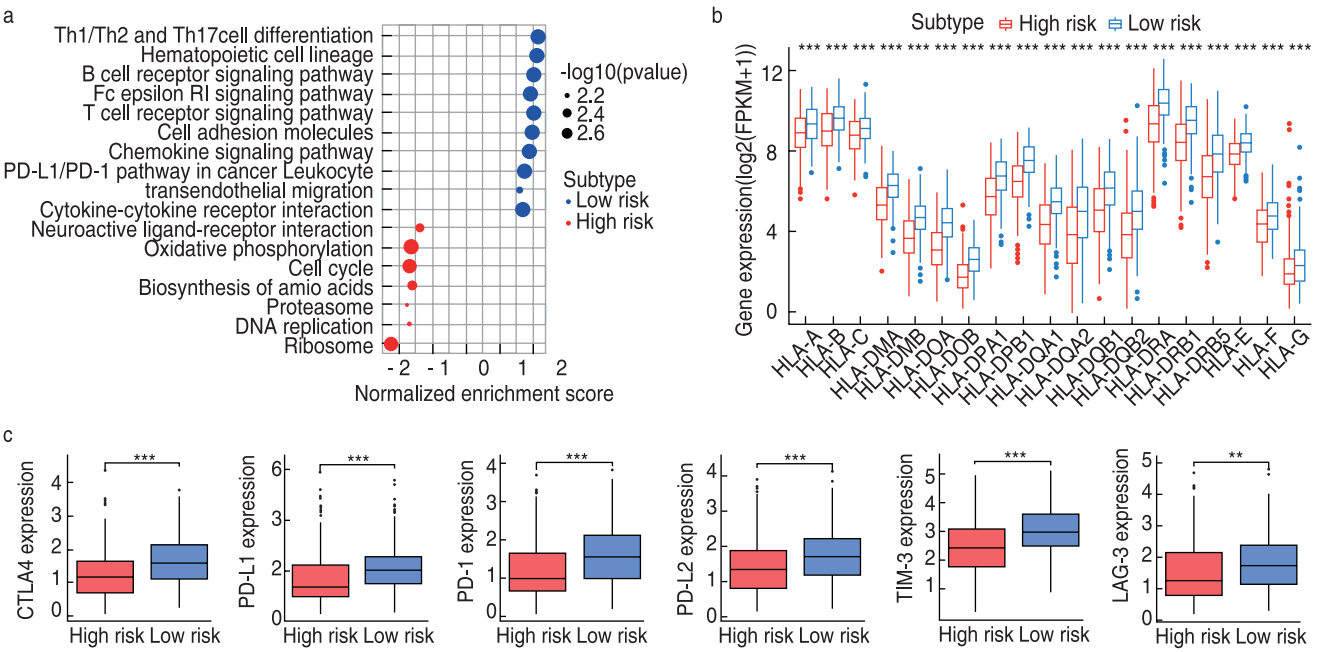


Fig. 7 Functional mechanisms of the TME-related model and association with immune checkpoint molecules. (a) KEGG pathways enriched in high- and low-risk samples; (b) The expression profiles of HLA genes of low- and high-risk groups. ***$P < 0.001$; **$P < 0.01$; *$P < 0.05$; (c) Comparison of expression levels of CTLA-4, PD-1,PD-L1, PD-L2,TIM-3 and LAG-3 between high-risk and low-risk groups (Wilcox rank-sum test)

(PD-L1/PD-1) have been used as crucial targets for immunotherapy in LUAD [16]. We explored the relationship between the model and immunotherapy response by analyzing the expression of common immune checkpoints in the high- and low-risk groups and found that the expression of CTLA-4, PD-1, PD-L1, PD-L2, TIM-3, and LAG-3 in the low-risk group was significantly higher than that in the high-risk group (all $P < 0.001$; Fig. 7c). This suggests that patients with low-risk scores might respond better to ICI treatment than those with high-risk scores because the expression of immune checkpoint molecules tends to be positively associated with immunotherapeutic responsiveness. Therefore, the construction of a risk cohort using our model could be a good stratification method for patients with LUAD regarding whether to conduct immunotherapy.

The mechanism of action of prognostic hub genes in the TME

To further analyze the potential function of the hub genes, our results were verified using the HPA and TIMER databases. We found that CD28 and INHA had significantly higher expression levels in LUAD samples than in normal lung samples, whereas BTK, PIK3CG, TLR4, and VEGFD had lower expression levels in tumor tissues (Fig. 8a). In terms of protein levels, the protein expression patterns of BTK, INHA, PIK3CG, TLR4, and VEGFD were consistent with their RNA-seq expression alterations (Fig. 8b). However, CD28 showed no significant difference. The PPI network also showed extensive interactions among BTK, CD28, PIK3CG, and TLR4.

GSEA suggested that BTK, CD28, PIK3CG, TLR4, and VEGFD were enriched in the same pathways. High expression of these five genes was mainly correlated with immune-related activities, such as antigen processing and presentation, the chemokine signaling pathway, and the JAK-STAT signaling pathway, whereas low expression of these genes was associated with metabolic pathways (Fig. 9a, 9b). In contrast, high expression of INHA was correlated with metabolic pathways, and low expression was involved in the activation of immune pathways (Fig. 9c, 9d). More importantly, the expression of BTK, CD28, PIK3CG, TLR4, and VEGFD positively correlated with the infiltration of CD4+ T cells, CD8+ T cells, B cells, neutrophils, dendritic cells, and macrophages (Fig. 10). However, INHA was negatively correlated with infiltration of the six immune cells. Collectively, these results suggest that these six hub genes affect the immune activity of the TME.

Discussion

In this study, we aimed to identify a novel TME-related prognostic model for LUAD. We embarked on TME-related DEGs generated by comparing the immune

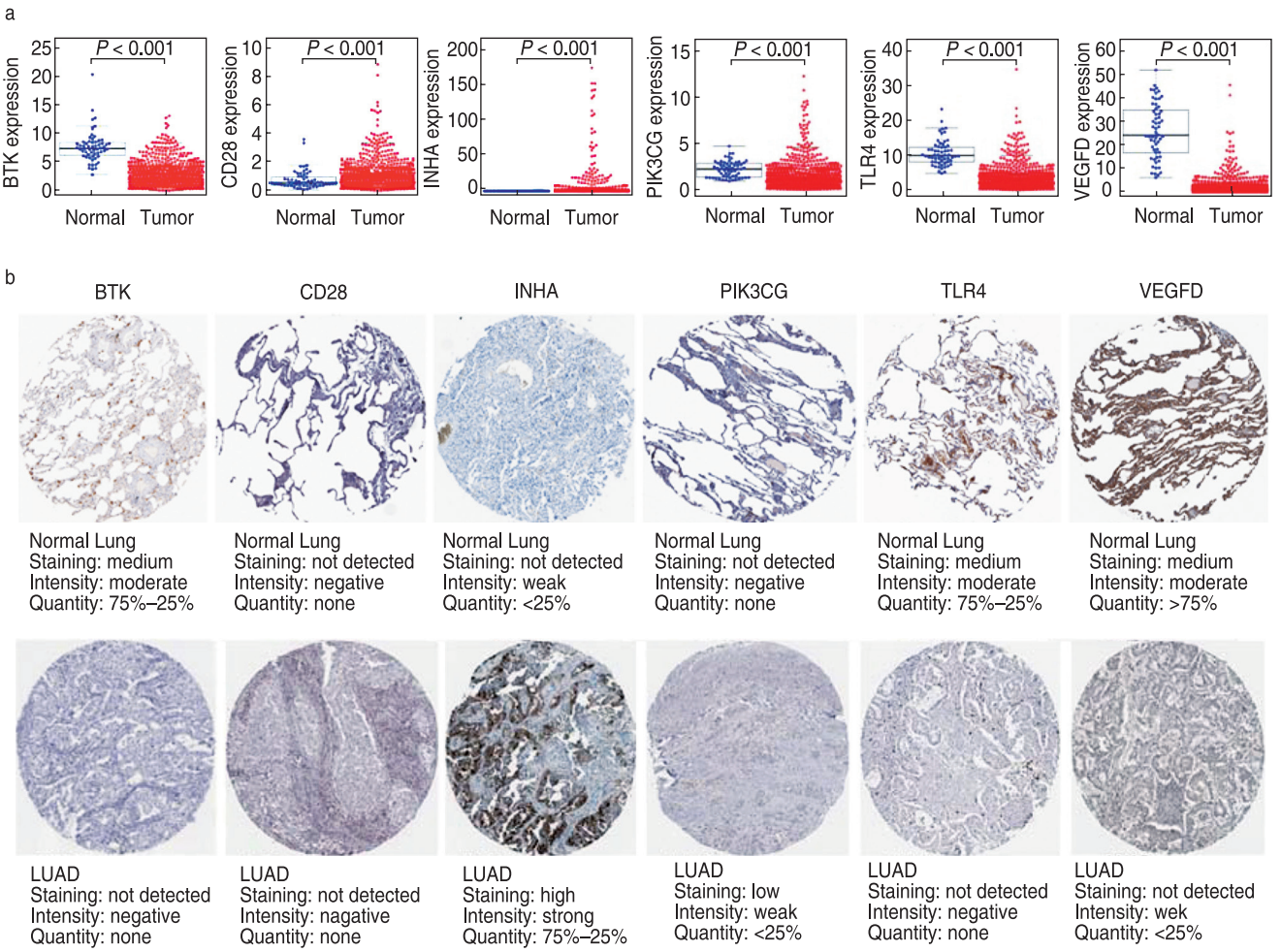


Fig. 8 Expression profiles of the six hub genes in the model. (a) The expression levels of the six prognostic genes in LUAD samples and normal lung samples in the TCGA database (BTK, CD28, INHA, PIK3CG, TLR4 and VEGFD). Wilcox test was used to calculate the significance level between the two groups; (b) The immunohistochemistry results reflecting the gene-encoding protein levels of the six hub genes in LUAD and normal lung tissues from the HPA database

and stromal scores in LUAD samples. Subsequently, a list of TME-related genes that contribute to the survival outcomes of patients with LUAD was extracted. Finally, a six-gene prognostic model, based on prognostic TME-related genes, was constructed. Both immune and stromal genes in LUAD samples were analyzed to better reflect the complete TME status. Furthermore, we validated its prognostic value in three testing sets from the GEO datasets. Kaplan-Meier and ROC analyses revealed the strong predictive ability of our model for LUAD prognosis in both the training and testing sets. Univariate and multivariate Cox analyses confirmed the independent prognostic value of the six-gene model. Accordingly, unlike those developed previously, the TME-related prognostic model developed herein could reflect the tumor immune microenvironment status and predict the prognosis of LAUD more accurately. Moreover, an enhanced understanding of the model and related hub genes would help to elucidate regulatory mechanisms of the TME and develop new treatment strategies.

Experimental and clinical studies have demonstrated that the immune and stromal components of the TME play significant roles in lung cancer development and progression [17]. The immune and stromal cells infiltrating the TME are composed of different cell types. As the most important immune cells in the TME, tumor-infiltrating T lymphocytes execute key effector cytotoxic functions and mediate responses to ICIs [18]. Tumor-associated macrophages are another class of immune cells that interact with lung cancer cells. Macrophage-tumor cell interactions lead to the release of pro-inflammatory cytokines, chemokines, and growth factors, which in turn recruit additional inflammatory cells to the microenvironment [19]. Other immune cells in the TME include B cells, NK cells, dendritic cells (DCs), T regulatory cells (Tregs), and B regulatory cells (Bregs). Cancer-associated fibroblasts

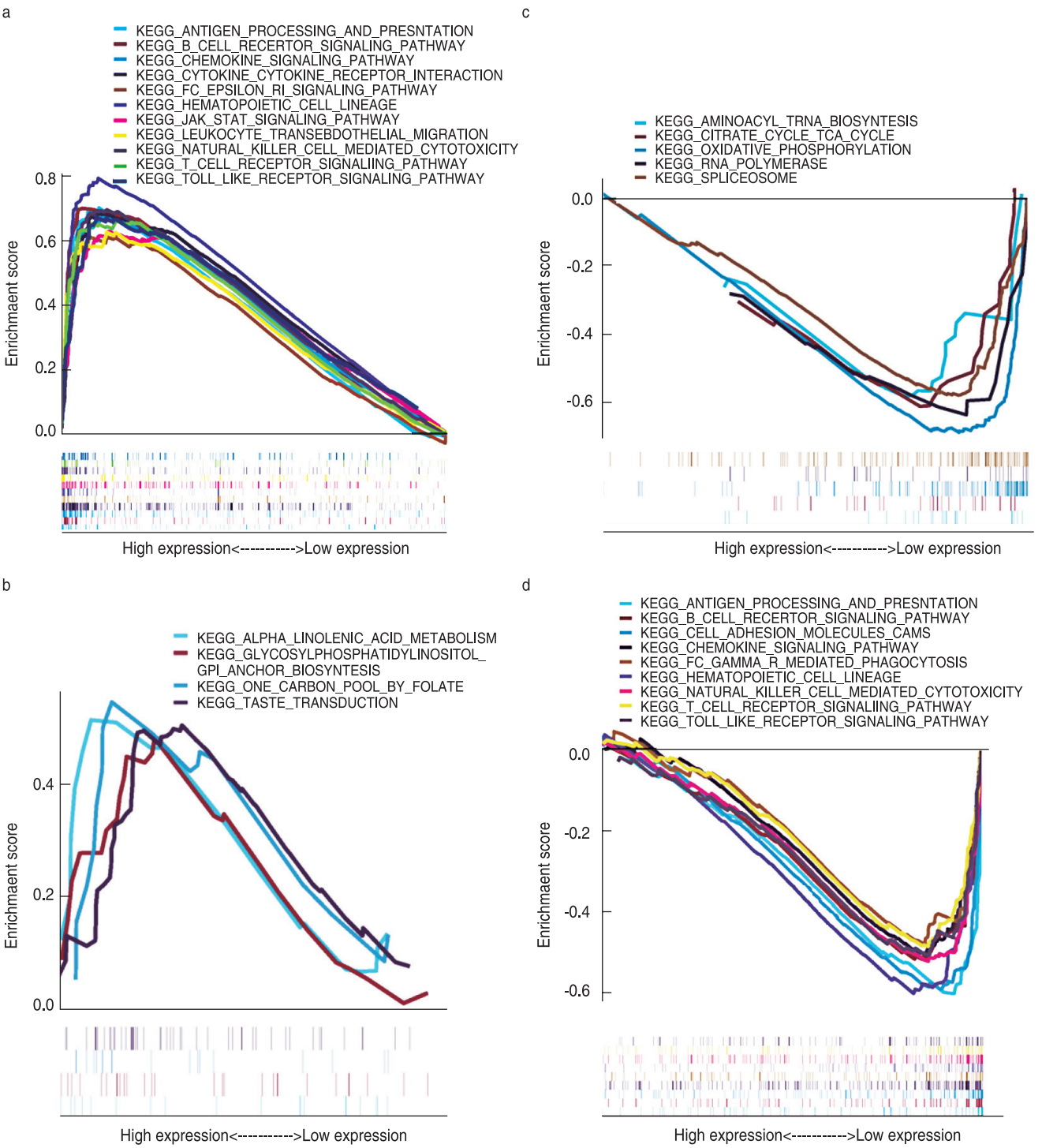


Fig. 9 The GSEA enrichment analysis of the hub prognostic genes. (a) The enrichment pathways of high expression of BTK, CD28, PIK3CG, TLR4 and VEGFD. Each line representing one particular pathway with unique color, only pathways with p and $q < 0.05$ were considered significant. And only several leading gene sets were displayed in the plot; (b) The enrichment pathways of low expression of the five genes; (c) Enrichment pathways of high INHA expression; (d) Enrichment pathways of the low INHA expression

are the most abundant stromal cells in the TME and play critical roles in the inflammatory response and immune suppression of tumors [20]. Fibroblasts promote tumor progression via multiple pathways, including regulation of the extracellular matrix, production of growth factors or cytokines, and promotion of angiogenesis, whereas

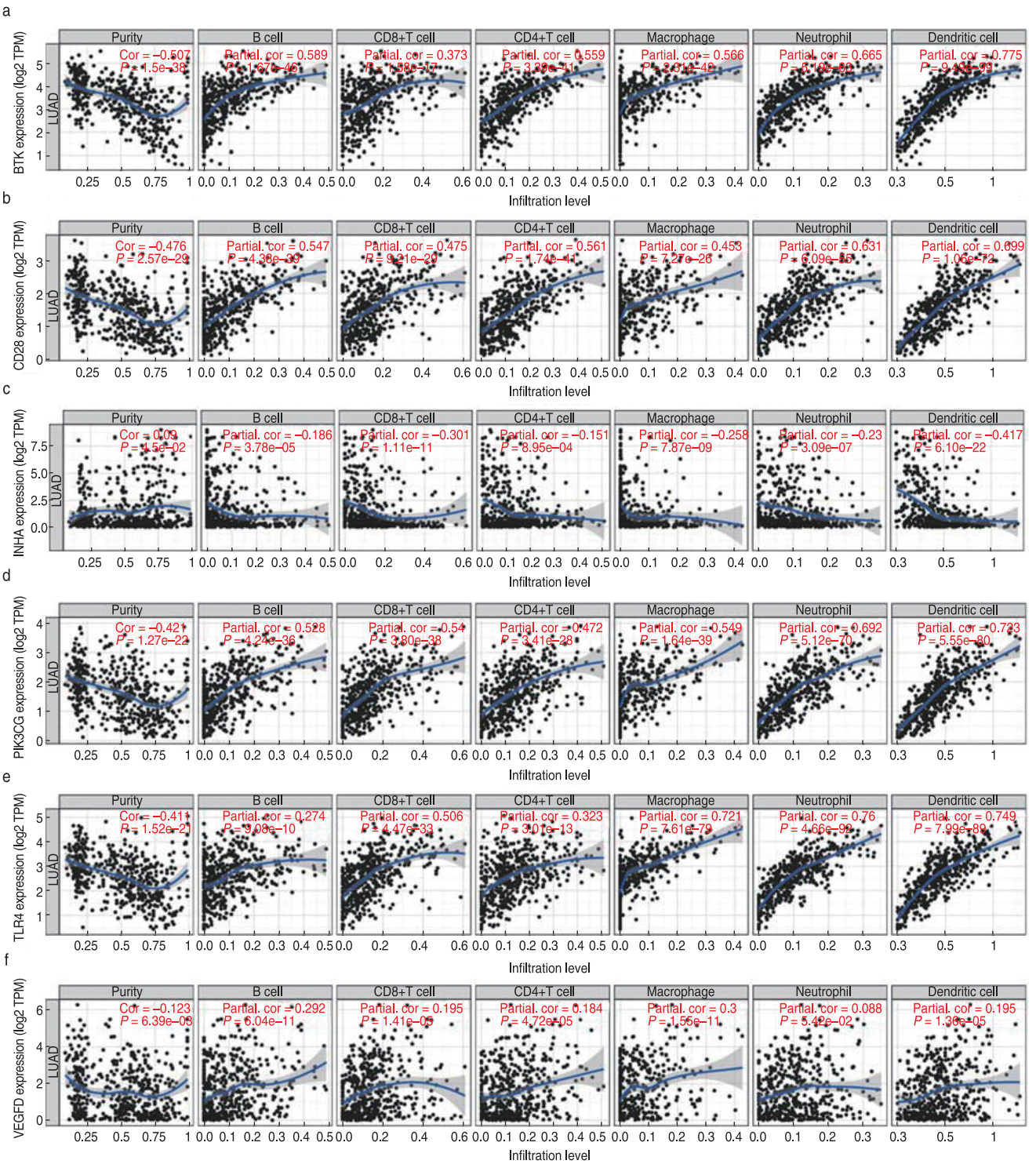


Fig. 10 Correlation between hub prognostic genes and immune cell infiltration. (a–f) The gene expression levels against tumor purity are displayed on the left-most panel

some fibroblast subtypes also show antitumor activities by secreting immunosuppressive cytokines [21].

To determine the distinct gene expression profiles in the TME with respect to immune and stromal components, DEGs based on immune and stromal scores were screened. Six prognostic hub genes among these

DEGs were identified (survival positive correlation: INHA; negative correlation: BTK, CD28, PIK3CG, TLR4, and VEGFD). Interestingly, functional analysis showed that BTK, CD28, PIK3CG, TLR4, and VEGFD could promote immune infiltration, while INHA inhibited immune cell infiltration. These results suggest that the six hub genes participate in the immune regulation of the TME and affect the prognosis of patients with LUAD, which might be potential immune prognostic markers and therapeutic targets for LUAD.

Among these six genes, Bruton tyrosine kinase (BTK) is a member of the Tec kinase family. As a key component of the B-cell antigen receptor signaling pathway, BTK plays a vital role in B lymphocyte development, differentiation, and signaling [22]. Ibrutinib is a small-molecule irreversible inhibitor of BTK that has been approved for the treatment of hematological malignancies and some solid tumors owing to its ability to inhibit tumor growth by modifying the tumor microenvironment and its potential synergistic activity with ICIs [23, 24]. CD28 is a key T-cell costimulatory molecule that binds to B7 molecules, which are involved in the regulation of T cell proliferation and activation, along with cytokine production [25]. Previous studies have demonstrated that CD28 can predict the response to anti-PD-1 therapy in patients with lung cancer [26]. PIK3CG encodes the PI3Kγ enzyme, which can activate the signaling molecule Akt and modulate various cell functions such as cell proliferation, migration, and adhesion [27]. Novel PI3K inhibitors are important for the treatment of hematologic malignancies [28]. The protein encoded by TLR4 is a member of the toll-like receptor family. Studies have shown that TLR4 is highly expressed in immune cells, such as monocytes and lymphocytes, but is expressed at low levels in epithelial, endothelial, and cancer cells. Thus, TLR4 agonists have been widely explored as potential immunotherapeutic agents for the treatment of cancer [29]. VEGFD belongs to the vascular endothelial growth factor family and can induce both angiogenesis and lymphangiogenesis [30]. Clinical studies have shown that low expression of VEGFD is a predictor of greater survival benefits from bevacizumab treatment in patients with CRC [31]. INHA encodes a member of the transforming growth factor-beta (TGF-beta) superfamily of proteins. However, its function in lung cancer remains unknown. Our results suggest that high expression of INHA is associated with a poor prognosis of LUAD. A possible mechanism might be that INHA is involved in vascularization and tumor metastasis, leading to a poor prognosis [32]. Further studies are required to clarify the role of these hub genes in the TME in the initiation and development of LUAD.

Finally, a TME-related prognostic model was developed using the six hub immune genes for survival prediction. The low-risk group showed higher expression of HLA genes. HLA-related genes play a significant role in immune regulation, and their expression is advantageous for immunotherapy efficacy [33]. Our results showed elevated antitumor immune activity in the low-risk group, which could explain why the low-risk group had more favorable clinical outcomes than the high-risk group. In addition to survival prediction, this TME-related signature was also a predictor of patient response to ICI treatment. To date, many biomarkers have been verified to indicate the efficacy of ICI treatment, including TMB, PD-L1 expression level, neoantigens, and gut microbiota [34]. Generally, most biomarkers reflect the status of the tumor immune microenvironment in a certain aspect. Thus, a prognostic model based on the TME may aid in the stratification of patients with LUAD to identify those responsive to immunotherapy. It is possible that patients with low-risk scores are more sensitive to immunotherapy than those with high-risk scores, since immune checkpoint molecules are more highly expressed in low-risk groups, and the increased levels of immune checkpoints indirectly indicate pre-existing T cell activation in the low-risk group.

This study has some limitations. The six-gene model was derived from retrospective data, and more prospective data are needed to validate our results. Second, this study lacked basic experiments to validate the function of the six hub genes and their association with immune cell infiltration. Third, patients receiving immunotherapy were not included in this study; therefore, the predictive ability of the model for immunotherapy response was evaluated indirectly.

Conclusions

In conclusion, we constructed a TME-related prognostic model to predict LUAD patient survival outcomes and responses to immunotherapy. Patients with low risk scores had better prognoses and were expected to benefit from ICI treatment. This model might be valuable for prognostic management and patient selection before immunotherapy and deserves further validation. A significant association was observed between the hub genes and patient prognosis and immune infiltration, providing novel insights for the development of new treatment strategies.

Acknowledgments

We are grateful to the TCGA database and GEO database for the source of data used in our study.

Conflicts of interest

The authors indicated no potential conflicts of interest.

References

1. Bray F, J Ferlay, I Soerjomataram, *et al*. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin, 2018, 68: 394–424.
2. Schneider BJ, N Ismaila, N Altorki. Lung cancer surveillance after definitive curative-intent therapy: ASCO Guideline Summary. JCO Oncol Pract, 2020, 16: 83–86.
3. Herbst RS, D Morgensztern, C Boshoff. The biology and management of non-small cell lung cancer. Nature, 2018, 553: 446–454.
4. Fenizia F, R Pasquale, C Roma, *et al*. Measuring tumor mutation burden in non-small cell lung cancer: tissue versus liquid biopsy. Transl Lung Cancer Res, 2018, 7: 668–677.
5. Zou W, JD Wolchok, L Chen. PD-L1 (B7-H1) and PD-1 pathway blockade for cancer therapy: Mechanisms, response biomarkers, and combinations. Sci Transl Med, 2016, 8: 328rv4.
6. Rotte A, JY Jin, V Lemaire. Mechanistic overview of immune checkpoints to support the rational design of their combinations in cancer immunotherapy. Ann Oncol, 2018, 29: 71–83.
7. Braun DA, KP Burke, EM Van Allen. Genomic Approaches to Understanding Response and Resistance to Immunotherapy. Clin Cancer Res, 2016, 22: 5642–5650.
8. Wood SL, M Pernemalm, PA Crosbie, *et al*. The role of the tumor-microenvironment in lung cancer-metastasis and its relationship to potential therapeutic targets. Cancer Treat Rev, 2014, 40: 558–566.
9. Hanahan D, LM Coussens. Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer Cell, 2012, 21: 309–322.
10. Quail DF, JA Joyce. Microenvironmental regulation of tumor progression and metastasis. Nat Med, 2013, 19: 1423–1437.
11. Guo X, Y Zhang, L Zheng, *et al*. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. Nat Med, 2018, 24: 978–985.
12. Bussard KM, L Mutkus, K Stumpf, *et al*. Tumor-associated stromal cells as key contributors to the tumor microenvironment. Breast Cancer Res, 2016, 18: 84.
13. Carter SL, K Cibulskis, E Helman, *et al*. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol, 2012, 30: 413–421.
14. Alonso MH, S Ausso, A Lopez-Doriga, *et al*. Comprehensive analysis of copy number aberrations in microsatellite stable colon cancer in view of stromal component. Br J Cancer, 2017, 117: 421–431.
15. Luo C, M Lei, Y Zhang, *et al*. Systematic construction and validation of an immune prognostic model for lung adenocarcinoma. J Cell Mol Med, 2020, 24: 123–144.
16. Nishino M, NH Ramaiya, H Hatabu, *et al*. Monitoring immune-checkpoint blockade: response evaluation and biomarker development. Nat Rev Clin Oncol, 2017, 14: 655–668.
17. Noonan DM, A De Lerma Barbaro, N Vannini, *et al*. Inflammation, inflammatory cells and angiogenesis: decisions and indecisions. Cancer Metastasis Rev, 2008, 27: 31–40.
18. Chen DS, I Mellman. Elements of cancer immunity and the cancer-immune set point. Nature, 2017, 541: 321–330.
19. Chanmee T, P Ontong, K Konno, *et al*. Tumor-associated macrophages as major players in the tumor microenvironment. Cancers (Basel), 2014, 6: 1670–1690.
20. Chen X, E Song. Turning foes to friends: targeting cancer-associated fibroblasts. Nat Rev Drug Discov, 2019, 18: 99–115.
21. Orimo A, PB Gupta, DC Sgroi, *et al*. Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. Cell, 2005, 121: 335–348.
22. Mohamed AJ, L Yu, CM Backesjo, *et al*. Bruton's tyrosine kinase (Btk): function, regulation, and transformation with special emphasis on the PH domain. Immunol Rev, 2009, 228: 58–73.
23. Gayko U, M Fung, F Clow, *et al*. Development of the Bruton's tyrosine kinase inhibitor ibrutinib for B cell malignancies. Ann N Y Acad Sci, 2015, 1358: 82–94.
24. Molina-Cerrillo J, T Alonso-Gordoa, P Gajate, *et al*. Bruton's tyrosine kinase (BTK) as a promising target in solid tumors. Cancer Treat Rev, 2017, 58: 41–50.
25. Esensten JH, YA Helou, G Chopra, *et al*. CD28 costimulation: from mechanism to therapy. Immunity, 2016, 44: 973–988.
26. Kamphorst AO, A Wieland, T Nasti, *et al*. Rescue of exhausted CD8 T cells by PD-1-targeted therapies is CD28-dependent. Science, 2017, 355: 1423–1427.
27. Lien EC, CC Dibble, A Toker. PI3K signaling in cancer: beyond AKT. Curr Opin Cell Biol, 2017, 45: 62–71.
28. Pei Y, N Hwang, F Lang, *et al*. Quassinoid analogs with enhanced efficacy for treatment of hematologic malignancies target the PI3Kgamma isoform. Commun Biol, 2020, 3: 267.
29. Maryam A Shetab Boushehri, Alf Lamprecht. TLR4-based immunotherapeutics in cancer: a review of the achievements and shortcomings. Mol Pharm, 2018, 15: 4777–4800.
30. Janiszewska M, DP Tabassum, Z Castano, *et al*. Subclonal cooperation drives metastasis by modulating local and systemic immune microenvironments. Nat Cell Biol, 2019, 21: 879–888.
31. Weickhardt AJ, DS Williams, CK Lee, *et al*. Vascular endothelial growth factor D expression is a potential biomarker of bevacizumab benefit in colorectal cancer. Br J Cancer, 2015, 113: 37–45.
32. Singh P, LM Jenkins, B Horst, *et al*. Inhibin Is a Novel Paracrine Factor for Tumor Angiogenesis and Metastasis. Cancer Res, 2018, 78: 2978–2989.
33. Chowell D, C Krishna, F Pierini, *et al*. Evolutionary divergence of HLA class I genotype impacts efficacy of cancer immunotherapy. Nat Med, 2019, 25: 1715–1720.
34. Patel SP, R Kurzrock. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. Mol Cancer Ther, 2015, 14: 847–856.

DOI 10.1007/s10330-021-0545-5
Cite this article as: Li Z, Feng YQ, Li P, *et al*. Development and validation of a tumor microenvironment-related prognostic signature in lung adenocarcinoma and immune infiltration analysis. Oncol Transl Med, 2021, 7: 253–268.

Oncology and Translational Medicine
DOI 10.1007/s10330-021-0509-9

December 2021, Vol. 7, No. 6, P269–P274

ORIGINAL ARTICLE

Correlation analysis of breast fibroadenoma and the intestinal flora based on 16S rRNA sequencing*

Bingdong Wang[1], Xin Liu[1], Yahong Bian[2], Guoxin Sun[1], Huizhe Wang[1], Jingjin Zhang[3], Zhengfu Zhang[2] (✉), Xiao Zou[1] (✉)

[1] Department of Breast Surgery, Qingdao Central Hospital Affiliated to Qingdao University, Qingdao 266042, China
[2] Qingdao Central Hospital Affiliated to Qingdao University, Qingdao 266003, China
[3] Linyi Central Hospital, Linyi 276400, China

Abstract

Objective To analyze the characteristics of the intestinal microflora in patients with breast fibroadenoma using 16S ribosomal RNA (rRNA) high-throughput sequencing.

Methods Fecal samples from 20 patients with breast fibroadenoma and 36 healthy subjects were randomly collected and analyzed using high-throughput sequencing technology for 16S rRNA V4 region sequencing, and the alpha diversity (Chao index, Shannon index) was calculated using Mothur (v.1.39.5) software. Beta diversity was analyzed using QIIME (v1.80). SPSS software (version 23.0) and the t-test of two independent samples were used to analyze differences in the abundance of bacteria between the two groups.

Results Compared with that in the healthy control group, the α diversity of the intestinal microflora in breast fibroadenoma patients increased, but the difference was not statistically significant ($P > 0.05$). At the phylum level, significant differences were observed between the two groups. The abundance of Firmicutes was higher in the breast fibroadenoma group ($P < 0.05$), whereas the abundance of Synergistetes was higher in the healthy control group ($P < 0.005$). A total of five bacterial genera showed significant differences between the two groups: the breast fibroadenoma group showed higher levels of Bautia ($P < 0.005$), Coprococcus ($P < 0.005$), Roseburia ($P < 0.05$), and Ruminococcus ($P < 0.005$), whereas Sutterella was more abundant in the healthy control group than in the breast fibroadenoma group ($P < 0.05$).

Conclusion The diversity and abundance of the intestinal flora in patients with breast fibroadenoma are significantly different from those in healthy subjects, suggesting that an imbalance in the intestinal flora is correlated with the occurrence of breast fibroadenoma.

Received: 15 July 2021
Revised: 31 August 2021
Accepted: 21 October 2021

Key words: intestinal flora; estrogen; breast fibroadenoma; 16S ribosomal RNA; high-throughput sequencing

Breast fibroadenoma is the most common benign breast tumor in women. It is primarily composed of proliferative breast fibrous tissue and ducts, and its occurrence may be related to an abnormal quality or quantity of estrogen receptors contained in fibroblasts; however, the precise etiology remains unclear [1].

The intestinal tract is the largest digestive organ of the human body, which contains a large number of bacteria and has a genome approximately 100 times that of humans [2].The human intestinal flora contains genes that encode thousands of microbial enzymes and metabolites [3–4]. The intestinal flora is closely related to the estrogen metabolism in the body. The intestinal microbes contain genes related to estrogen metabolism and encode β-glucuronidase. When the content of this enzyme in the intestine increases, the glucuronidase-estrogen conjugate is decomposed; estrogen returns to the free state and is re-absorbed into the blood through the hepatointestinal circulation, thus leading to an increase in endogenous estrogen levels [5].Therefore, changes in estrogen levels caused by the imbalance of intestinal flora may be an important factor in the occurrence of breast

✉ Correspondense to: Zhengfu Zhang. Email: sqzzf1964@163.com; Xiao Zou, Email:1392365491@qq.com
* Supported by a grant from the Qingdao Pharmaceutical Research Guidance Plan 2019 (No. 2019-WJZD140).

fibroadenoma.

A number of studies have confirmed that the intestinal flora can affect the occurrence and development of breast cancer through estrogen metabolism, immune regulation, and generation of short-chain fatty acids (SCFAs) [5–8]; however, a correlation between breast fibroadenoma and intestinal microbes has not been reported. This study involved the collection of stool samples from 20 female patients with breast fibroadenoma and 36 healthy adult women. Through 16S rRNA sequencing of the V4 area and variance analysis, the intestinal flora diversity and composition of the samples were evaluated. These results provide a new theoretical basis for the diagnosis and prevention of breast fibroadenoma.

Materials and methods

Case selection

Patients admitted to the Breast Surgery Department of Qingdao Central Hospital and diagnosed with breast fibroadenoma via postoperative paraffin section pathology in the Pathological Diagnostic Center and healthy adult females without any breast-related diseases, as confirmed by the Physical Examination Center, were selected. All subjects had a normal body mass index (BMI) and had not used antibiotics, probiotics, antacids, gastrointestinal motility agents, or other drugs that could affect the intestinal flora in the 6 months before enrollment. The subjects did not have hypertension, coronary heart disease, diabetes, cirrhosis, malignant tumors, or other primary diseases. A total of 56 female subjects were included in this study, including 20 patients with breast fibroadenoma and 36 healthy adult females. All subjects signed an informed consent form and volunteered to participate in the study.

Specimen collection

Fresh fecal samples (no less than 10 g) from the 56 subjects were collected with sterile cotton swabs, placed in a sterile container, and immediately stored in a refrigerator at –80°C for low-temperature preservation. All of the above procedures were performed under sterile conditions.

Amplifier sequencing

The collected fecal samples were cryopreserved and sent to Qingdao BGI Institute for gene sequencing. The process was as follows: (1) Genomic DNA extraction: The cetyltrimethylammonium bromide (CTAB) or sodium dodecyl sulfate (SDS) method was used to extract genomic DNA from the samples, and agarose gel electrophoresis was used to detect the purity and concentration of the DNA. An appropriate amount of the samples was placed in a centrifuge tube, and the samples were diluted to 1 ng/μL with sterile water. (2) PCR amplification: Diluted genomic DNA was used as a template. According to the selection of the sequencing region, specific primers with Barcode were used; the 16S V4 primer was 515F-806R. The Phusion® High-Fidelity PCR Master Mix and GC Buffer from New England Biolabs were used. PCR was performed using high-efficiency and high-fidelity enzymes to ensure amplification efficiency and accuracy. The PCR was conducted on the Bio-Rad T100 gradient PCR instrument. PCR products were detected by electrophoresis on a 2% agarose gel. (3) Mixing and purification of PCR products: The PCR products were mixed and purified according to the concentration of the PCR products, and the PCR products were mixed at the same concentration. After thorough mixing, the PCR products were purified by electrophoresis with a 1× TAE concentration of 2% agarose, and the target bands were recovered by gelling, using the Thermo Scientific Genejet Gel Recovery Kit. (4) Library construction and computer sequencing: The Illumina TruSeq DNA PCR-Free Library Preparation Kit was used to construct the library. After Qubit quantification and library testing, NovaSeq 6000 was used for computer sequencing of the qualified library.

Bacterial community information analysis

The software Mothur v.1.39.5 was used to remove all the redundant tags, and the software USEARCH (v7.0.1090) was used to cluster the spliced tags into operational taxonomic units (OTUs). After the OTU representative sequence was obtained, species annotation was carried out by comparing the OTU representative sequence with the Greengenes database using RDP Classifier (V2.2) software, and the confidence threshold was set to 0.8. Alpha diversity was calculated using Mothur (v.1.39.5) software, and beta diversity was analyzed using QIIME (v1.80).

Statistical analysis

SPSS software (version 23.0) was used for data analysis, and the *t*-test of two independent samples was used to analyze differences in the abundance of bacteria between the two groups. Statistical significance was set at $P < 0.05$.

Results

Sequencing data, sample out, and diversity analysis

A total of 5004,192 high-quality sequences were obtained from 56 samples in the two groups, with an average sequence length of approximately 252 bp, and a total of 3911 OTUs were generated. The sequencing coverage of all samples reached 99.9%. The dilution curve reflects whether the sequencing quantity of the sample was sufficient. If the curve flattens or reaches the

plateau stage, the sequencing depth can be considered to have covered all the species in the sample. The contrary means that the species diversity in the sample is high and there are more species that have not been detected by sequencing. As shown in Fig. 1, the Chao index dilution curve gradually flattened with the increase in sequencing data, as did the Shannon index dilution curve. Therefore, it can be concluded that the sequencing depth covered all species.

The α diversity index was calculated based on the species and abundance of OTUs.

The Chao index reflects the richness of the community in the sample; the higher the index, the richer the species. The Shannon index reflects the diversity of the community; the larger the Shannon index, the greater the diversity of the community, as shown in Fig. 2a and 2b. There was no significant difference in the abundance of intestinal microbial species ($P > 0.05$), but there was an increasing trend of α diversity in the breast fibroadenoma group. In addition, principal component analysis (PCA) was conducted based on the OTU level (Fig. 2c). The samples from the breast fibroadenoma group (shown in green in Fig. 2c) and the healthy control group (shown in blue in Fig. 2c) were compared and analyzed. The results showed that the bacterial flora compositions of the two groups were different.

Analysis of flora structure and relative abundance

In this study, the structure and composition of the intestinal flora of the breast fibroadenoma group and the healthy adult female group were analyzed at the phylum and genus levels, respectively.

(1) Relative abundance analysis of the intestinal flora at the phylum level

At the phylum level, the top 13 strains were selected

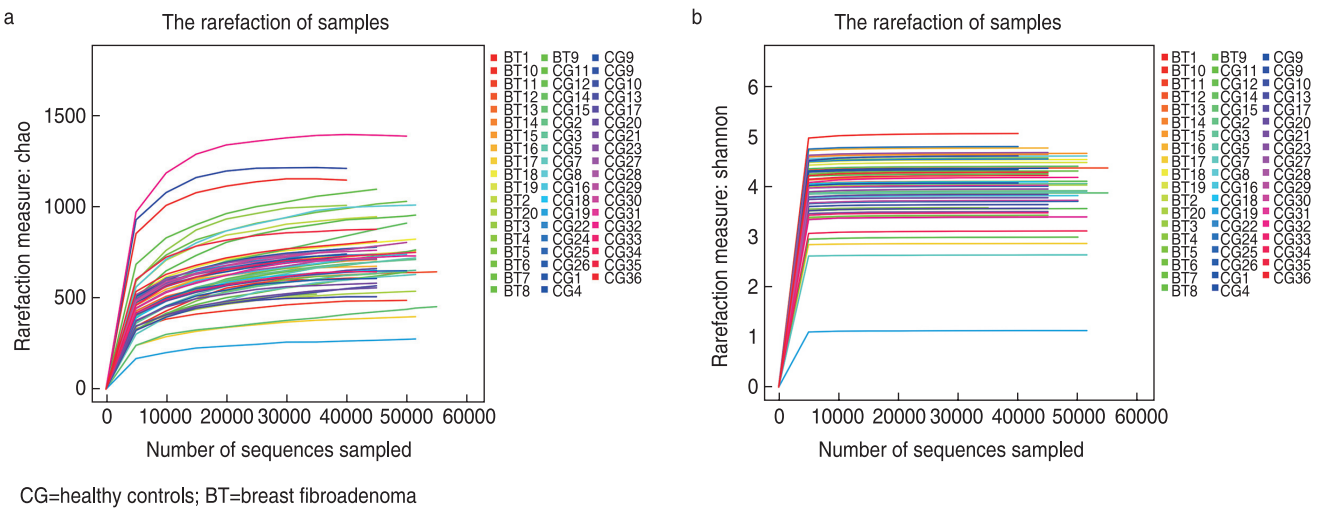


Fig. 1 Chao index and Shannon index dilution curve

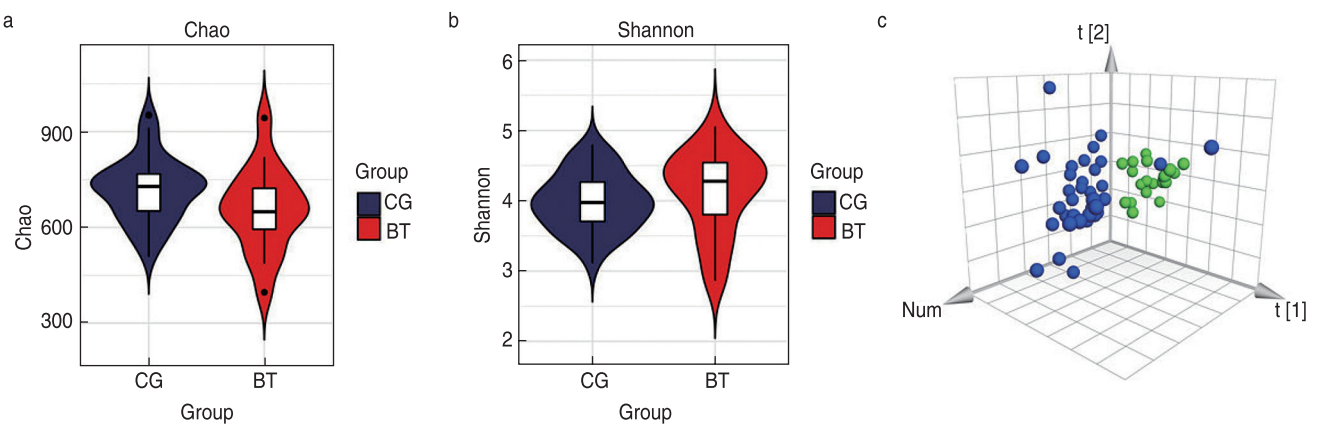


Fig. 2 α-diversity analysis and principal component analysis (PCA)

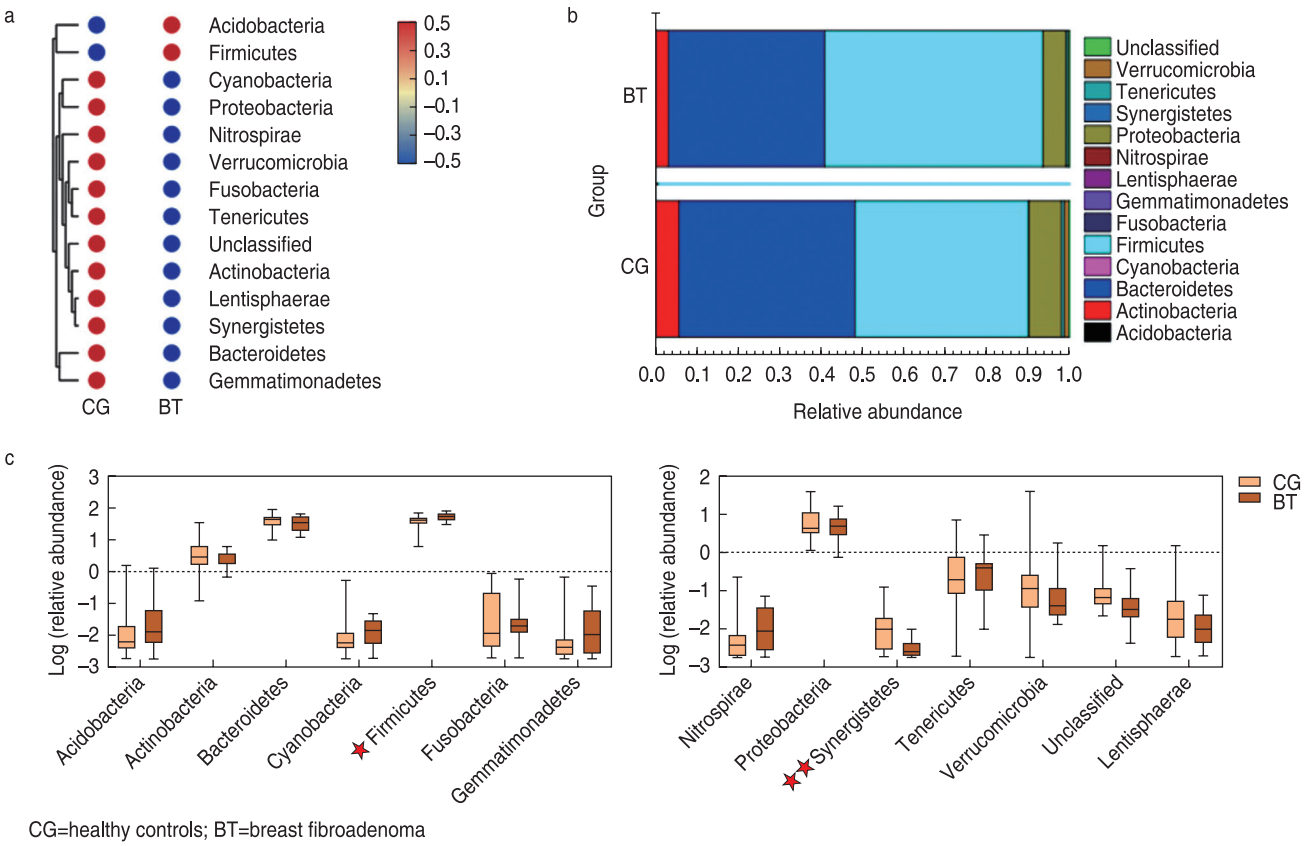
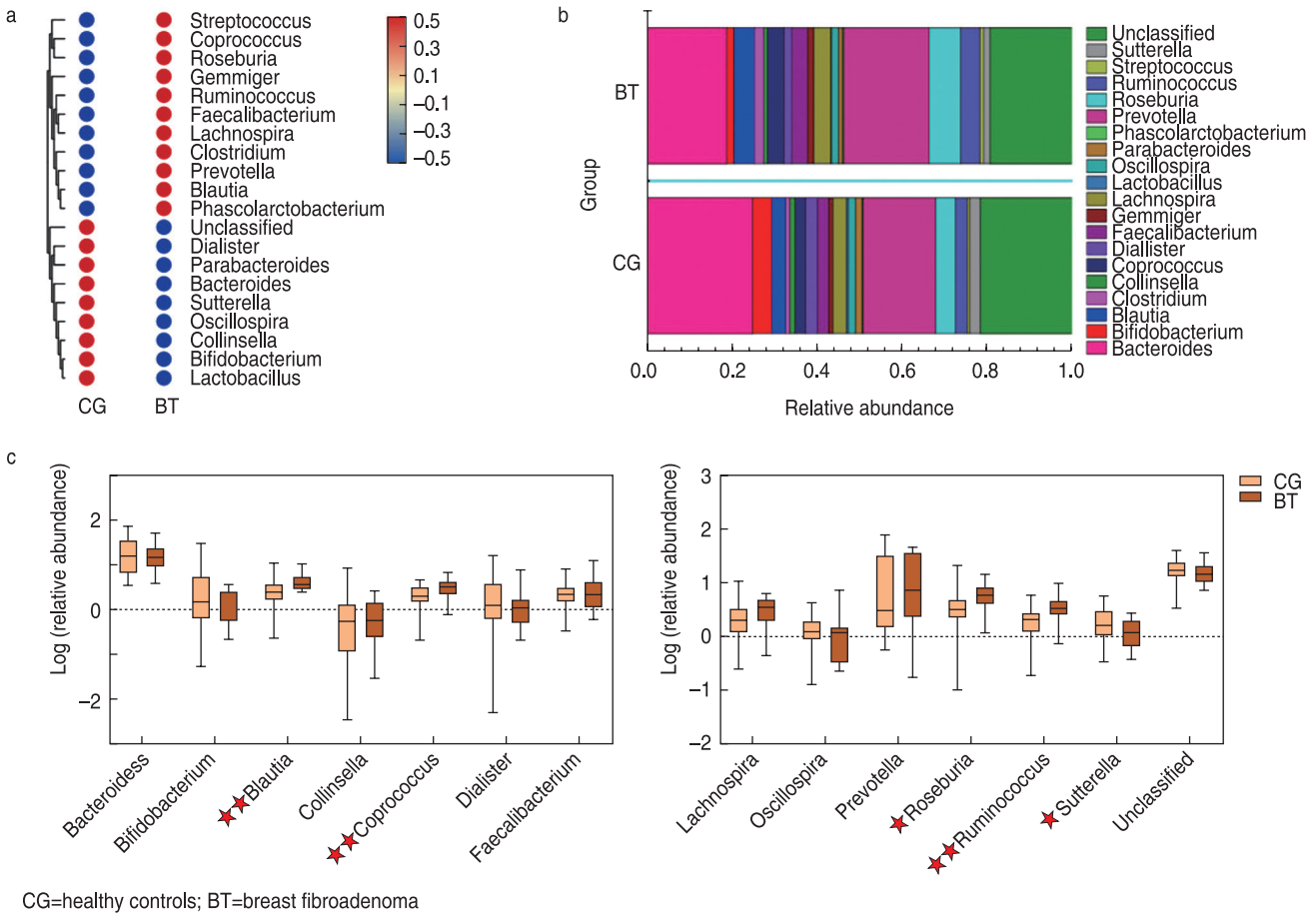


Fig. 3 Analysis of relative abundance of the intestinal flora at the phylum level between the healthy control group and the breast fibroadenoma group

for relative abundance analysis, as shown in Fig. 3a. The dominant bacterial phyla in the healthy control group were Cyanobacteria, Proteobacteria, Nitrospirae, Verrucomicrobia, and Fusobacteria, Tenericutes, Actinobacteria, Lentisphaerae, Synergistetes, Bacteroidetes and Gemmatimonadetes, whereas the dominant phyla in the breast fibroadenoma group were Acidobacteria and Firmicutes. As shown in Figure 3B, at the phylum level, the dominant species in both groups were Firmicutes, Bacteroidetes, Verrucomicrobia, and Actinobacteria. However, there were significant differences in species composition between the healthy control group and the breast fibroadenoma group. Based on the abundance data of the two groups, the *t*-test of two independent samples was used to analyze the species with different phylum levels in the intestinal flora of the breast fibroadenoma group and the healthy adult female group. As shown in Fig. 3c, there were significant differences in the two categories between the two groups. The abundance of Firmicutes was higher in the breast fibroadenoma group than in the healthy control group ($P < 0.05$). Synergistetes were more abundant in the healthy control group than in the breast fibroadenoma group ($P < 0.005$). According to these results, the two abovementioned phyla may be correlated with the occurrence of breast fibroadenoma; however, further analysis is necessary.

(2) Relative abundance analysis of the intestinal flora at the genus level

The top 19 bacterial genera were selected for further analysis of the relative abundance. As shown in Fig. 4a, the dominant genera in the healthy control group were Dialister, Parabacteroides, Bacteroides, Sutterella, Oscillospira, Collinsella, Bifidobacterium, and Lactobacillus. The dominant species in the breast fibroadenoma group were *Streptococcus*, *Coprococcus*, *Roseburia*, *Gemmiger*, *Ruminococcus*, *Facecalibacterium*, *Lachnospira*, *Clostridium*, *Prevotella*, *Blautia*, and *Phascolarctobacterium*. As shown in Fig. 4b, the dominant species in both groups were Bacteroides, Prevotella, and Roseburia. Based on the abundance data of the two groups, the t-test of two independent samples was used to analyze the species with a different intestinal flora at the genus level between the breast fibroadenoma group and the healthy adult female group. As shown in Fig. 4c, a total of five bacterial genera showed significant differences between the two groups of samples. The breast fibroadenoma group and the healthy control group



CG=healthy controls; BT=breast fibroadenoma

Fig. 4 Analysis of relative abundance of the intestinal flora at the genus level between the healthy control group and the breast fibroadenoma group

had comparable levels of Bautia ($P < 0.005$), Coprococcus ($P < 0.005$), Roseburia ($P < 0.05$) and Ruminococcus ($P < 0.005$). The abundance of Sutterella in the healthy control group was higher than that in the fibroadenoma group ($P < 0.05$).

Discussion

Breast fibroadenoma is the most common breast fibroepithelial tumor in women. These tumors are hormone dependent; they increase in size due to factors such as estrogen, progesterone, prolactin, and pregnancy and decrease after menopause [9, 10]. Currently, there are no effective preventive measures for breast fibroma, and the treatment primarily consists of surgical resection, which is associated with a risk of recurrence [11, 12]. Studies have shown that breast fibroadenoma is an independent risk factor for breast cancer, and the risk of breast cancer 20 years later in patients with breast fibroadenoma is twice that in healthy women [13]. Studies have confirmed that the intestinal flora plays an important role in the occurrence and progression of breast cancer [6–8]. However, it is still unclear whether there is a correlation between breast fibroadenoma and the intestinal flora. Therefore, this study used 16S rRNA high-throughput sequencing to evaluate the intestinal flora of patients with breast fibroadenoma. The results of this study showed that compared with that of the healthy control group, the intestinal microflora of the patients with breast fibroadenoma showed an increased α diversity, indicating an imbalance of the intestinal microflora in patients with breast fibroadenoma. The dominant phyla in the breast fibroadenoma group and the healthy control group were Firmicutes, Bacteroidetes, Verrucomicrobia, and Actinobacteria. However, the abundance of Firmicutes was higher in the breast fibroadenoma group than in the healthy control group, and the abundance of Synergistetes was higher in the healthy control group than in the breast fibroadenoma group. A total of five bacterial genera showed significant differences between the two groups. Compared with that in the healthy control group, the abundance of Blautia, Coprococcus, Roseburia and Ruminococcus in the breast fibroadenoma group was higher. The abundance of

Sutterella was higher in the healthy control group than in the fibroadenoma group. Chan *et al*. [14] pointed out that Firmicutes, Proteobacteria, and Bacteroidetes have β-glucuronidase activity. Exogenous estrogen levels are closely related to β-glucuronidase [15]. In this study, we found that the abundance of Firmicutes was higher in the breast fibroadenoma group than in the healthy control group, and the difference was statistically significant ($P < 0.05$). Therefore, the significant differences in the abundance of Firmicutes suggest that the imbalance of the intestinal flora may influence the development of breast fibroadenoma by affecting estrogen metabolism, which is consistent with the previous hypothesis. Patients with breast fibroadenoma have a relatively high Prevotella content; Prevotella can induce intestinal mucosal inflammation [16]. Therefore, patients with breast fibroadenoma may show intestinal mucosal injury. In addition, the content of SCFA-producing bacteria, such as Streptococcus, Coprococcus, Ruminococcus, Lachnospira, and Clostridium, in breast fibroadenoma patients was relatively high. SCFAs, primarily acetate, propionate, and butyrate, are bacterial fermentation products derived from soluble dietary fiber in the colon. A growing body of evidence suggests that SCFAs play a key role in maintaining the intestinal barrier by stabilizing specific transcription factors, promoting the composition of tight junctions and the secretion of mucins [17]. SCFAs also regulate the differentiation of T cells into effector cells or regulatory T cells and are considered potential predictors of immunotherapeutic responses in some cancers [18]. Therefore, the results of this study suggest that the intestinal flora of patients with breast fibroadenoma may be associated with an immune response. In conclusion, patients with breast fibroadenoma show an imbalance of the intestinal flora.

This study had the following limitations: a small sample size, analysis by 16S rRNA level observational studies, lack of validation using large samples and animal experiments. Nevertheless, when combined with relevant clinical indicators, the findings of this study might provide important theoretical guidance for the prevention, diagnosis, and treatment of breast fibroadenoma.

Conflicts of interest

The authors indicated no potential conflicts of interest.

References

1. Houssami N, Cheung MN, Dixon JM. Fibroadenoma of the breast. Med J Aust, 2001, 174: 185–188.
2. Filippo CD, Cavalieri D, Paola MD, *et al*. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A, 2010, 107: 14691–14696.
3. Bull MJ, Plummer NT. Part 1: The Human Gut Microbiome in Health and Disease. Integr Med, 2014, 13: 17–22.
4. Rath CM, Dorrestein PC. The bacterial chemical repertoire mediates metabolic exchange within gut microbiomes. Curr Opin Microbiol, 2012, 15: 147–154.
5. Plottel C, Blaser M. Microbiome and Malignancy. Cell Host Microbe, 2011, 10: 324–335.
6. Flores R, Shi J, Fuhrman B, *et al*. Fecal microbial determinants of fecal and systemic estrogens and estrogen metabolites: a cross-sectional study. J Transl Med, 2012, 10: 253.
7. Goedert JJ, Hua X, Bielecka A, *et al*. Postmenopausal breast cancer and oestrogen associations with the IgA-coated and IgA-noncoated faecal microbiota. Br J Cancer, 2018, 118: 471–479.
8. He C, Liu Y, Ye S, *et al*. Changes of intestinal microflora of breast cancer in premenopausal women. Eur J Clin Microbiol Infect Dis, 2020, 40: 503–513.
9. Loke BN, Nasir NDM, Thike AA, *et al*. Genetics and genomics of breast fibroadenomas. J Clin Pathol, 2018, 71: 381–387.
10. Funahashi S, Okazaki Y, Nagai H, *et al*. Twist1 was detected in mesenchymal cells of mammary fibroadenoma and invasive components of breast carcinoma in rats. J Toxicol Pathol, 2019, 32: 19–26.
11. Sun C, Zhang W, Ma H, *et al*. Main traits of breast fibroadenoma among adolescent girls. Cancer Biother Radiopharm, 2020, 35: 271–276.
12. Lo Martire N, Nibid A, Farello G, *et al*. Giant fibroadenoma of the breast in an adolescent: a case report. Ann Ital Chir, 2002, 73: 631–634.
13. Dupont WD, Page DL, Parl FF, *et al*. Long-term risk of breast cancer in women with fibroadenoma. N Engl J Med, 1994, 331: 10–15.
14. Chan AA, Bashir M, Rivas MN, *et al*. Characterization of the microbiome of nipple aspirate fluid of breast cancer survivors. Sci Rep, 2016, 6: 28061.
15. Baker JM, Al-Nakkash L, Herbst-Kralovetz MM. Estrogen-gut microbiome axis: Physiological and clinical implications. Maturitas, 2017, 103: 45–53.
16. Su T, Liu R, Lee A, *et al*. Altered intestinal microbiota with increased abundance of prevotella is associated with high risk of diarrhea-predominant irritable bowel syndrome. Gastroenterol Res Pract, 2018, 2018: 6961783.
17. Xie QS, Zhang JX, Liu M, *et al*. Short-chain fatty acids exert opposite effects on the expression and function of p-glycoprotein and breast cancer resistance protein in rat intestine. Acta Pharmacol Sin, 2021, 42: 470–481.
18. Malczewski AB, Navarro S, Coward JI, *et al*. Microbiome-derived metabolome as a potential predictor of response to cancer immunotherapy. J Immunother Cancer, 2020, 8: e001383.

DOI 10.1007/s10330-021-0509-9
Cite this article as: Wang BD, Liu X, Bian YH, *et al*. Correlation analysis of breast fibroadenoma and the intestinal flora based on 16S rRNA sequencing. Oncol Transl Med, 2021, 7: 269–274.

Oncology and Translational Medicine
DOI 10.1007/s10330-021-0532-2
December 2021, Vol. 7, No. 6, P275–P278

ORIGINAL ARTICLE

Effect of radiotherapy on tumor markers and serum immune-associated cells in patients with esophageal cancer*

Wei Gao[1], Xiaoxiao Liu[2], Hongbing Ma[2] (✉)

[1] Department of Internal Medicine, The Hospital of Xi 'an Architectural Science and Technology University, Xi'an 710055, China
[2] Department of Oncology Radiotherapy, The Second Affiliated Hospital of Xi 'an Jiaotong University, Xi'an 710004, China

Abstract

Objective This study aimed to investigate the effect of radiotherapy on serum immune-associated cells and tumor markers in patients with esophageal cancer.

Methods A total of 87 patients with esophageal cancer admitted to our hospital between October 2016 and July 2020 were selected as the observation group, and all patients received radiotherapy. A total of 87 healthy volunteers who underwent physical examination at our hospital during the same period were selected as the control group in order to compare the changes in serum immune-associated cells and tumor markers between the two groups.

Results The levels of carcinoembryonic antigen (CEA), cancer antigen (CA) 125, CA72-4, C-terminus of cytokeratin (CYFRA) 21-1, and squamous cell carcinoma (SCC) antigen in the observation group before radiotherapy were higher than those in the control group, and the differences were significant ($P < 0.05$). The levels of CEA, CA125, CA72-4, CYFRA21-1, and SCC antigen in the research group after radiotherapy were significantly lower than those before radiotherapy, but were still significantly higher than those in the control group ($P < 0.05$). The levels of $CD3^+$, CD4+, CD4+/CD8+, and natural killer cells in the research group before and after radiotherapy were significantly lower, while the levels of Treg and CD8+ cells were significantly higher than those in the control group ($P < 0.05$). The levels of CD3+, CD4+, and CD4+/CD8+ cells in the observation group after radiotherapy were lower, while the levels of CD8+ cells were significantly higher than those before radiotherapy ($P < 0.05$).

Conclusion Radiotherapy can effectively reduce the level of serum tumor markers in patients with esophageal cancer; these antigens and cells can be used as tumor markers of esophageal cancer in order to determine its prognosis. However, radiotherapy has adverse effects on the immune function of the body. The reasons behind this need to be further studied and analyzed.

Received: 6 November 2021
Revised: 21 November 2021
Accepted: 29 November 2021

Key words: radiotherapy; esophageal cancer; tumor markers, immune-associated cells

Esophageal cancer is a common malignant tumor of the digestive system, which gravely threatens human health and safety. Most esophageal cancer patients are already in the middle or late stages of the disease by the time they seek treatment and are no longer suitable to undergo surgery, and have a 5-year survival rate of less than 20% [1]. Radiotherapy has become an important method of treatment for middle and advanced esophageal cancer [2–3]; however, due to the different sensitivities of each individual to radiotherapy, identification of effective indicators is needed to determine the therapeutic effect of radiotherapy. Serum tumor markers are rapid, simple, and less invasive detection indicators. Radiotherapy can treat esophageal cancer, but may also affect the normal immune function of the body. To further understand the clinical effect of radiotherapy on esophageal cancer and its influence on the immune function, changes in serum tumor markers and T cell subsets in patients with

✉ Correspondence to: Hongbing Ma. Email: mhbxian@126.com
* Supported by a grant from the National Natural Science Foundation of China (No. 81872471)

esophageal cancer were determined before and after radiotherapy and analyzed in this study.

Materials and methods

Data

A total of 87 patients with esophageal cancer admitted in our hospital between October 2016 and July 2020 were selected as the observation group, based on the following inclusion and exclusion criteria.

(1)Inclusion criteria: Patients who underwent endoscopic examination, ultrasound examination, and pathological examination to confirm the presence of TNM stage III; who had esophageal cancer metastasis; who refused to undergo surgical treatment or were unable to undergo surgery prior to radiotherapy; with no history of radiation or chemotherapy; with no contraindications to radiotherapy or chemotherapy; who had an expected survival time of not less than 6 months; who had a quality of life score of > 60 points; and who signed an informed consent, were included in the study.

(2)Exclusion criteria: Patients with severe heart, liver, and renal insufficiencies; mental disorders; and immune system diseases were excluded.

The observation group comprised 55 men and 32 women, with ages ranging from 43 to 82 years (mean age: 61.7 ± 8.2 years); with regard to the TNM stage, 37 patients had stage III disease, while 50 had stage IV disease. In terms of the pathological type, 73 patients had squamous cell carcinoma (SCC), while 14 had adenosquamous cell carcinoma. Ninety healthy volunteers who underwent physical examination in our hospital during the same period were selected as the control group, which included 53 men and 34 women. Their ages ranged from 41 to 83 years (mean age: 62.4 ± 9.2 years). No significant difference was observed in the sex, age, and other basic data between the two groups ($P > 0.05$), thus indicating comparability (Table 1).

Table 1 Characteristics of patients in observation group

Factor	*n*	%
Age (years)		
≤ 60	21	24.1
> 60	66	75.9
Sex		5
Male	55	63.2
Female	32	36.8
Smoking		
No	29	33.3
Yes	58	66.7
Pathological types		
Squamous	73	83.9
Adenosquamous cell carcinoma	14	16.1
Tumour site		
Upper 1/3	30	34.5
Middle 1/3	41	47.1
Lower 1/3	16	18.4
T stage		
II	2	2.3
III	35	40.2
IV	50	57.5
N stage (/)		
0	55	63.2
1	32	36.8
Clinic stage		
II	4	4.6
III	33	37.9
IV	50	57.5

Methods

Treatment

All patients underwent computed tomography (CT). Continuous spiral CT scanning was performed with the upper boundary at the upper edge of the fourth cervical vertebra and the lower boundary at the lower edge of the second lumbar vertebra. The scanning images were transmitted to the three-dimensional (3D) treatment planning system; the target area was determined according to the examination results, and 3–5 coplanar fields were selected for irradiation. The radiotherapy techniques used were Varian linear accelerator 3D conformal radiotherapy, image-guided intensity-modulated radiotherapy, or volume rotation intensity-modulated radiotherapy. A total radiotherapy dose of 60 Gy (2 Gy/fraction) was delivered, and the radiotherapy was performed for 30–32 cycles.

Fasting venous blood

Fasting venous blood was collected before and 1 month after radiotherapy in the observation group. Fasting venous blood was collected from the control group in the morning of the physical examination day and centrifuged at 3,000 r/min for 10 min; the serum was separated and stored at 2 °C to 6 °C for detection. Levels of the following serum tumor markers were measured using an Abbott I2000 chemiluminescence analyzer and the associated reagents: carcinoembryonic antigen (CEA), carbohydrate antigen 19-9 (CA19-9), carbohydrate antigen 72-4 (CA72-4), cytokeratin 19 fragment (CYFRA21-1), and SCC antigen. The levels of T cell subsets, including CD3+, CD4+, CD8+, and CD4+/ CD8+ cells, were measured using a FACS Canto II flow cytometer.

Statistical analysis

The SPSS22.0 software was used to perform all data analyses. The measurement data were expressed as mean ± standard. Independent-sample *t* test was used to perform a between-group comparison, while paired *t*-test

was used to perform a within-group comparison.

Results

Comparison of serum tumor markers

The CEA, CA19-9, CA72-4, CYFRA21-1, and SCC antigen levels before radiotherapy in the observation group were higher than those in the control group, and the differences were significant ($P < 0.05$). The CEA, CA19-9, CA72-4, CYFRA21-1, and SCC antigen levels in the observation group after radiotherapy were lower than those before radiotherapy, but were still higher than those of the control group; the differences were significant ($P < 0.05$; Table 2).

Comparison of T cell subsets

The levels of CD3+, CD4+, and CD4+/CD8+ cells before and after radiotherapy in the observation group were lower than those in the control group, while the levels of CD8+ were higher than those in the control group; the differences were significant ($P < 0.05$). The levels of CD3+, CD4+, and CD4+/CD8+ cells in the observation group after radiotherapy were lower than those before radiotherapy, while the levels of CD8+ cells were higher than those before radiotherapy; the differences were significant ($P < 0.05$; Table 3).

Discussion

Esophageal cancer is a common malignant tumor of the digestive system and poses a serious threat to the life and health of Chinese residents. At present, surgery is the primary treatment method for early-stage esophageal cancer; however, for patients with middle and advanced stage esophageal cancer, surgical resection is no longer effective, and thus, radiotherapy is often preferred for the clinical treatment of middle and advanced stage esophageal cancer [4]. In recent years, continual development in 3D conformal radiotherapy technology has allowed focus on the target area of tumor cells for irradiation, while reducing unnecessary damage to the surrounding normal tissues. However, radiotherapy may inhibit the immune function of the body, thus affecting its therapeutic effects. Therefore, it is of great clinical significance to explore effective detection indicators to evaluate the outcomes of radiotherapy and detect changes in immune function.

Comprehensive treatment is recommended for patients with inoperable advanced esophageal cancer. In this group, patients with medical diseases, of older age, or those unwilling to synchronize radiotherapy and chemotherapy prior to radiotherapy were selected. This study aimed to understand the effects of radiotherapy alone on tumor markers and immune cells and to provide a preliminary understanding of the mode of radiotherapy combined with other treatments. For patients with long-term follow-up, comprehensive treatment was preferred; for patients with advanced stage esophageal cancer, chemotherapy or immunotherapy combined with traditional Chinese medicine was provided.

Serum tumor markers are expressed in malignant tumor cells or are generated after the tumor tissue is stimulated; they play a predictive role in the occurrence and development of tumors [5]. CEA is a broad-spectrum tumor marker, and its high level of expression in the serum has suggestive effects on esophageal cancer, breast cancer, gastric cancer, among others [6]. CA19-9 is a glycoprotein tumor marker, and its serum level is increased in various types of cancer [7]. CA72-4 dynamic monitoring can be performed to assist in the clinical

Table 2 Comparison of serum tumor markers

Groups	*n*	CEA (ng/mL)	CA19-9 (U/mL)	CA72-4 (U/L)	CYFRA21-1 (ng/mL)	SCC (ng/mL)
Observation group	87					
Before radiotherapy		4.25 ± 0.98[a]	15.25 ± 3.98[a]	5.64 ± 1.08[a]	4.45 ± 0.98[a]	2.45 ± 0.78[a]
After radiotherapy		2.37 ± 0.86[ab]	11.67 ± 2.86[ab]	3.67 ± 0.89[ab]	2.27 ± 0.56[ab]	0.87 ± 0.32[ab]
Control group	87	1.02 ± 0.73	7.52 ± 2.53	1.11 ± 0.33	1.02 ± 0.73	0.18 ± 0.05

a: Compared with the control group, $P < 0.05$; b: compared with that before radiotherapy, $P < 0.05$

Table 3 Comparison of T cell subsets

Groups	*n*	CD^{3+}	CD^{4+}	CD^{8+}	CD^{4+}/CD^{8+}	NK cells	Treg cells
Observation group	87						
Before radiotherapy		62.25 ± 7.58[a]	35.95 ± 8.13[a]	27.45 ± 6.42[a]	1.31 ± 0.29[a]	15.64 ± 2.23[a]	17.31 ± 2.29[a]
After radiotherapy		55.32 ± 8.75[ab]	32.21 ± 6.17[ab]	32.07 ± 7.69[ab]	0.81 ± 0.16[ab]	20.01 ± 0.53[ab]	10.48 ± 1.76[ab]
Control group	87	66.42 ± 6.01	40.32 ± 5.53	23.89 ± 2.57	1.69 ± 0.73	26.19 ± 2.93	3.37 ± 0.43

a: Compared with the control group, $P < 0.05$; b: compared with that before radiotherapy, $P < 0.05$

diagnosis of esophageal cancer and observation of the curative effects of treatments [8]. CYFR21-1 can be used as a tumor marker to predict the tumor changes in lung and breast cancers, and to effectively distinguish patients with cancer from those without cancer [9]. SCC antigen is a tumor-related glycoprotein fragment that can be detected in the tissues of patients with esophageal, lung, and cervical cancer. Although CA19-9, CA125, and CA72-4 have been used as tumor markers in the clinical detection of digestive tract adenocarcinoma, some studies have reported that esophageal SCC antigen can also be used as a tumor marker to help clinically judge the prognosis of esophageal SCC; therefore, this marker was also included in the current study [8]. In the observation group, the levels of CEA, CA19-9, CA72-4, CYFRA21-1, and SCC antigen after radiotherapy were lower than those before radiotherapy, which was evidently related to the response of the tumor to radiotherapy. The results from this study showed that the levels of CEA, CA19-9, CA72-4, CYFRA21-1, and SCC antigen in the observation group before and after radiotherapy were higher than those in the control group; moreover, radiotherapy could inhibit the proliferation of cancer cells. However, whether the effect on the level of tumor markers is different from that after chemotherapy has not been reported.

T cell subsets are the main markers that reflect the immune function of the body. The co-receptors of CD3+ T cells are common markers on the surface of T lymphocytes. CD4+ cells can induce the differentiation of lymphocytes and production of antibodies, to induce an immune response. CD8+ cells can act as inhibitory T cells, and often show cytotoxic activity, while inhibiting the secretion of antibodies. Tumor cells can directly activate or induce the increase in the expression of CD8+ T cells to inhibit cellular immune responses, resulting in a decrease in the ratio of CD3+, CD4+, and CD4+/CD8+ cells, and a decrease in the immune function of the body, thus allowing the tumor cells to evade immune surveillance and grow progressively [10]. The results of this study showed that the levels of CD3+, CD4+, and CD4+/CD8+ cells in the observation group before and after radiotherapy were lower, while the levels of CD8+ cells were higher than those in the control group; this indicated that the immune function of esophageal cancer patients was significantly reduced, which was consistent with the above results. In addition, the levels of CD3+, CD4+, CD4+/CD8+, and NK cells in the observation group after radiotherapy were lower, and the levels of CD8+ and Treg cells were higher than those before radiotherapy; this suggested that the immune function of patients with esophageal cancer was suppressed after undergoing radiotherapy, which may be caused by the significant non-selective killing effect of radiotherapy on normal tissues.

In conclusion, radiotherapy can effectively reduce the level of serum tumor markers in patients with esophageal cancer, but has adverse effects on the immune function of the body; hence, further clinical studies are needed to obtain a better clinical efficacy.

Conflicts of interest

The authors indicated no potential conflicts of interest.

References

1. Siegel RL, Miller KD, Fuchs HE, *et al*. Cancer Statistics, 2021. CA Cancer J Clin, 2021, 71: 7–33.
2. Watanabe M, Otake R, Kozuki R, *et al*. Recent progress in multidisciplinary treatment for patients with esophageal cancer. Surg Today, 2020, 50: 12–20.
3. Wang J, Liu F, Wu YY, *et al*. Survival outcomes of patients with cervical esophageal cancer who received definitive radiotherapy: a retrospective study conducted in a single institution. Oncol Transl Med, 2020, 6: 135–142.
4. Smyth EC, Lagergren J, Fitzgerald RC, *et al*. Oesophageal cancer. Nat Rev Dis Primers, 2017,3: 17048.
5. Zhang J, Zhu Z, Liu Y, *et al*. Diagnostic value of multiple tumor markers for patients with esophageal carcinoma. PLoS One, 2015, 10: e0116951.
6. Scarpa M, Noaro G, Saadeh L, *et al*. Esophageal cancer management: preoperative CA19.9 and CEA serum levels may identify occult advanced adenocarcinoma.World J Surg, 2015, 39: 424–432.
7. Silsirivanit A. Glycosylation markers in cancer. Adv Clin Chem, 2019, 89: 189–213.
8. Yang YC, Huang XZ, Zhou LK, *et al*. Clinical use of tumor biomarkers in prediction for prognosis and chemotherapeutic effect in esophageal squamous cell carcinoma. BMC Cancer, 2019, 19: 526.
9. Ma R, Xu H, Wu J, *et al*. Identification of serum proteins and multivariate models for diagnosis and therapeutic monitoring of lung cancer. Oncotarget, 2017, 8: 18901–18913.
10. Wang XB, Wu DU, Chen WP, *et al*. Impact of radiotherapy on immunological parameters, levels of inflammatory factors, and clinical prognosis in patients with esophageal cancer. J Radiat Res, 2019, 60: 353–363.

DOI 10.1007/s10330-021-0532-2
Cite this article as: Gao W, Liu XX, Ma HB. Effect of radiotherapy on tumor markers and serum immune-associated cells in patients with esophageal cancer. Oncol Transl Med, 2021, 7: 275–278.

Oncology and Translational Medicine
DOI 10.1007/s10330-021-0503-3

December 2021, Vol. 7, No. 6, P279–P285

ORIGINAL ARTICLE

Malnutrition as a predictor of prolonged length of hospital stay in patients with gynecologic malignancy: A comparative analysis*

Yongning Chen[1, 2], Runrong Li[2] (Co-first author), Li Zheng[2], Wenlian Liu[2], Yadi Zhang[2], Shipeng Gong[2] (✉)

[1] Department of Gynecology, The First Affiliated Hospital of Jinan University, Guangzhou 510630, China
[2] Department of Obstetrics and Gynecology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China

Abstract

Objective To explore the consistency of the Patient-generated Subjective Global Assessment (PG-SGA) and Nutritional Risk Screening-2002 (NRS-2002) for nutritional evaluation of patients with gynecologic malignancy and their predictive effect on the length of hospital stay (LOS).

Methods We recruited 147 hospitalized patients with gynecologic malignancy from Nanfang Hospital in 2017. Their nutritional status was assessed using the PG-SGA and NRS-2002. The consistency between the two assessments was compared via the Kappa test. The relationship between malnutrition and LOS was analyzed using crosstabs and Spearman's correlation.

Results The PG-SGA demonstrated that 66.7% and 54.4% of patients scoring ≥ 2 and ≥ 4 were malnourished, respectively. Furthermore, the NRS-2002 indicated that 55.8% of patients were at nutritional risk. Patients with ovarian cancer had a relatively high incidence of malnutrition. However, this was only significant for patients who scored ≥ 4 in the PG-SGA (P = 0.001 and P = 0.019 for endometrial carcinoma and cervical cancer, respectively). The PG-SGA and NRS-2002 showed good consistency in evaluating the nutritional status of patients with gynecologic malignancy (0.689, 0.643 for PG-SGA score ≥ 2, score ≥ 4 and NRS-2002, respectively). Both the scores of PG-SGA and NRS-2002 were positively correlated with LOS. Furthermore, prolonged LOS was higher in patients with malnutrition than in those with adequate nutrition.

Conclusion The PG-SGA and NRS-2002 shared a good consistency in evaluating the nutritional status of patients with gynecologic malignancy. Both assessments could be used as predictors of LOS.

Received: 14 June 2021
Revised: 21 August 2021
Accepted: 11 November 2021

Key words: malnutrition; patient-generated subjective global assessment; nutritional risk screening-2002; length of hospital stay; gynecologic malignancy

Malnutrition in hospitalized patients is a crucial issue and has been related to higher rates of morbidity and mortality [1]. Several studies have reported that the prevalence of malnutrition among those with cancer ranges from 31%-97% [2–3]. The association between malnutrition and hospitalization has been established for some diseases, in particular, malignant diseases [4]. Hence, it is important to identify malnourished patients. Knowing the patient's nutritional status may help improve patient outcomes during hospitalization. The assessment of nutritional status may be directed to several nutrition features as further discussed below.

The Patient-generated Subjective Global Assessment (PG-SGA) is a further modification of the SGA. The PG-SGA was developed specifically for cancer patients with a number of different conditions, and adapted by Ottery [5] for cancer patients. The PG-SGA as a patient's nutritional assessment has been used in various cancers, including colorectal cancer [6], head and neck cancer [7], esophageal cancer, and gynecological cancer [8]. It provides a numerical score, which translates as the level of nutrition intervention required. A higher score

✉ Correspondence to: Shipeng Gong. Email: gsp2103@163.com
* Supported by grants from the Guangdong Medical Research Fund (No. A2021054) and Nanfang Hospital President's Fund (No. 2019B019).

indicates a greater risk for malnutrition. Rodrigues *et al* [9]. showed that the PG-SGA could be used as a major predictor of prognosis and mortality in patients with gynecologic cancer. The Nutritional Risk Screening-2002 (NRS-2002) is a simple process for triaging at-risk patients indicated for nutrition interventions by assessing body mass index, appetite, weight loss, and severity of the disease. The NRS-2002 has been reported to effectively predict the nutritional risk for gynecologic patients. According to these studies, malnourished patients as determined by the NRS-2002, showed a significantly higher complication rate and longer LOS [4]. Malnutrition identified by the PG-SGA and NRS-2002 may reflect the patient prognosis and has been frequently used as an outcome measure [10]. Hence, the PG-SGA and NRS-2002 are useful for detecting the nutritional status of patients with cancer [11–12]. Additionally, the PG-SGA and NRS-2002 are considered the best validated tools for oncology patients [13]. However, studies have consistently shown the inadequacy of any single assessment tool in accurately determining a patient's nutritional status [14]. Therefore, we used the PG-SGA and NRS-2002 in combination for assessing patients with gynecologic malignancy. The assessment tools were applied before patients showed any signs of malnutrition and nutritional risk. To our knowledge, no study has yet evaluated the nutritional status of patients with gynecologic malignancy using both PG-SGA and NRS-2002.

The length of hospital stay (LOS) is used as the surrogate marker of a patient's recovery [15] and as an indicator of resource consumption [16]. Predicting LOS helps to minimize costs and maximize hospital resources [17] and facilitates an effective health care plan [15]. Guaitoli *et al* [15]. have shown that malnutrition as evaluated by the PG-SGA and the risk of malnutrition as evaluated by NRS-2002 are associated with a prolonged LOS.

The study aimed to evaluate the consistency of the PG-SGA and NRS-2002 in the nutritional evaluation of patients with gynecologic malignancy. The study also investigated if nutritional status as assessed by both can predict LOS.

Materials and methods

Participants and setting

All patients were recruited from Nanfang Hospital, Southern Medical University. The inclusion criterion was patients with histologically verified malignant gynecologic tumors. The exclusion criteria included patients who did not sign informed consent, patients who declined nutritional assessment, and patients younger than 18 years of age. From January 2017 to December 2017, 147 patients met the inclusion criteria. Patients were categorized according to their cancer sites: (1) cervical cancer (88 cases); (2) endometrial carcinoma (26 cases), and 3) ovarian cancer (33 cases).

Instruments

PG-SGA

The PG-SGA was used as previously reported to assess nutritional status 5, based on features of the physical examination and patient history. It consists of two sections including (1) a questionnaire about recent weight loss, food intake, and symptoms (such as nausea, diarrhea, and vomiting), and (2) information about the patient's disease and metabolic needs. Based on the global rating, those with a score < 2 were classified as well-nourished; a score between 2 and 4 as moderately malnourished or suspected of being malnourished; and ≥ 4 as severely malnourished 5. For analysis, each patient was classified as well-nourished (score < 2) or malnourished (score ≥ 2) 5. We also identified those with a malnutrition score of ≥ 4 to distinguish the patients who were in critical need of nutritional intervention [18].

NRS-2002

The NRS-2002 evaluates recent unintentional weight loss, appetite, and disease severity, and was recommended by the European Society of Parenteral and Enteral Nutrition as a preferred method of nutritional risk screening in hospital patients [19]. The final NRS-2002 score was between 0 and 7, and a score of ≥ 3 was classified as having nutritional risk [19]. The NRS-2002 examiners were not aware of the experimental test results at the time of the assessment.

Prolonged LOS

To explore whether the PG-SGA and NRS-2002 scores could predict the LOS of patients with gynecologic malignancy, prolonged LOS was defined as more than the median hospitalization day20, and the patients were divided into two categories, surgery and chemotherapy patients.

Data collection

By the time the patients were admitted to the hospital, our researchers had already obtained basic information from the nurses' station. Within 48 h after admission, we described the purpose of our study to potential patients and recruited those who were willing to participate in the study and provide informed consent. Subsequently, the investigators were trained by a nutritional specialist from our hospital and informed of relevant precautions when completing the PG-SGA and NRS-2002. Furthermore, whether the patients underwent surgery or chemotherapy, and their LOS, were determined after discharge from the hospital.

Statistical analysis

Measurement data were expressed as medians (*P25*, *P75*) and analyzed using the Mann-Whitney *U* test. The Kappa test was used to analyze the consistency of nutritional assessment via the PG-SGA and NRS-2002. Additionally, the receiver operator curve was plotted on the basis of the ability of the PG-SGA to evaluate the diagnostic value of NRS-2002. Crosstabs and Spearman's correlation were used to evaluate the relationship between malnutrition and LOS. Statistical analysis was performed using the SPSS statistics version 20.0 (IBM Corp., Armonk, NY). $P < 0.05$ was considered statistically significant.

Results

Patient characteristics

In our retrospective analysis, the patient's age, previous anti-tumor treatment, type of tumor, treatment methods, and the most recent LOS of the 147 recruited patients are shown in Tables 1–3.

Nutritional status assessed by the PG-SGA and NRS-2002

The PG-SGA median score was 4 (1, 7), and the NRS-2002 score was 3 (1, 3 Based on the PG-SGA); 98 patients (66.7%) scored ≥ 2, and 80 patients (54.4%) scored ≥ 4. In the NRS-2002 assessment, 82 patients (55.8%) scored ≥ 3 (Table 2).

The incidence of malnutrition in patients with cervical cancer, endometrial carcinoma, ovarian cancer, and other cancers such as gynecologic malignancy was assessed using the PG-SGA and NRS-2002 scores (Table 2). The results showed that patients with ovarian cancer have a relatively high incidence of malnutrition (78.8%, PG-SGA ≥ 2; 75.8%, PG-SGA ≥ 4; 69.7%, NRS-2002 ≥ 3). In contrast, patients with endometrial have the lowest incidence of malnutrition (53.8%, PG-SGA ≥ 2; 34.6%, PG-SGA ≥ 4; 42.3%, NRS-2002 ≥ 3). Only the incidence of malnutrition (PG-SGA≥ 4) was significantly different between patients with ovarian cancer and those with endometrial carcinoma or cervical cancer ($P = 0.001$ and $P = 0.019$, respectively).

Consistency between the PG-SGA and NRS-2002

The Kappa test was used to assess the consistency of the two instruments for assessing malnutrition. When the PG-SGA score ≥2 was set as the standard for a diagnosis of malnutrition, we found that the positive rate of PG-SGA was significantly consistent with the NRS-2002 for all patients ($k = 0.689$) and patients with cervical cancer ($k = 0.626$), endometrial carcinoma ($k = 0.772$), or ovarian

Table 1 Basic clinical characteristics of patients

Item				
Previous anti-tumor treatment	Untreated (*n* = 52)	Neoadjuvant treatment (*n* = 30)	Surgery (*n* = 20)	Postoperative chemotherapy (*n* = 45)
Treatment received in our hospital	Chemotherapy patients (*n* = 82)		Surgical Patients (*n* = 65)	
Length of hospital stay (days)	4 (3, 6)		10 (9, 13)	
Age (years)	47.5 (41.75, 55)		48 (40.5, 54.5)	

Table 2 The incidence of malnutrition in gynecologic malignant patients according to the score of PG-SGA and NRS-2002 [*n* (%)]

Reference method		Total patients (*n* = 147)	Cervical cancer patients (*n* = 88)	Endometrial carcinoma patients (*n* = 26)	Ovarian cancer patients (*n* = 33)
PG-SGA ≥ 2	Well	49 (33.4)	30 (34.1)	12 (46.2)	7 (21.2)
	Malnutrition	98 (66.7)	58 (65.9)	14 (53.8)	26 (78.8)
PG-SGA ≥ 4	Well	67 (45.6)	42 (47.7)	17 (65.4)	8 (24.2)
	Malnutrition	80 (54.4)	46 (52.3)	9 (34.6)	25 (75.8)
NRS-2002 ≥ 3	Well	65 (44.2)	40 (45.5)	15 (57.7)	10 (30.3)
	Malnutrition	82 (55.8)	48 (54.5)	11 (42.3)	23 (69.7)

Table 3 Consistency of NRS-2002 and PG-SGA (score ≥ 2 or 4) in gynecologic malignant patients (*k* value)

Reference method / Pathological classification	NRS-2002 ≥ 3			
	Total patients (*n* = 147)	Cervical cancer patients (*n* = 88)	Endometrial carcinoma patients (*n* = 26)	Ovarian cancer patients (*n* = 33)
PG-SGA ≥ 2	0.689	0.626	0.772	0.765
PG-SGA ≥ 4	0.643	0.589	0.516	0.848

cancer (k = 0.765) (Table 3). When the PG-SGA score ≥4 was set as the standard for the diagnosis of malnutrition, the result was similar for all patients (k = 0.643) and patients with cervical cancer (k = 0.589), endometrial carcinoma (k = 0.516), or ovarian cancer (k = 0.848) (Table 3).

When a PG-SGA score ≥ 2 was set as the "gold standard" to calculate the sensitivity and specificity of the NRS-2002 score ≥ 3, the sensitivity was 80.6% and the specificity 93.9% for all the patients (Fig. 1). When a PG-SGA score ≥ 4 was set as the "gold standard" to calculate the sensitivity and specificity of a NRS-2002 score ≥3, the sensitivity was 85.0% and the specificity was 79.1% for all the patients (Fig. 1).

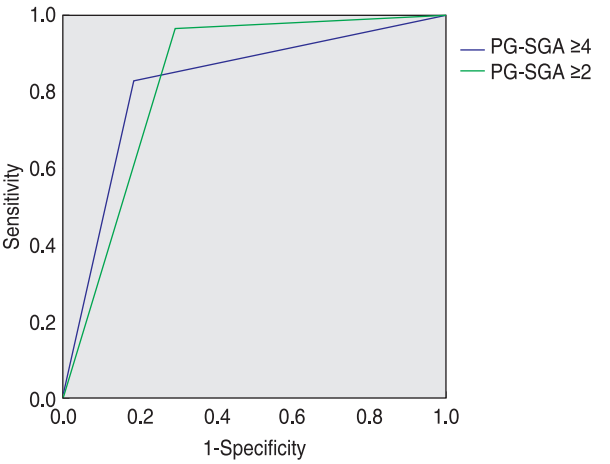


Fig. 1 Receiver-operating characteristic curve comparing NRS-2002 (score ≥ 3) to PG-SGA (score ≥ 2) or PG-SGA (score ≥ 4) in gynecologic malignant patients at admission

Association between nutritional scores and LOS

The nutritional scores of PG-SGA and NRS-2002 were positively correlated with LOS in the surgery group (Table 4) and chemotherapy group (Table 5). The specific manifestation presented a significantly higher proportion of prolonged LOS in malnourished patients than in those with normal nutritional status as assessed by either the PG-SGA or NRS-2002.

Discussion

Nutritional screening is the first step in developing an effective nutritional plan during admission. In our study, we explored the value of using the PG-SGA and NRS-2002 in assessing nutritional status and their predictive effects on LOS in a series of 147 gynecologic malignancy patients.

In this study, based on the PG-SGA, over 66.7% (PG-SGA score ≥ 2) and 54.4% (PG-SGA score ≥ 4) of patients had poor nutritional status. Using the NRS-2002 (score ≥ 3), we found 55.8% of patients at nutritional risk. The above results demonstrated that hospitalized patients with gynecologic malignancy had a substantial malnutrition or nutritional risk. These findings concurred with other studies, in which the prevalence of malnutrition was 62.4% as evaluated by the PG-SGA in those with gynecologic cancer [9, 18]. According to the PG-SGA, only 23.7% were classified as malnourished [21]. Moreover, using the NRS-2002, 35.8% were identified as

Table 4 Comparison of Prolonged LOS in Surgical Patients Evaluated by PG-SGA and NRS-2002 [*n* (%)]

Reference method		LOS		Spearman's Coefficients	*P* value
		Normal LOS	Prolonged LOS		
PG-SGA ≥ 2	Well (n = 29)	23 (79.3)	6 (20.7)	0.666	< 0.001
	Malnutrition (n = 36)	12 (33.3)	24 (66.7)		
PG-SGA ≥ 4	Well (n = 39)	31 (79.5)	8 (20.5)		< 0.001
	Malnutrition (n = 26)	4 (15.4)	22 (84.6)		
NRS-2002 ≥ 3	Well (n = 31)	27 (87.1)	4 (12.9)	0.071	< 0.001
	Malnutrition (n = 34)	8 (23.5)	26 (76.5)		

Table 5 Comparison of Prolonged LOS in Chemotherapy Patients Evaluated by PG-SGA and NRS-2002 [*n* (%)]

Reference Method		LOS		Spearman's Coefficients	*P* value
		Normal LOS	Prolonged LOS		
PG-SGA ≥2	Well (n = 20)	16 (80.0%)	4 (20.0%)	0.734	< 0.001
	Malnutrition (n = 62)	23 (37.1%)	39 (62.9%)		
PG-SGA ≥4	Well (n = 28)	23 (82.1%)	5 (17.9%)		< 0.001
	Malnutrition (n = 54)	16 (29.6%)	38 (70.4%)		
NRS-2002 ≥3	Well (n = 34)	28 (82.4%)	6 (17.6%)	0.728	< 0.001
	Malnutrition (n = 48)	11 (22.9%)	37 (77.1%)		

having nutritional risk [4]. Malnutrition or nutritional risk is also related to perioperative fasting, surgical trauma stress responses, increased metabolism, and decreased intake caused by radiotherapy and chemotherapy [12, 22–23]. In our study, not only patients who were previously untreated, but also those who had received surgery and (or) chemotherapy were included. This may be the main reason for the higher malnutrition rate or nutritional risk in this study.

The prevalence of malnutrition may be affected by different evaluation tools and tumor sites. Orell-Kotikangas *et al*. found that 69.5% of patients with multiple types of malignant tumors had nutritional risks as evaluated by NRS-2002 [24]. Another study reported 20-88% of patients with gynecological cancer had some degree of malnutrition 23. We also observed malnutrition in patients with malignant gynecologic tumors in different sites. We found that patients with ovarian cancer had a relatively high incidence of malnutrition, while patients with endometrial carcinoma had a relatively low incidence of malnutrition. Rodrigues *et al*. 18 also found that patients with endometrial carcinoma showed a significantly lower median score compared to those with cervical and ovarian tumors. Additionally, Zorlini *et al*. reported a significantly higher prevalence of malnutrition in patients with endometrium cancer as opposed to those with cancer at other sites [25]. Laky and colleagues found that patients with ovarian cancer were more susceptible to nutritional status alterations, whereas those with endometrial and uterine cancers comprise a less predisposed group to such alterations [8]. This discrepancy may be related to (1) differences in sample size; (2) regional differences resulting in different dietary patterns that may influence the population nutritional status; (3) complications caused by cancer; and (4) different previous treatment regimens. Furthermore, the rate of malnutrition in patients with cancer seems to depend on multiple factors, including tumor sites, treatment, staging, and histology.

A general concordance and agreement (k value = 0.523) were observed between the PG-SGA and NRS-2002 in the diagnosis of malnutrition among patients with cervical cancer [13]. In our study, we also detected a high concordance and agreement (k statistic was 0.689 and 0.643 when the PG-SGA score was ≥ 2 and 4, respectively) between the two assessments when used for patients with gynecologic malignancy. Concordance between the PG-SGA and NRS-2002 was also supported by Helena in a study of patients with head and neck cancer [24]. The concordance between the PG-SGA and NRS-2002 was also observed in different gynecologic tumor sites. Despite the lack of homogeneity studies, both the PG-SGA and NRS-2002 are currently recommended for nutritional screening of patients with gynecologic malignancy. Although there are other nutrition assessment tools, there is a lack of consensus on which tool is the most suitable for patients with malignant tumors. Our findings demonstrated a high concordance between the two assessment tools and supported the use of the NRS-2002 and PG-SGA in patients with gynecologic cancer.

Good nutritional screening tools should show good specificity and sensitivity [14]. In our study, the NRS-2002 cut-off score of ≥ 3 compared with the PG-SGA showed high specificity and sensitivity in patients with gynecologic cancer. As mentioned before, this concurs with the findings from a large oncology study in patients with head and neck squamous cell carcinoma by Helena *et al* [24]. In particular, a PG-SGA nutritional status score of 7.5 predicted febrile neutropenia, with a sensitivity of 100% and a specificity of 60% in patients with gynecologic cancer, suggesting that these patients may have a higher baseline PG-SGA score [26]. A higher baseline provides a more accurate identification of malnourished patients. Our results showed that a NRS-2002 cut-off score of ≥ 3 and PG-SGA score≥ 2 or 4 are suitable for predicting the nutritional status of patients with gynecologic cancer.

Many nutritionally at-risk patients present with complications during admission. The effect of poor nutritional status on early readmissions and the development of complications have been previously demonstrated [27–28]. We also showed a positive correlation between LOS and compromised nutritional status as per the PG-SGA or NRS-2002. Further analysis revealed that a prolonged LOS is more common in patients with nutrition risk or those who are undernourished than patients with a good nutritional status. The PG-SGA has been validated as an assessment of nutritional status, which can be used to indicate a longer length of stay in patients with multiple types of cancer [15, 29]. A longer LOS was also observed in surgical patients with nutritional risks as identified by the NRS-2002 [30]. Overall, the LOS increased significantly in cancer patients with severe malnutrition and nutritional risk as identified by the PG-SGA or NRS-2002 [10, 31]. In patients with gynecologic malignancy, an association between malnutrition and LOS based on the PG-SGA score was found by Laky and colleagues [20]. In that study, the medial hospitalization time of patients with malnutrition as assessed by the NRS-2002 (score ≥ 3) was increased from 7 to 10 days [4]. The PG-SGA and NRS-2002 shared similar validity and good consistency in predicting the LOS of patients with gynecologic malignancy. This suggests that they could be used for nutritional screening at the time of admission of patients with gynecologic malignancy. The PG-SGA and NRS-2002 can be completed in a few minutes, unlike the Mini Nutritional Assessment, which is the most time-consuming tool (410 min) [32]. However, LOS is influenced by many factors other than nutritional

status, such as illness severity, disease, and age. Therefore, related research about LOS may have been biased because these studies did not address all the potential contributing factors [33].

It may be necessary for trained physicians to improve their competency in using the PG-SGA properly. NRS-2002 requires less training and is more convenient than the PG-SGA. In addition, several patient-related factors are influential to LOS, such as diagnosis, age, and hospital procedures such as elective surgeries. Therefore, further studies should 1) increase the number of research samples, 2) reduce population heterogeneity, and 3) apply the same treatment regimen as for other patients with cancer and specifically define the associations with age, complications, mortality, costs, and so on, in patients with gynecologic cancer.

In summary, our findings suggest that a high prevalence of moderate and severe malnutrition or nutritional risks are common among patients with gynecologic malignancy based on evaluations using the PG-SGA and NRS-2002. Furthermore, the PG-SGA and NRS-2002 correlated with each other. Either assessment can be used to predict prolonged LOS in patients with gynecologic malignancy.

Ethics approval

An ethics committee approval was obtained from Nanfang Hospital.

Conflicts of interest

The authors indicated no potential conflicts of interest.

References

1. Bechard LJ, Duggan C, Touger-Decker R, *et al*. Nutritional status based on body mass index is associated with morbidity and mortality in mechanically ventilated critically Ill children in the PICU. Crit Care Med, 2016, 44: 1530–1537.
2. Malihi Z, Kandiah M, Chan YM, *et al*. Nutritional status and quality of life in patients with acute leukaemia prior to and after induction chemotherapy in three hospitals in Tehran, Iran: a prospective study. J Hum Nutr Diet, 2013, 26 Suppl 1: 123–131.
3. Zhang L, Lu Y, Fang Y. Nutritional status and related factors of patients with advanced gastrointestinal cancer. Br J Nutr, 2014, 14, 111: 1239–1244.
4. Hertlein L, Kirschenhofer A, Fürst S, *et al*. Malnutrition and clinical outcome in gynecologic patients. Eur J Obstet Gynecol Reprod Biol, 2014, 174: 137–140.
5. Ottery FD. Definition of standardized nutritional assessment and interventional pathways in oncology. Nutrition, 1996, 12: S15–19.
6. Daudt HM, Cosby C, Dennis DL, *et al*. Nutritional and psychosocial status of colorectal cancer patients referred to an outpatient oncology clinic. Support Care Cancer, 2012, 20: 1417–1423.
7. Paccagnella A, Morello M, Da Mosto MC, *et al*. Early nutritional intervention improves treatment tolerance and outcomes in head and neck cancer patients undergoing concurrent chemoradiotherapy. Support Care Cancer, 2010, 18: 837–845.
8. Laky B, Janda M, Bauer J, *et al*. Malnutrition among gynaecological cancer patients. Eur J Clin Nutr, 2007, 61: 642–646.
9. Rodrigues CS, Lacerda MS, Chaves GV. Patient Generated Subjective Global Assessment as a prognosis tool in women with gynecologic cancer. Nutrition, 2015, 31: 1372–1378.
10. Mendes J, Alves P, Amaral TF. Comparison of nutritional status assessment parameters in predicting length of hospital stay in cancer patients. Clin Nutr, 2014, 33: 466–470.
11. Kwag SJ, Kim JG, Kang WK, *et al*. The nutritional risk is a independent factor for postoperative morbidity in surgery for colorectal cancer. Ann Surg Treat Res, 2014, 86: 206–211.
12. Dong W, Liu X, Zhu S, *et al*. Selection and optimization of nutritional risk screening tools for esophageal cancer patients in China. Nutr Res Pract, 2020, 14: 20–24.
13. Tian M, Fu H, Du J. Application value of NRS2002 and PG-SGA in nutritional assessment for patients with cervical cancer surgery. Am J Transl Res, 2021, 13: 7186–7192.
14. Zhang Z, Wan Z, Zhu Y, *et al*. Prevalence of malnutrition comparing NRS2002, MUST, and PG-SGA with the GLIM criteria in adults with cancer: A multi-center study. Nutrition, 2021, 83: 111072.
15. Guerra RS, Fonseca I, Pichel F, *et al*. Usefulness of six diagnostic and screening measures for undernutrition in predicting length of hospital stay: a comparative analysis. J Acad Nutr Diet, 2015, 115: 927–938.
16. Zoungrana B, Sawadogo PS, Somda NS, *et al*. Performance et coût de la prise en charge de la malnutrition aiguë sévère avec complications à Kaya, Burkina Faso [Effectiveness and cost of management of severe acute malnutrition with complications in Kaya, Burkina Faso]. Pan Afr Med J, 2019, 34: 145. French.
17. Nguyen C, Hernandez-Boussard T, Davies SM, *et al*. Cleft palate surgery: an evaluation of length of stay, complications, and costs by hospital type. Cleft Palate Craniofac J, 2014, 51: 412–419.
18. Rodrigues CS, Chaves GV. Patient-generated Subjective Global Assessment in relation to site, stage of the illness, reason for hospital admission, and mortality in patients with gynecological tumors. Support Care Cancer, 2015, 23: 871–879.
19. Kondrup J, Rasmussen HH, Hamberg O, *et al*. Ad Hoc ESPEN Working Group. Nutritional risk screening (NRS 2002): a new method based on an analysis of controlled clinical trials. Clin Nutr, 2003, 22: 321–336.
20. Laky B, Janda M, Kondalsamy-Chennakesavan S, *et al*. Pretreatment malnutrition and quality of life – association with prolonged length of hospital stay among patients with gynecological cancer: a cohort study. BMC Cancer, 2010, 10: 232.
21. Laky B, Janda M, Cleghorn G, *et al*. Comparison of different nutritional assessments and body-composition measurements in detecting malnutrition among gynecologic cancer patients. Am J Clin Nutr, 2008, 87: 1678–1685.
22. Weycker D, Li X, Tzivelekis S, *et al*. Burden of chemotherapy-induced febrile neutropenia hospitalizations in US clinical practice, by use and patterns of prophylaxis with colony-stimulating factor. Support Care Cancer, 2017, 25: 439–447.
23. Obermair A, Simunovic M, Isenring L, *et al*. Nutrition interventions in patients with gynecological cancers requiring surgery. Gynecol Oncol, 2017, 145: 192–199.
24. Orell-Kotikangas H, Österlund P, Saarilahti K, *et al*. NRS-2002 for pre-treatment nutritional risk screening and nutritional status assessment in head and neck cancer patients. Support Care Cancer, 2015, 23: 1495–1502.

25. Zorlini R, Akemi Abe Cairo A, Salete Costa Gurgel M. Nutritional status of patients with gynecologic and breast cancer. Nutr Hosp, 2008, 23: 577–583.
26. Phippen NT, Lowery WJ, Barnett JC, *et al*. Evaluation of the Patient-Generated Subjective Global Assessment (PG-SGA) as a predictor of febrile neutropenia in gynecologic cancer patients receiving combination chemotherapy: a pilot study. Gynecol Oncol, 2011, 123: 360–364.
27. Schneider S, Armbrust R, Spies C, *et al*. Prehabilitation programs and ERAS protocols in gynecological oncology: a comprehensive review. Arch Gynecol Obstet, 2020, 301: 315–326.
28. Bisch S, Nelson G, Altman A. Impact of Nutrition on Enhanced Recovery After Surgery (ERAS) in Gynecologic Oncology. Nutrients. 2019, 11: 1088.
29. Hsueh SW, Liu KH, Hung CY, *et al*. Predicting postoperative events in patients with gastric cancer: a comparison of five nutrition assessment tools. In Vivo, 2020, 34: 2803–2809.
30. Almeida AI, Correia M, Camilo M, *et al*. Length of stay in surgical patients: nutritional predictive parameters revisited. Br J Nutr, 2013, 109: 322–328.
31. Santos IM, Mendes L, Carolino E, *et al*. Nutritional status, functional status, and quality of life – What is the impact and relationship on cancer patients? Nutr Cancer, 2020, 29: 1–14.
32. Velasco C, García E, Rodríguez V, *et al*. Comparison of four nutritional screening tools to detect nutritional risk in hospitalized patients: a multicentre study. Eur J Clin Nutr, 2011, 65: 269–274.
33. van Bokhorst-de van der Schueren MA, Guaitoli PR, Jansma EP, *et al*. Nutrition screening tools: does one size fit all? A systematic review of screening tools for the hospital setting. Clin Nutr, 2014, 33: 39–58.

DOI 10.1007/s10330-021-0503-3
Cite this article as: Chen YN, Li RR, Zheng L, *et al*. Malnutrition as a predictor of prolonged length of hospital stay in patients with gynecologic malignancy: A comparative analysis. Oncol Transl Med, 2021, 7: 279–285.

Oncology and Translational Medicine
DOI 10.1007/s10330-021-0500-0
December 2021, Vol. 7, No. 6, P286–P293

ORIGINAL ARTICLE

GFPT2 pan-cancer analysis and its prognostic and tumor microenvironment associations*

Jiachen Zhang[1, 2], Ting Wang[1, 3], Siang Wei[1, 2], Shujia Chen[1, 2], Juan Bi [1](✉)

[1] Department of Pharmacy, The First Affiliated Hospital, Naval Medical University, Shanghai 200002, China
[2] Jinzhou Medical University, Liaoning 121000, China
[3] Department of Pharmacy, The First Affiliated Hospital of Anhui Medical University, Anhui 230022, China

Abstract

Objective Glutamine fructose-6-phosphate transaminase 2 (*GFPT2*) is involved in a wide range of biological functions in human cancer. However, few studies have comprehensively analyzed the correlation between *GFPT2* and different cancer prognoses and tumor microenvironments (TMEs).

Methods We evaluated the expression level and prognostic value of *GFPT2* using updated public databases and multiple comprehensive bioinformatics analysis methods and explored the relationship between *GFPT2* expression and immune infiltration, immune neoantigens, tumor mutational burden (TMB), and microsatellite instability in pan-cancer.

Results *GFPT2* was highly expressed in five cancers. *GFPT2* expression correlates with the prognosis of several cancers from The Cancer Genome Atlas (TCGA) and is significantly associated with stromal and immune scores in pan-cancer. High *GFPT2* expression in BLCA, BRCA, and CHOL was positively correlated with the infiltration of immune cells, such as B-cells, CD4+ T, CD8+ T cells, dendritic cells, neutrophils, and macrophages.

Conclusion High *GFPT2* expression may modify the outcomes of patients with BLCA, BRCA, or CHOL cancers by increasing immune cell infiltration. These findings may provide insights for further investigation into *GFPT2* as a potential target in pan-cancer.

Received: 10 June 2021
Revised: 21 August 2021
Accepted: 25 September 2021

Key words: Glutamine fructose-6-phosphate transaminase 2 (*GFPT2*); pan-cancer, prognosis, immune, microenvironment

According to the global cancer statistics in 2018, it was estimated that there would be 18.1 million new cancer cases and 9.6 million cancer deaths in 2018 [1]. The World Health Organization estimated that the number of cancer cases worldwide is likely to increase by 60% over the next 20 years [2]. Cancer incidence and mortality is rapidly increasing worldwide. The reasons are complex, but they reflect population aging and growth and changes in the prevalence and distribution of the major cancer risk factors associated with socioeconomic development [3]. Cancer is associated with various genes, and the accumulation of molecular modifications in the somatic genome is fundamental to cancer progression. Traditional therapies, including surgery, radiotherapy, and chemotherapy, remain the first treatments for most cancer patients. However, breakthroughs in targeted therapy and immune checkpoint blockade therapy have significantly improved cancer patient survival [4–6].

Thus far, many studies have investigated how microenvironments and immune cell infiltration contribute to cancer development. Cancer tissue contains not only cancer cells but also non-cancer cells, such as stromal and immune cells [7]. Non-cancer cells dilute cancer cell purity and play an important role in cancer biology [8]. Under different purity conditions, the generally accepted prediction index is no longer valid. Therefore, the composition and proportion of stromal cells and immune cells in a tumor may determine the clinical prognosis of patients. In colon cancer, low tumor purity is associated with poor prognosis because of the high mutation frequency of key pathways and purity-related microenvironment changes [9]. In these biological processes, immune-related genes may affect cancer patient prognosis by affecting the abundance of infiltrating

✉ Correspondence to: Juan Bi. E-mail: bjfclcys@163.com
*Supported by a grant from the National Natural Science Foundation of China (No. 81700256).

immune cells. Therefore, it is necessary to identify the immune-related genes of a specific tumor phenotype to clarify its exact mechanism and find biomarkers or targets for tumor diagnosis and treatment.

As a key factor in the hexosamine biosynthesis signaling pathway, glutamine fructose-6-phosphate transaminase 2 (*GFPT2*) protein phosphorylation promotes glycosylation of downstream protein O-GlcNAc and mediates various physiological and pathological cell activities. Recent studies have confirmed that GFPT family proteins play an important role in the occurrence and development of various cancers. However, the relationship between *GFPT2* and cancer immune cells is still unclear, limiting our understanding of the specific function of *GFPT2* in the occurrence and development of cancer and the implementation of therapeutic measures.

In the current study, we comprehensively analyzed the prognostic value of *GFPT2* in pan-cancer via multiple databases, including the GTEx, Cancer Genome Atlas (TCGA), TIMER, and PrognoScan. We also evaluated the potential association between *GFPT2* expression and the tumor microenvironment (TME). Furthermore, we examined the relationship between *GFPT2* expression and immune score and matrix score in pan-cancer. We comprehensively evaluated the expression level and prognostic value of *GFPT2* based on multiple public resources and integrated bioinformatics analysis.

Materials and methods

Analyzing *GFPT2* expression in various cancers

Considering there are a relatively small number of normal samples in TCGA database, we analyzed the expression differences of 27 kinds of tumors based on the data of normal tissues in the GTEx database and TCGA tumor tissues. In the figure, * indicates $P < 0.05$, ** indicates $P < 0.01$, and *** indicates $P < 0.001$.

Prognostic analysis of *GFPT2* expression in cancer patients

We used gene expression profile data to analyze gene expression and prognosis in tumors. Considering that there may be non-tumor death factors during follow-up, the researchers analyzed the relationship between gene expression and prognosis of disease-specific survival (DSS) in 33 tumors of TCGA database and made a forest map and Kaplan–Meier (KM) curve of tumor prognosis.

Gene expression and immune relationship in various tumor cells

Tumor-infiltrating lymphocytes are an independent predictor of the status and presence of cancer in sentinel lymph nodes. We studied whether gene expression is related to the level of immune cell infiltration in different types of cancer. Firstly, we downloaded the score data of six kinds of immune infiltrating cells in 33 types of cancer from the TIMER database, and the correlations between gene expression and immune cell score were analyzed, respectively. Three tumors in which GFPT2 is most closely related to immune cells are demonstrated. Then, the immune scores and matrix scores of various tumor samples were analyzed by using the R software package estimate. The relationship between gene expression and immune score, gene expression, and matrix score the most significant first three tumors were observed in 33 tumors. Finally, we collected more than 40 common immune checkpoint genes, analyzed the relationship between gene expression and immune checkpoints, extracted these immune checkpoint genes, respectively, and calculated the correlation with the target gene's expression. * represents $P < 0.05$, ** represents $P < 0.01$, and *** represents $P < 0.001$.

Results

Gene expression in pan-cancer

Combined with the data analysis of normal tissue in the GTEx database and TCGA tumor tissue, *GFPT2* was differentially expressed in 27 types of cancer. Among them, *GFPT2* was highly expressed in CHOL, GBM, HNSC, LAML, LGG, and PAAD but was significantly lower in BLCA, BRCA, CESC, COAD, ESCA, KICH, LUAD, LUSC, OV, PRAD, SKCM, TGCT, THCA, UCEC, and UCS (Fig. 1). Therefore, *GFPT2* can be used as a biomarker to detect these 21 kinds of cancer.

Prognosis analysis of genes in pan-cancer

Prognosis analysis (Fig. 2) showed that there was a significant correlation between gene expression and cancer hazard ratios (HRs), including KICH [1.12 (1.02–1.21), $P = 1.2e-02$], KIRC [1.03 (10.2–1.04), $P = 2.5e-19$], OV [1.03 (1.01–1.06), $P = 1.4e-01$], and THCA [1.07 (1.02–1.12), $P = 6.4e-03$]. The HR value of *GFPT2* in pan-cancer and the prognosis analysis showed that the ratio of *GFPT2* to the risk function of renal cancer was higher, which may be related to the larger energy demands of the kidneys, and *GFPT2* played an important role in the regulation of glucose metabolism. According to the prognosis survival curve, high and low *GFPT2* expression was related to prognosis survival period intervention in many cancer patients. The results of the survival curve showed that *GFPT2* over- and under-expression was significant in 12 cancers, namely, BLCA, GBM, KICH, KIRC, KIRP, LUAD, LUSC, MESO, OV, THGA, UCEC, and UVM (Fig. 3).

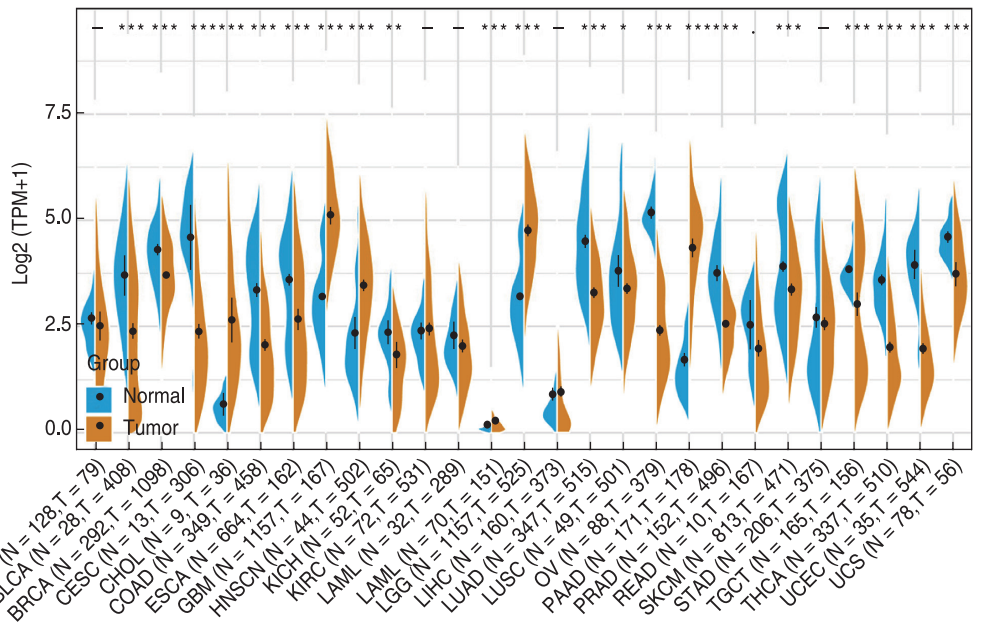


Fig. 1 *GFPT2* expression in normal and cancer tissues

	HR	*P* value
ACC	1 (0.97~1.04)	8.1e–01
BLCA	1.01 (1~1.01)	2.6e–03
BRCA	1 (0.98~1.02)	8.6e–01
CESC	1 (0.97~1.04)	7.7e–02
CHOL	1 (0.96~1.04)	1.0e+00
COAD	1.01 (0.99~1.03)	5.0e–01
DLBC	0.88 (0.73~1.08)	2.2e–01
ESCA	1.01 (0.98~1.05)	4.1e–01
GBM	1 (1~1.01)	3.1e–02
HNSC	1.01 (1~1.02)	2.1e–01
KICH	1.12 (1.02~1.21)	1.2e–02
KIRC	1.03 (1.02~1.1)	2.5e–19
KIRP	1.05 (1.01~1.1)	1.5e–02
LAML	NA (NA~NA)	
LGG	1 (1~1.01)	4.8e–01
LIHC	1.02 (0.98~1.07)	3.2e–01
LUAD	1.02 (1.01~1.03)	5.5e–07
LUSC	1.02 (1~1.03)	7.9e–03
MESO	1 (1~1)	9.9e–03
OV	1.03 (1.01~1.06)	3.9e–03
PAAD	1.01 (1~1.02)	1.4e–01
PCPG	1 (0.79~1.25)	9.8e–01
PRAD	0.75 (0.49~1.15)	1.9e–01
READ	0.95 (0.84~1.09)	4.8e–01
SARC	1 (1~1)	2.3e–01
SKCM	1 (0.99~1)	3.4e–01
STAD	1.01 (0.99~1.02)	3.1e–01
TGCT	1.01 (0.94~1.09)	7.7e–01
THCA	1.07 (1.02~1.12)	6.4e–03
THYM	0.98 (0.83~1.17)	8.6e–01
UCEC	1.02 (1.01~1.03)	2.0e–05
UCS	1.02 (0.99~1.05)	2.5e–01
UVM	0.89 (0.8~1)	4.7e–02

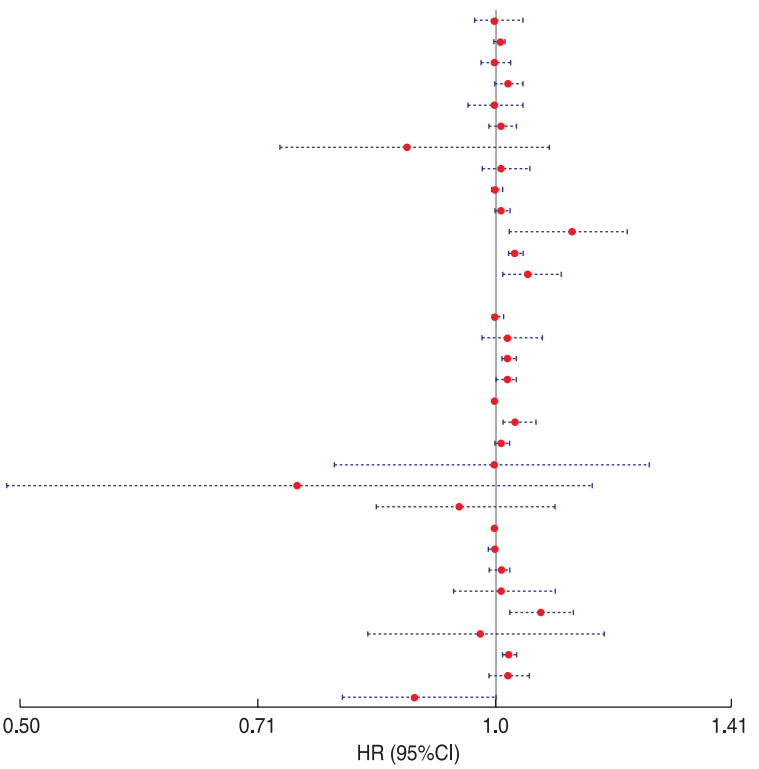


Fig. 2 Correlation between *GFPT2* expression and cancer risk

Relationship between gene expression and immunity in various tumors

The immune system allows the human body to defend against foreign pathogens. It can identify adverse agents and attack and eliminate pathogenic microorganisms, such as bacteria, viruses, and molds [10]. Studies have shown that some cancer cells can actively induce immune cells to secrete growth factors, thus, promoting cancer cell growth and metastasis themselves [11–12]. Based on the correlation between gene expression and immune cells, the latter mainly including B, CD4+, CD8+, and dendritic cells and neutrophils and macrophages, BLCA, BRCA,

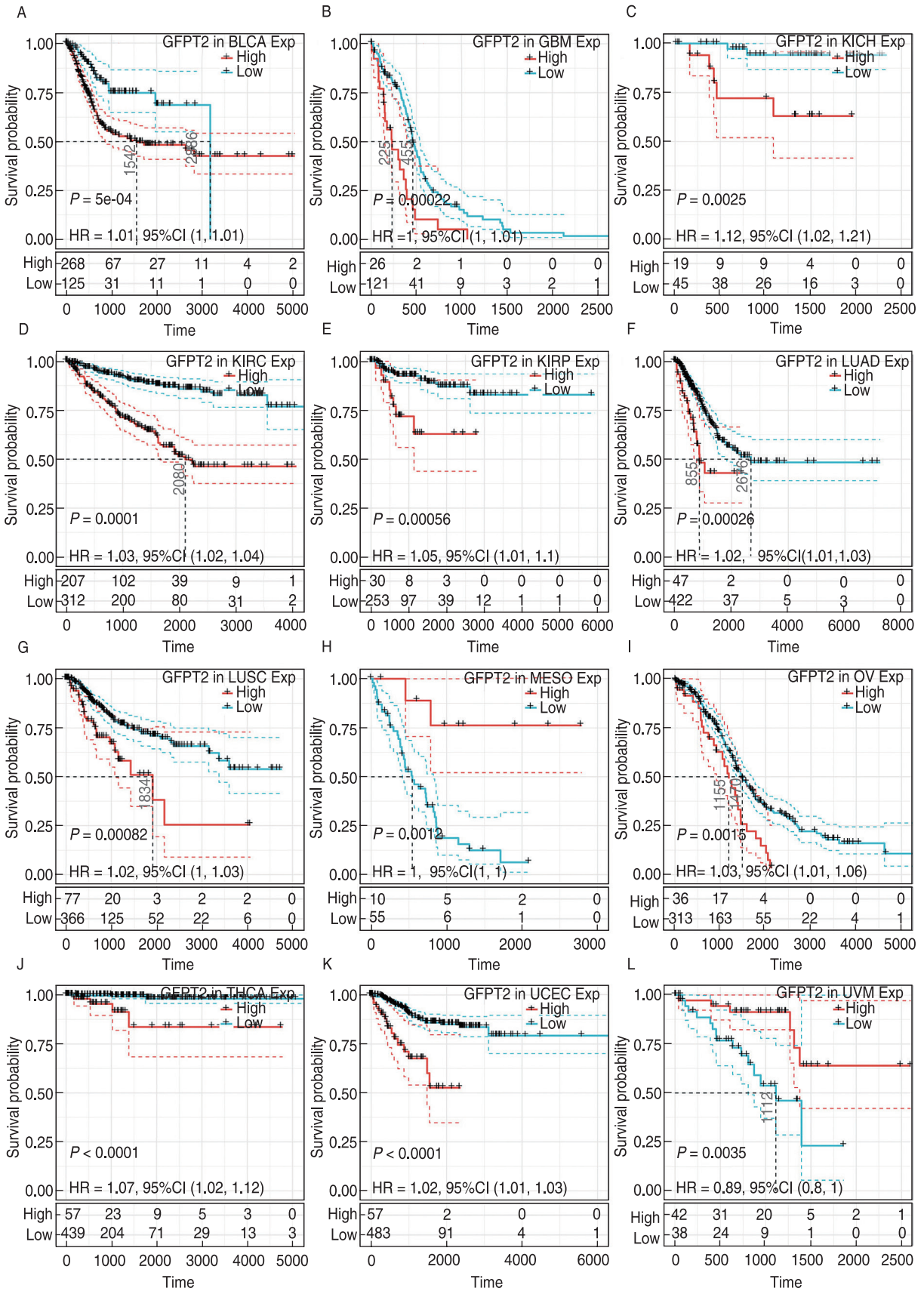


Fig. 3 ROC curve analysis of GFPT2 expression in pan-cancer. Fig. (a–l) showed the relationship between GFPT2 expression and survival possibility of patients with bladder urothelial carcinoma (BLCA), glioblastoma multiforme (GBM), kidney chromophobe (KICH) kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), mesothelioma (MESO), ovarian serous cystadenocarcinoma (OV), thyroid carcinoma (THCA), uterine corpus endometrial carcinoma (UCEC) and uveal melanoma (UVM).

and CHOL were the most significantly correlated tumors. The Spearman's correlation coefficients between *GFPT2* and the above three types of cancer were as follows: B cells (R = -0.153, P = 0.0019; R = 0.28, P = 4.06e-21; R = 0.56, P = 0.000475), CD4+ cells (R = 0.452, P = 6.64e-22; R = 0.492, P = 1.01e-67; R = 0.576, P = 0.000304), CD8+ cells (R = 0.496, P = 1.14e-26; R = 0.466, P = 6.28e-60; R = 0.504, P = 0.00196), dendritic cells (R = 0.654, P = 0; R = 0.564, P = 6.38e-93; R = 0.673, P = 1.19e-05), neutrophils (R = 0.615, P = 0; R = 0.547, P = 1.58e-86; R = 0.689, P = 6.41e-06), and macrophages (R = 0.527, P = 1.58e-30; R = 0.498, P = 1.41e-69; R = 0.57, P = 0.000365) (Fig. 4).

R software package estimate analyzed the immune scores and matrix scores of the gene and tumor samples. The three most significant tumors were BLCA (R = 0.826, P = 0), CESC (R = 0.504, P = 0), and COAD (R = 0.885, P = 0; Fig. 5).

Discussion

Using independent data sets from the GTEx and TCGA, we investigated *GFPT2* expression in 27 different cancer types and tumor or normal tissues. Previous research has shown that activated *GFPT2* binds to many signaling proteins, stimulating the activation of several signaling pathways and contributing to human cancers. Analysis of 27 cancer datasets from the GTEx and TCGA was consistent with those in previous studies that demonstrated that *GFPT2* was significantly overexpressed in five types of cancer compared to that in normal tissues, while *GFPT2* expression was downregulated in 15 types of cancer (Fig. 1). Therefore, our research provides insights into the application of *GFPT2* as a pan-cancer prognostic marker in the context of oncology, thereby potentiating the development of targeted therapy research for *GFPT2*.

Our current study also identified the relationship between *GFPT2* expression level and pan-cancer prognosis in the GTEx and TCGA databases (Fig. 3). The high expression level of *GFPT2* is significantly correlated with an improved overall survival (OS) in MESO and UVM (Fig. 2 and 3).

GFPT2 expression is related to reduced treatment response and poor outcomes in non-small-cell lung cancer [13]. Likewise, increased *GFPT2* expression is related to poor

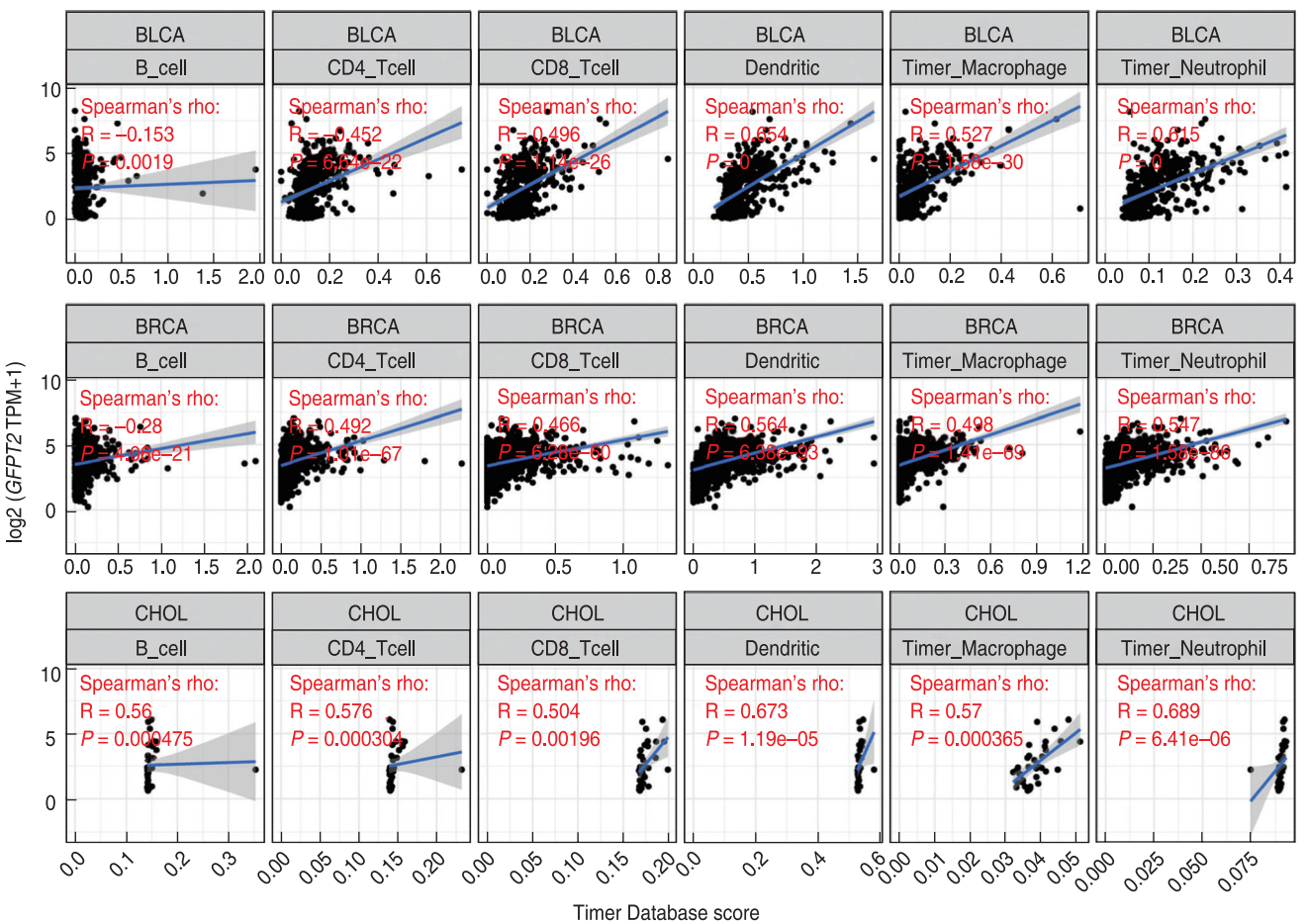


Fig. 4 Correlation between *GFPT2* expression and immune cells in BLCA, BRCA, and CHOL

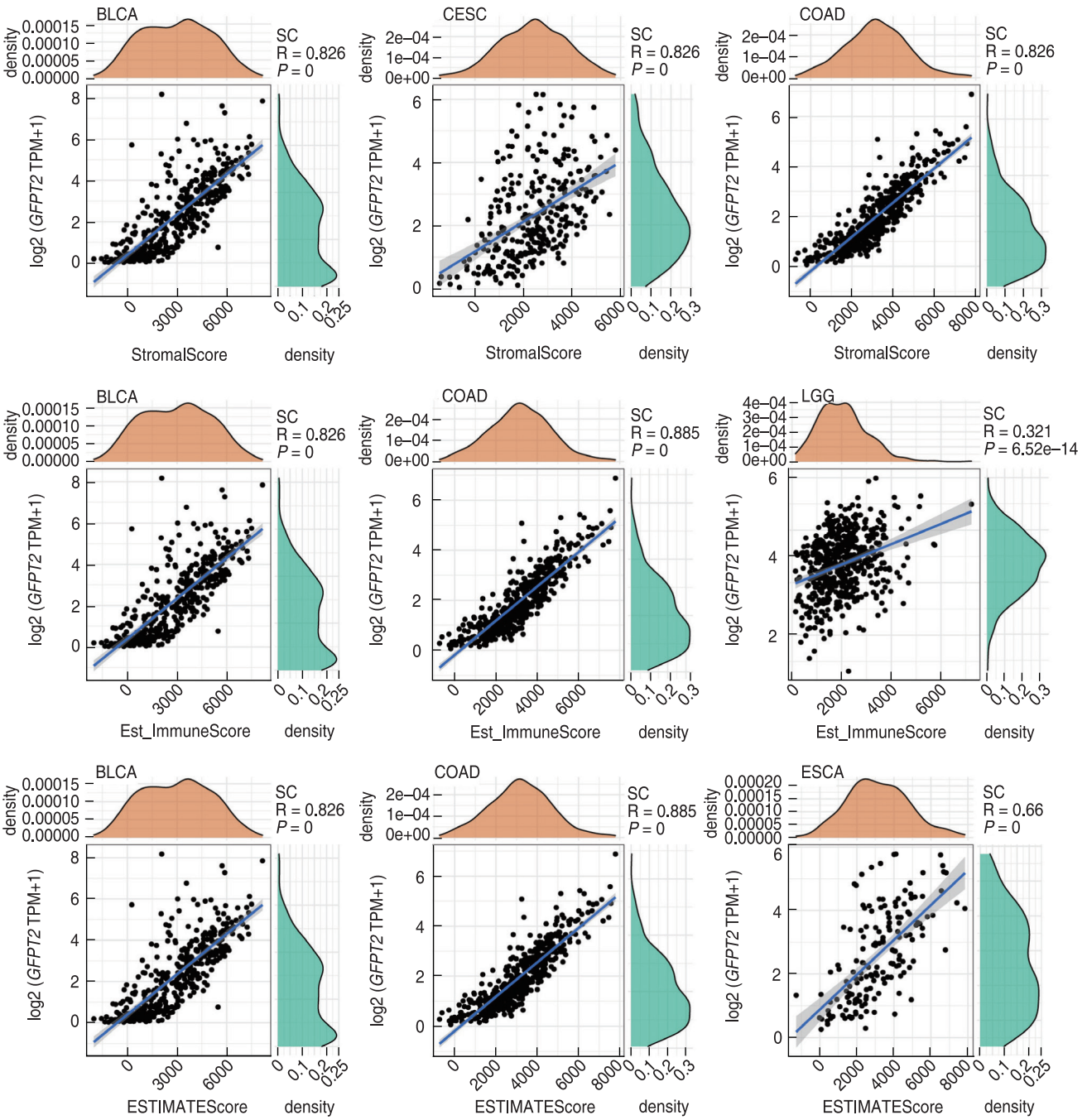


Fig. 5 TME of *GFPT2* expression in BLCA, COAD, and ESCA

outcomes, including decreased OS, locoregional relapse, and treatment failure in UVM [14]. These data contradicted our current results. *GFPT2* had a detrimental effect in MESO and UVM, which is consistent with the results of previous studies that the OS of UVM patients with high *GFPT2* expression is significantly lower than that of patients with low *GFPT2* expression [15]. In summary, these findings strongly suggest that *GFPT2* can be used as a prognostic marker for pan-cancer.

The presence of *GFPT2* in lung tumors has been shown to predict adequate diagnosis [16]. Recent studies have shown that lung function decline is related to the downregulation of *GFPT2*-regulated immune microenvironments and lung microenvironments present favorable anti-tumor immune response features [17]. High *GFPT2* expression can reduce the inflammatory response of macrophages [18]. These studies confirmed that *GFPT2* might improve or worsen the disease by regulating

immune-related cells and microenvironments. In this study, the Spearman's correlation coefficient increased as *GFPT2* expression increased (Fig. 4). High *GFPT2* expression significantly enhanced the body's immune ability, providing a precise target for the treatment of patients with BLCA, BRCA, and CHOL.

Another essential finding in this study was that *GFPT2* expression was related to TMEs in pan-cancer (Fig. 5). TMEs act in tumorigenesis and progression [19–21]. The ESTIMATE algorithm is based on single sample Gene Set Enrichment Analysis and generates three scores: stromal score, which captures the presence of stroma in tumor tissue; immune score, which represents immune cell infiltration in tumor tissue; and estimate score, which infers tumor purity [22]. Exploring potential therapeutic targets can help reshape the TME and promote it from tumor-friendly metastasis to tumor-suppressive metastasis. Many studies have revealed the importance of the immune microenvironment in tumorigenesis [23–27]. The results of our transcriptome analysis on the pan-cancer data from TCGA database show that the immune components in the TME contribute to patient prognosis. In particular, the ratio of stromal and immune components in the TME is significantly related to BLCA, CESC, and COAD (Fig. 5). These results emphasize the importance of exploring the interaction between tumor cells and immune cells to provide new insights for developing more effective treatment options. It is also crucial to distinguish the inherent stemness of cancer stem cells from the dedifferentiation caused by the TME. However, to solve this problem, other genome data sets and/or laboratory experiments need to be used for further verification, which is beyond the scope of this paper.

In conclusion, *GFPT2* was screened as a key immune-related gene in BLCA, BRCA, CHOL, CESC, and COAD. The present study data suggest that *GFPT2* might predict unfavorable cancer outcomes. The effect of tumor purity and immune cell infiltration on prognosis should also be considered in cancer research.

Conflict of interest

The authors indicate no potential conflicts of interest.

References

1. Bray F, Ferlay J, Soerjomataram I, *et al*. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA, 2018, 68: 394–424.
2. Arnold M, Abnet CC, Neale R E, *et al*. Global burden of 5 major types of gastrointestinal cancer. Gastroenterology, 2020, 159: 335–349.
3. Feng RM, Zong YN, Cao SM, *et al*. Current cancer situation in China: good or bad news from the 2018 Global Cancer Statistics? Cancer commun, 2019, 39: 22–30.
4. Lee JH, An CH, Yoo NJ, *et al*. Mutational intratumoral heterogeneity of a putative tumor suppressor gene RARRES3 in colorectal cancers. Pathol Res Pract, 2018, 10: 1–9.
5. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. Nature, 2018, 553: 446–454.
6. Barry K C, Hsu J, Broz M L, *et al*. A natural killer–dendritic cell axis defines checkpoint therapy–responsive tumor microenvironments. Nat Med, 2018, 24: 1178–1191.
7. Viúdez A, Carvalho F, Maleki Z, *et al*. Correction: A new immunohistochemistry prognostic score (IPS) for recurrence and survival in resected pancreatic neuroendocrine tumors (PanNET). Oncotarget, 2017, 8: 18617–18625.
8. Ireland L, Santos A, Ahmed MS, *et al*. Chemoresistance in pancreatic cancer is driven by stroma-derived insulin-like growth factors. Cancer Res, 2016, 76: 1–15.
9. Katheder NS, Khezri R, O'Farrell F, *et al*. Microenvironmental autophagy promotes tumour growth. Nat, 2017, 541: 417–420.
10. Chovancová Z. Secondary immunodeficiency as a consequence of chronic diseases. Vnitr Lek, 2019, 65: 117–124.
11. Schlesinger M. Role of platelets and platelet receptors in cancer metastasis. J Hematol Oncol, 2018, 11: 125–132.
12. Wang W, Yang X, Dai J, *et al*. Prostate cancer promotes a vicious cycle of bone metastasis progression through inducing osteocytes to secrete GDF15 that stimulates prostate cancer growth and invasion. Oncogene, 2019, 38: 4540–4559.
13. Szymura SJ, Zaemes JP, Allison DF, *et al*. NF-κB upregulates glutamine-fructose-6-phosphate transaminase 2 to promote migration in non-small cell lung cancer. Cell Commun Signal , 2019, 17: 24–30.
14. Carvajal RD, Schwartz GK, Tezel T, *et al*. Metastatic disease from uveal melanoma: treatment options and future prospects. Br J Ophthalmol, 2017, 101: 38–44.
15. Chattopadhyay C, Kim DW, Gombos DS, *et al*. Uveal melanoma: From diagnosis to treatment and the science in between. Cancer, 2016, 122: 2299–2312.
16. Zhang W, Bouchard G, Yu A, *et al*. GFPT2-expressing cancer-associated fibroblasts mediate metabolic reprogramming in human lung adenocarcinoma. Cancer Res, 2018, 78: 3445–3457.
17. Xu X, Qiao D, Mann M, *et al*. Respiratory syncytial virus infection induces chromatin remodeling to activate growth factor and extracellular matrix secretion pathways. Viruses, 2020, 12: 5–13.
18. Lu L, Wei R, Bhakta S, *et al*. Weighted gene co-expression network analysis identifies key modules and hub genes associated with mycobacterial infection of human macrophages. Antibiotics, 2021, 10: 145–153.
19. Ang KK, Berkey BA, Tu X, *et al*. Impact of epidermal growth factor receptor expression on survival and pattern of relapse in patients with advanced head and neck carcinoma. Cancer Res, 2002, 62: 7350–7356.
20. Ganly I, Talbot S, Carlson D, *et al*. Identification of angiogenesis/metastases genes predicting chemoradiotherapy response in patients with laryngopharyngeal carcinoma. J Clin Oncol, 2007, 25: 1369–1376.
21. Takikita M, Xie R, Chung JY, *et al*. Membranous expression of Her3 is associated with a decreased survival in head and neck squamous cell carcinoma. J Transl Med, 2011, 9: 126–136.
22. Yoshihara K, Shahmoradgoli M, Martínez E, *et al*. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun, 2013, 4: 2612–2618.
23. Yan H, Qu J, Cao W, *et al*. Identification of prognostic genes in the acute myeloid leukemia immune microenvironment based on TCGA data analysis. Cancer Immunol Immun, 2019, 68: 1971–1978.
24. Binnewies M, Roberts EW, Kersten K, *et al*. Understanding the tumor

immune microenvironment (TIME) for effective therapy. Nat Med, 2018, 24: 541–550.
25. Brouwer-Visser J, Cheng WY, Bauer-Mehren A, *et al*. Regulatory T-cell genes drive altered immune microenvironment in adult solid cancers and allow for immune contextual patient subtypin. Cancer Epidem Biomar, 2018, 27: 103–112.
26. Taube J M, Galon J, Sholl L M, *et al*. Implications of the tumor immune microenvironment for staging and therapeutics. Modern pathol, 2018, 31: 214–234.
27. Aixia C, Zhao S N, Zhou F, *et al*. Identification of potential immune-related prognostic biomarkers of lung cancer using gene co-expression network analysis. Oncol Transl Med, 2020, 6: 247–257.

DOI 10.1007/s10330-021-0500-0
Cite this article as: Zhang JC, Wang T, Wei SA, *et al*. *GFPT2* pan-cancer analysis and its prognostic and tumor microenvironment associations. Oncol Transl Med, 2021, 7: 286–293.

Oncology and Translational Medicine
DOI 10.1007/s10330-021-0515-5

December 2021, Vol. 7, No. 6, P294–P304

ORIGINAL ARTICLE

Effect of *UBR5* on the tumor microenvironment and its related mechanisms in cancer*

Guangyu Wang[1], Sutong Yin[2], Justice Afrifa[3], Guihong Rong[4], Shaofeng Jiang[5], Haonan Guo[4] (✉), Xianliang Hou[1] (✉)

[1] Central Laboratory, Guangxi Health Commission Key Laboratory of Glucose and Lipid Metabolism Disorders, The Second Affiliated Hospital of Guilin Medical University, Guilin 541199, China
[2] Reproductive Medical Center, No. 924 Hospital of Chinese People's Liberation Army, Guilin 541001, China
[3] Department of Medical Laboratory Science, School of Allied Health Sciences, University of Cape Coast, Cape Coast, Ghana
[4] Department of Clinical Laboratory, The Affiliated Hospital of Guilin Medical University, Guilin 541001, China
[5] Guangxi Key Laboratory of Tumor Immunology and Microenvironmental Regulation, Guilin Medical University, Guilin 541199, China

Abstract

Objective *UBR5*, recently identified as a potential target for cancer therapeutics, is overexpressed in multiple malignant tumors. In addition, it is closely associated with the growth, prognosis, metastasis, and treatment response of multiple types of cancer. Although emerging evidence supports the relationship between *UBR5* and cancer, there are limited cancer analyses available.

Methods In this study, online databases (TIMER2, GEPIA2, UALCAN, c-BioPortal, STRING) were employed to comprehensively explore expression levels and prognostic values of the *UBR5* gene in cancer, using bioinformatic methods.

Results We found that various characteristics of the *UBR5* gene such as gene expression, survival value, genetic mutation, protein phosphorylation, immune infiltration, and pathway activities in the normal tissue were remarkably different from those in the primary tumor. Furthermore, "protein processing in spliceosome" and "ubiquitin mediated proteolysis" have provided evidence for their potential involvement in the development of cancer.

Conclusion Our findings may provide insights for the selection of novel immunotherapeutic targets and prognostic biomarkers for cancer.

Received: 31 August 2021
Revised: 20 September 2021
Accepted: 25 October 2021

Key words: *UBR5*; cancer; tumor; prognosis; biomarker

Despite its declining incidence in many developed countries, cancer remains the most common cause of death across the globe. More than nineteen million people were diagnosed and over nine million people died as a result of cancer in 2020 alone. Cancer is characterized by a high degree of malignancy, rapid development and poor prognosis [1–3]. In the wake of the rapid strides being made by scientists and clinicians to explore novel prognostics, diagnostics and therapeutic options, cancer still remains one of the most elusive diseases in terms of treatment and management.

The human *UBR5* gene, which is widely expressed in various cell types, has 59 exons encoding approximately 10 kb of mRNA and > 300 kDa of protein [4]. It is highly conserved in metazoans, has unique structural features, and has been implicated in the regulation of the DNA damage response, metabolism, transcription, and apoptosis [5–7]. *UBR5* is a key regulator of cell signaling

✉ Correspondence to: Xianliang Hou. E-mail: houxl115@126.com; Haonan Guo. E-mail: guohaonan@glmc.edu.cn
*Supported by grants from the Guangxi Natural Science Foundation (No. 2019GXNSFBA245032), the Guangxi Science and Technology Plan Project (No. Gui Ke AD20238021), the open funds of the Guangxi Key Laboratory of Tumor Immunology and Microenvironmental Regulation (No. 203030302008, 203030302018 and 2020KF010), Guilin Science Research and Technology Development Project (No. 20190218-5-5), and the Basic Ability Enhancement Program for Young and Middle-aged Teachers of Guangxi (No. 2021KY0496).

related to the field of tumor biology. Recent studies have primarily demonstrated that *UBR5* plays an important role in the development of many tumors, and its expression may be closely associated with growth and proliferation of malignant tumors [8–9]. For example, in breast cancer, *UBR5* is coamplified with Myelocytomatosis (MYC) to limit MYC-dependent apoptosis by encoding a ubiquitin ligase [10]. Also in breast cancer, others have shown that triple-negative breast cancer (TNBC) metastasis and cisplatin resistance may be mediated by elevated *UBR5* expression [11–13]. Similarly, in colorectal cancer, Xie *et al*. concluded that an elevated *UBR5* levels play an oncogenic role and may be a potential prognostic marker [14]. The trends elucidated in the various studies point to the likelihood of *UBR5* as an oncogenic mediator in most cancers. However, contrary to other cancers, inactivating mutations have been observed in the *UBR5* gene, as is the case in approximately 18% of mantle cell lymphoma cases [15]. It is therefore clear that *UBR5* is a key cell signaling regulator that has been strongly associated with cancer; however, its function as a promoter or inhibitor of tumorigenesis still remains inconclusive.

In this study, we conducted an in-depth and comprehensive bioinformatic analysis of the expression of the *UBR5* gene and evaluated its potential as a therapeutic target and prognostic biomarker. Findings from this study will provide a better understanding of this gene, and help clinicians select appropriate therapeutic drugs and more accurately prognose long-term outcomes in cancer patients.

Materials and methods

TIMER2

TIMER2 (http://timer.cistrome.org/) is a reliable tool that provides the expression status of *UBR5* across The Cancer Genome Atlas (TCGA) datasets from different tumor tissues and adjacent normal tissues. It also provides a robust estimation of immune infiltration levels for TCGA or user-provided tumor profiles using six state-of-the-art algorithms. In this study, the expression status of *UBR5* across the TCGA dataset, and the correlation between the infiltration of immune cells and *UBR5* expression was evaluated [16–18].

GEPIA2

GEPIA2 (http://gepia2.cancer-pku.cn/#index) is a tool for analyzing the RNA sequencing expression data of 9,736 tumors and 8,587 normal samples from TCGA [19]. GEPIA2 was employed in this study to perform a differential *UBR5* expression analysis of tumor and adjacent normal tissue, expression of *UBR5* total protein, pathological stage analysis, and correlative prognostic analysis of the *UBR5* gene.

UALCAN

UALCAN (http://ualcan.path.uab.edu/index.html) is an interactive web resource that provides analysis based on TCGA and MET500 cohort data [20]. It allows analysis of relative expression of query genes across tumor and normal samples, as well as in various tumor sub-groups based on individual cancer stages, tumor grade or other clinicopathological features. Protein expression analysis was conducted using Clinical Proteomic Tumor Analysis Consortium (CPTAC) and the available datasets of six tumors were selected in our study.

cBioPortal

The cBioPortal (http://cbioportal.org) is a web resource for exploring, visualizing, and analyzing multidimensional cancer genomics data. This web resource provides the option of querying a single cancer study or querying across multiple cancer studies. It is also possible to view relevant genomic alterations in cancer samples and analyze multidimensional cancer genomics data [21]. The alteration frequency, type of alterations of *UBR5* and copy number alterations are shown in our study. In addition, we aimed to assess the genetic alterations of *UBR5* and its correlation with survival values in cancer patients, using data from TCGA.

STRING

STRING (https://string-db.org/) is a web resource that integrates all known and predicted associations between proteins [22]. We conducted a protein-protein interaction network analysis of differentially expressed levels of the *UBR5* gene, to explore the interactions among them with STRING.

Results

Aberrant expression of *UBR5* in patients with cancer

To understand the oncogenic role of human *UBR5*, we examined its expression status across the TCGA dataset from different cancer types using the TIMER2 approach. The expression level of *UBR5* in the tumor tissues of breast invasive carcinoma (BRCA), cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), pheochromocytoma and paraganglioma (PCPG), prostate adenocarcinoma (PRAD), rectum adenocarcinoma (READ), and stomach adenocarcinoma (STAD) was higher than in normal tissues. On the contrary, the expression level of kidney chromophobe (KICH), kidney renal clear cell carcinoma (KIRC), thyroid carcinoma (THCA), and uterine corpus

endometrial carcinoma (UCEC) was lower than in normal tissues (Fig. 1a).

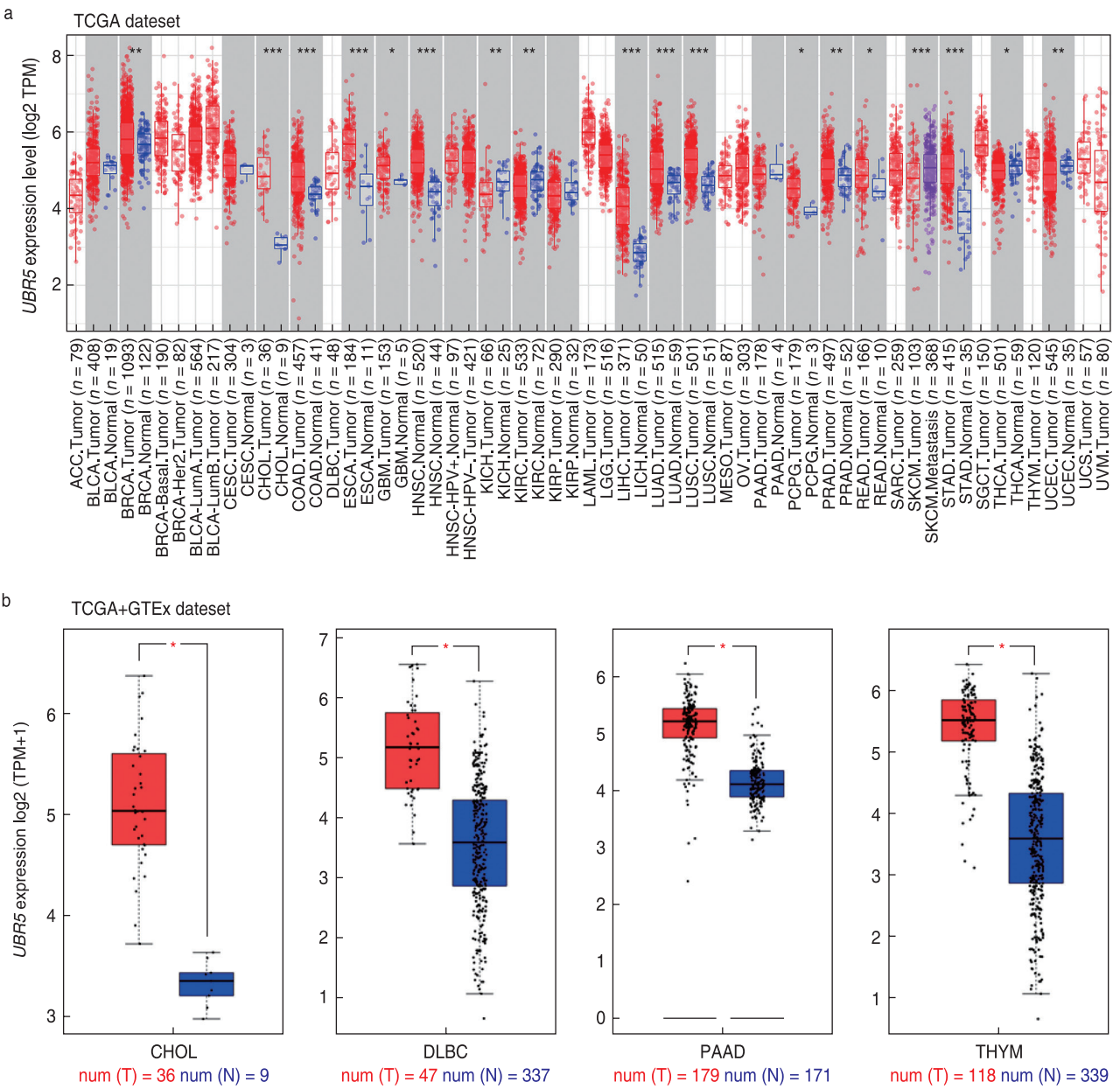
Based on clinical data extracted from the GTEx dataset, we compared the differential expression level of *UBR5* in tumor tissues with that in matched normal tissues of CHOL, lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), pancreatic adenocarcinoma (PAAD), and thymoma (THYM). The results showed a significantly elevated *UBR5* expression among the tumor tissues ($P < 0.05$) compared to that in normal tissues (Fig. 1b).

In the CPTAC dataset, we observed significantly higher expression of *UBR5* total protein in the primary tissues of breast cancer ($P < 0.001$), clear cell RCC ($P < 0.001$), colon cancer ($P < 0.001$), LUAD ($P < 0.001$), UCEC ($P < 0.001$) and ovarian cancer ($P < 0.05$), than in normal tissues (Fig. 1c).

To assess the association between *UBR5* expression and the pathological stages of cancer, the "Pathological Stage Plot" module of GEPIA2 was employed to analyze pathological data from COAD ($P < 0.001$), ESCA ($P < 0.05$) and KICH ($P < 0.001$) patients in the TCGA database (Fig. 1d).

Prognostic value of *UBR5* in patients with cancer

Patients were grouped into high-expression and low-expression groups. We examined the association between *UBR5* expression and the prognosis of patients with



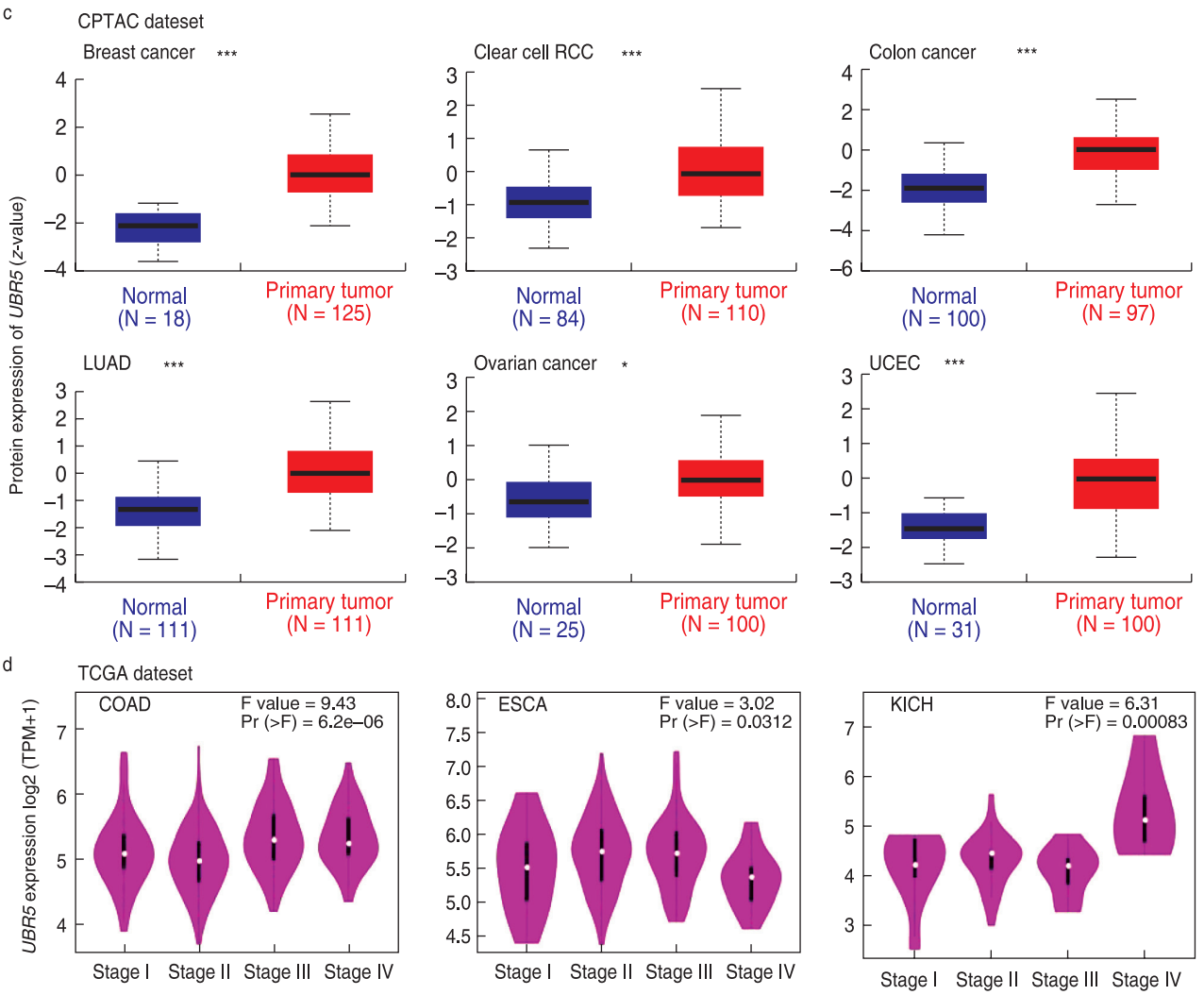


Fig. 1 Expression levels of the *UBR5* gene in different cancer samples. (a) *UBR5* expression status varies in different cancers through TIMER2, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; (b) Differential expression of *UBR5* between the normal tissues and the tumor tissues through GEPIA2, $^{*}P < 0.05$; (c) Higher expression of *UBR5* total protein in the primary tissues through UALCAN, all $P < 0.05$; (d) Expression levels of *UBR5* in different pathological stages through GEPIA2, $P < 0.05$.

cancer using TCGA and GEO datasets. *UBR5* expression was linked to cancer prognosis: the Fig. 2 plot showing overall survival (OS) for BRCA ($P = 0.041$) within the TCGA project indicates that higher *UBR5* expression is linked to a poor prognosis. Disease-free survival (DFS) for PRAD ($P = 0.013$) and READ ($P = 0.037$) values also supported this conclusion.

Genomic alterations of *UBR5* in cancer

The cBioPortal was used to determine the genetic alteration status of *UBR5* in cancer, based on TCGA datasets. As shown in Fig. 3a, the highest alteration frequency was found in the bladder urothelial carcinoma tumor with "Amplification" as the primary type, whereas PCPG exhibited the lowest alteration among all of the cancer samples queried. It is noted that all uveal melanoma cases with genetic alterations showed copy number amplifications of *UBR5*. As shown in Fig. 3b, 519 mutations were identified in patients. Out of the alterations, 377 missense mutations, 126 truncating mutations, 4 in-frame mutations and 12 fusion mutations were detected. Missense mutations of *UBR5* were identified as the main type of genetic alteration, and E2121Kfs*28/E2121Rfs*13/K2120Rfs*13 alteration was predicted to induce a frame shift mutation of the *UBR5* gene. The 3D structure of the *UBR5* protein can be observed in Fig. 3c. In addition, genetic alterations have been found in patients with different types of cancer, which is related to survival prognoses. As shown in Fig. 3d, STAD cases with altered *UBR5* showed better

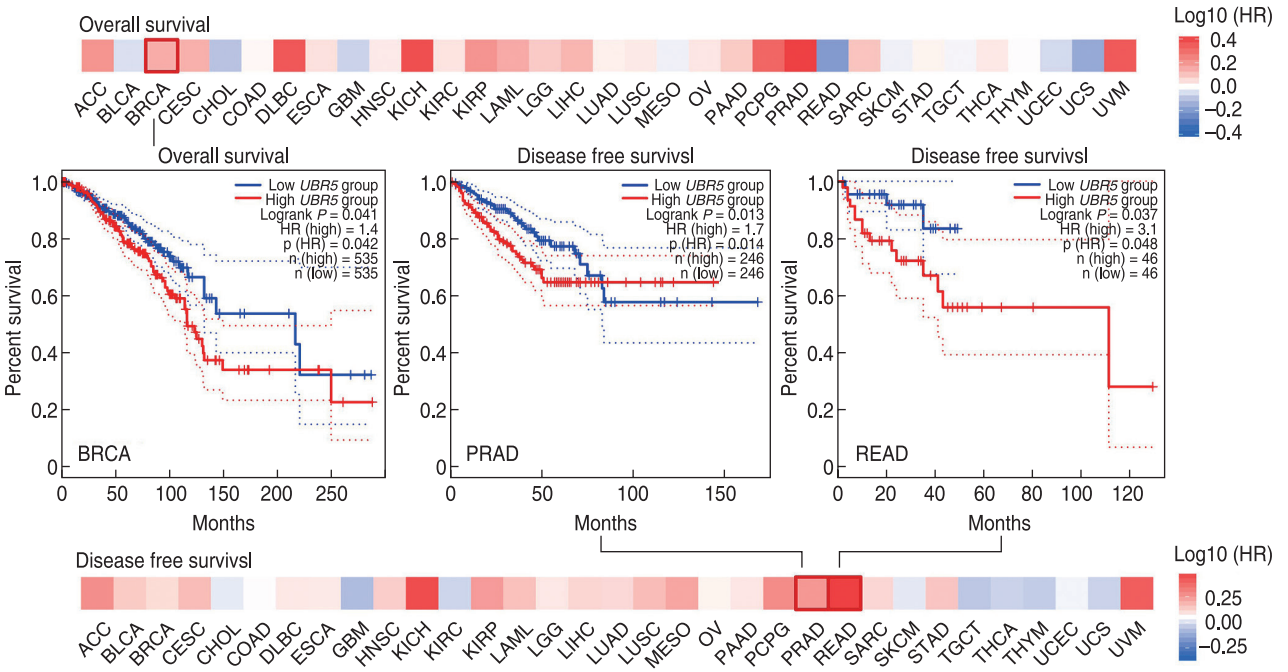
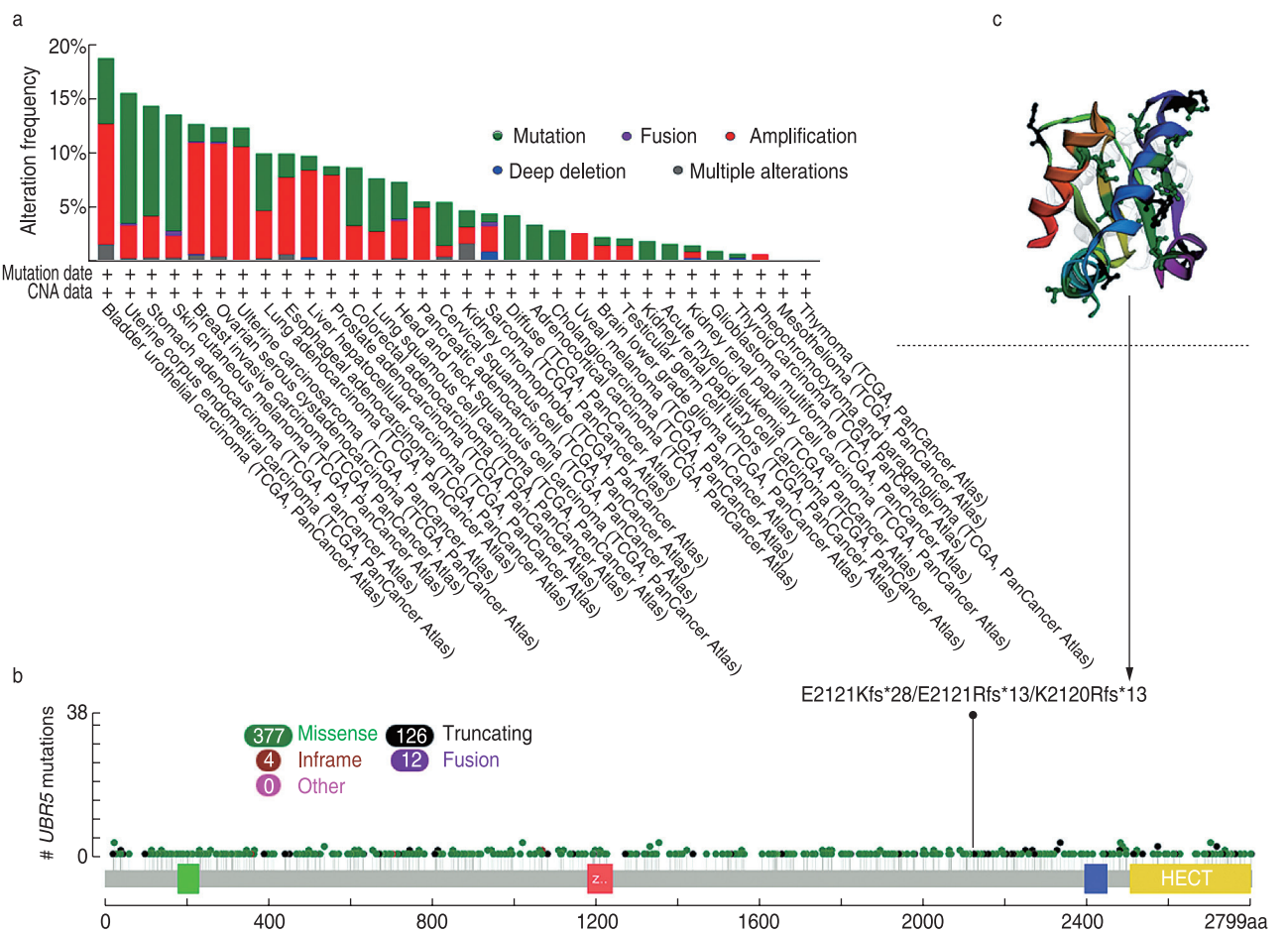


Fig. 2 Relationship between *UBR5* gene expression level and survival in cancer patients using GEPIA2. Clinical survival curves of BRCA (OS), PRAD (DFS), and READ (DFS) are presented, P < 0.05.



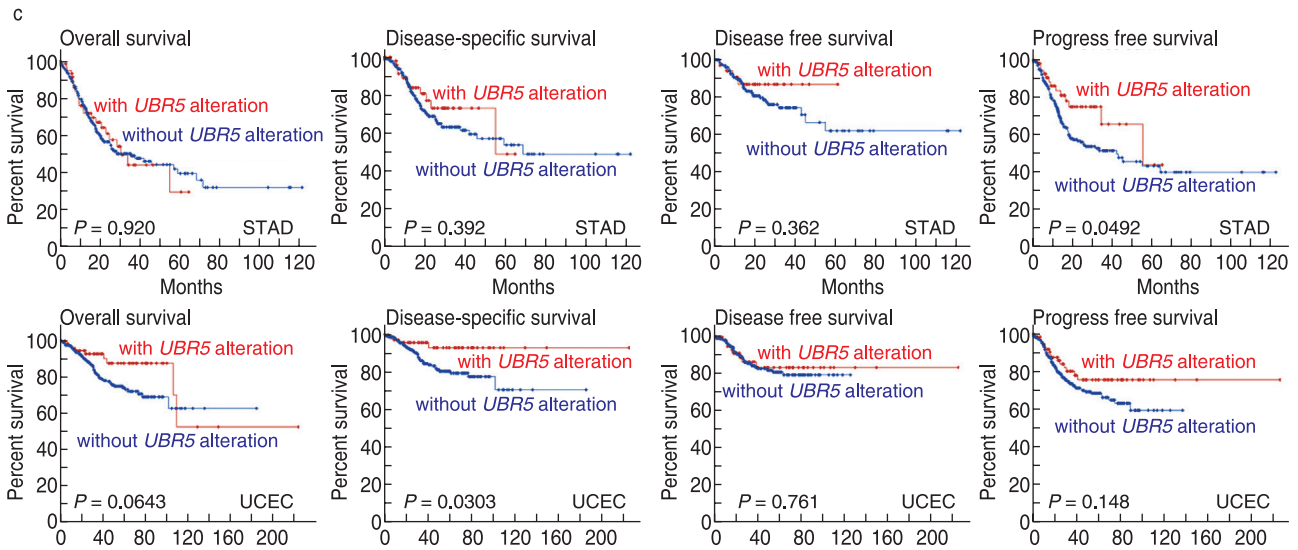


Fig. 3 Genetic features of mutations of *UBR5* in different tumors (cBioPortal). (a) Alteration frequency in different tumor samples; (b) Sites and case number of *UBR5* genetic alterations; (c) 3D structure of the *UBR5* protein; (d) Clinical survival curve of STAD and UCEC.

prognosis in progression-free (P = 0.0492), but not overall (P = 0.920), disease-specific (P = 0.392) and disease-free (P = 0.362) survival, compared with cases without *UBR5* alteration. UCEC cases with altered *UBR5* showed better prognosis in disease-specific (P = 0.0492), but not overall (P = 0.0643), disease-free (P = 0.761) and progression-free survival (P = 0.148), compared with cases without *UBR5* alteration.

Protein phosphorylation of *UBR5* in patients with cancer

We also investigated *UBR5* phosphorylation levels using the CPTAC dataset. Clear cell clear cell renal cell carcinoma (RCC), ovarian cancer, LUAD, UCEC and breast cancer were analyzed. The analysis of *UBR5* phosphoprotein expression level is presented in Fig. 4a. The clinical data showed a higher phosphorylation level of the S327 locus in all primary tumor tissues compared with that seen in normal tissues (Fig. 4b–f, all $P < 0.05$), followed by a lower phosphorylation level of the S636 locus for colon cancer (Fig. 4f, P = 1.2e-06), LUAD (Fig. 4d, P = 1.8e-05), colon cancer (Fig. 4f, P = 1.3e-14) and the S1549 locus for ovarian cancer (Fig. 4c, P = 5.2e-03),

Immune cell infiltration of *UBR5* in patients with cancer

Next, we used the TIMER2, TIDE, XCELL, MCPCOUNTER and EPIC algorithms to assess the correlations of *UBR5* expression with immune infiltration levels. Heat map of different expressed *UBR5* gene are further presented in Fig. 5a. We found a significant negative correlation between *UBR5* expression and the estimated infiltration value of cancer-associated fibroblasts for Testicular Germ Cell Tumor (TGCT). (Fig. 5b, *cor* = –0.242, P = 3,11e-03)

Enrichment analysis of *UBR5*-related partners

In an attempt to investigate the potential enrichment of particular molecular mechanisms in tumorigenesis, we attempted to screen out targeting *UBR5*-binding proteins and *UBR5* expression-related genes using STRING and GEPIA2. Fig. 6a showed the interaction network of 50 *UBR5*-binding proteins supported by experimental evidence. There were significant positive correlations between the expression level of *UBR5* and that of cell division cycle and apoptosis regulator 1 (*CCAR1*) (R = 0.56), Arginine/serine-rich coiled-coil 2 (*RSRC2*) (R = 0.52), Suppressor of mek1 (*SMEK1*) (R = 0.52), Ubiquitin specific peptidase 7 (*USP7*) (R = 0.52) and Zinc finger protein 7 (*ZNF7*) (R = 0.68) genes (all $P < 0.001$; Fig. 6b). As shown in Fig. 6c, the heatmap also revealed that the above-mentioned genes were positively correlated with *UBR5* in the majority of types of tumor. An intersection analysis of 50 *UBR5*-binding proteins and 100 *UBR5* expression-related genes showed one common member, namely, SRSF1 (Fig. 6d). In addition, the KEGG data suggested that “spliceosome” and “ubiquitin mediated proteolysis” pathways were involved in cancer progression. (Fig. 6e).

Discussion

It is understood that *UBR5* is a tumor-related gene that affects the biological behavior of tumors in many aspects,

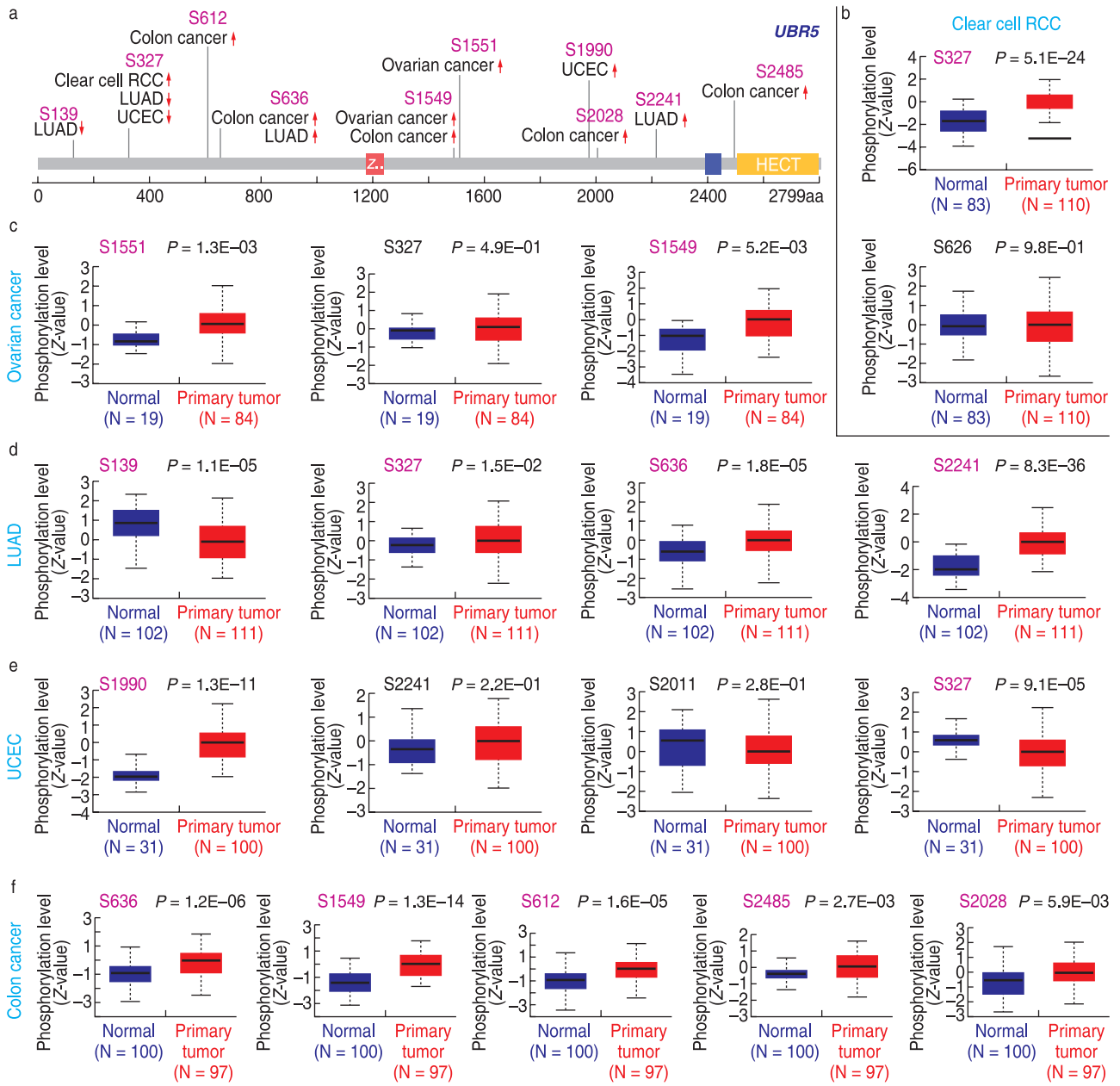


Fig. 4 Phosphorylation analysis of different tumors. (a) Analysis of *UBR5* phosphoprotein expression level based on the CPTAC dataset, S139, S327, S612, S626, S1549, S1551, S1990, S2028, S2241, and S2485; The box plots for different cancers, including clear cell RCC (b), ovarian cancer (c), LUAD (d), UCEC (e) and breast cancer (f), all $P < 0.05$.

such as cell cycle regulation, apoptosis regulation, tumor suppressor gene regulation, invasion and metastasis regulation [23]. Some studies have reported a correlation between the *UBR5* gene, tumor microenvironment, and cancer immunotherapy, suggesting that the gene may modulate tumor progression and provide an immunotherapeutic effect [24–25]. However, the prognostic value and the biological function of the *UBR5* gene in cancer has not been well-characterized. With further investigation into this gene, knowledge regarding its regulatory mechanism in cancer will become increasingly clear, which will aid in the molecular diagnosis and targeted therapy of cancer, as well as improving prognostic assessments in cancer patients. Thus, we present a comprehensive overview of the *UBR5* gene based on data from TCGA, CPTAC and GEO databases.

We first explored expression of the *UBR5* gene and its correlation with the pathological cancer stage. We found that 17 genes were differentially expressed in cancerous tissues compared with the corresponding control tissues

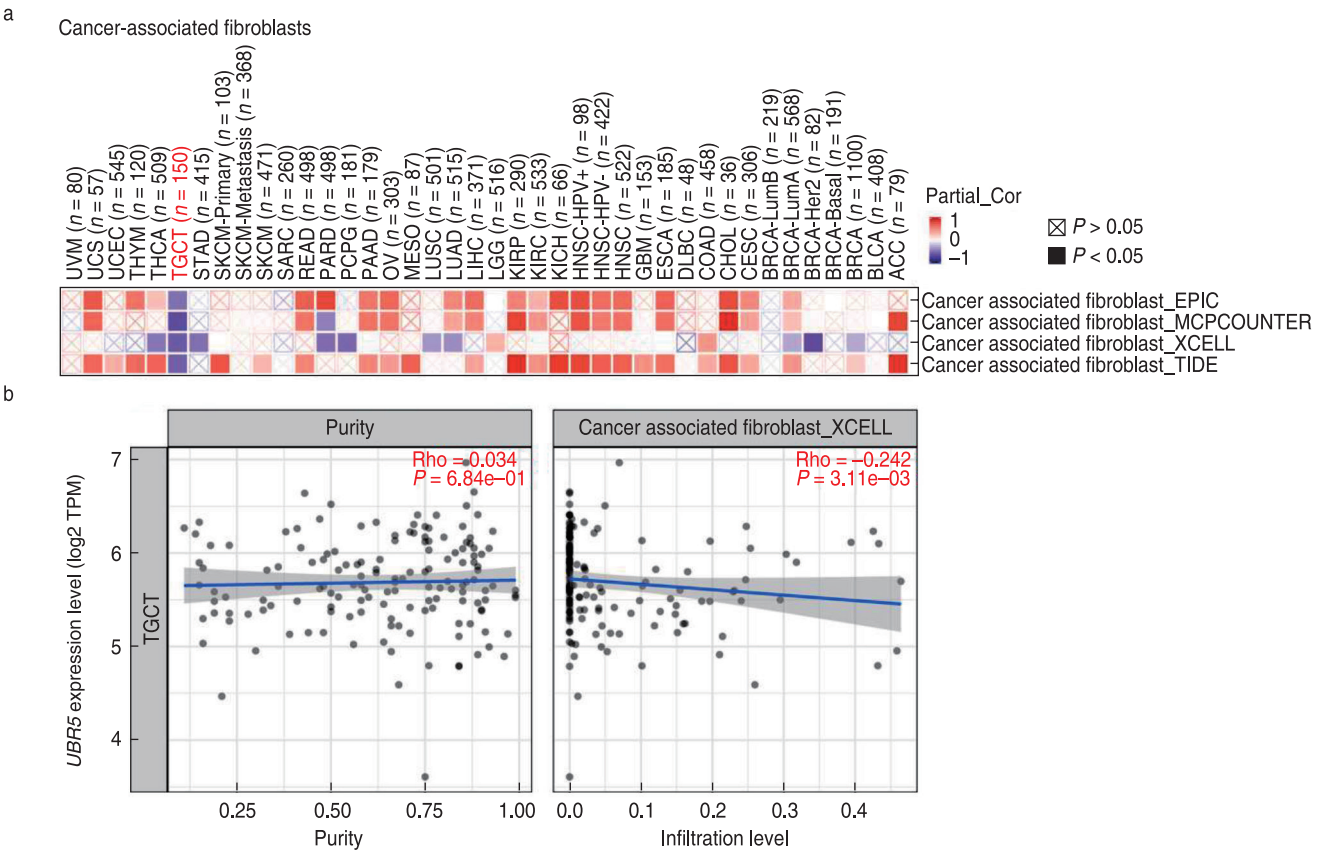
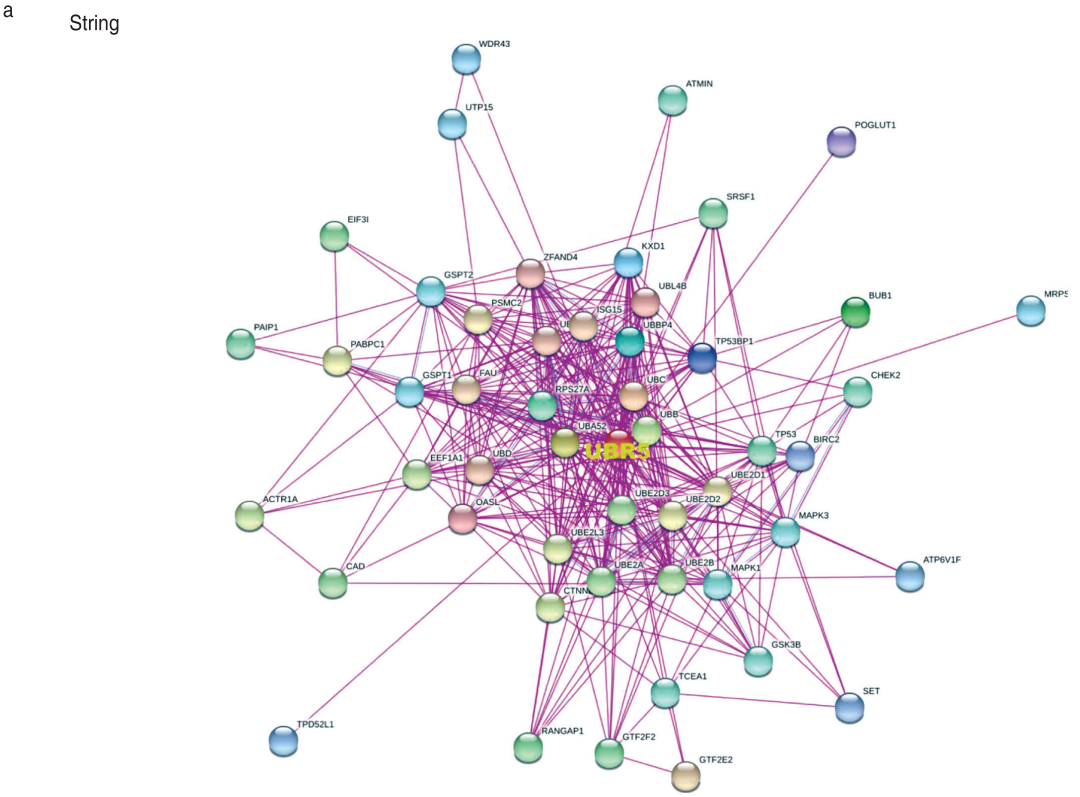


Fig. 5 Correlation analysis of *UBR5* in the tumor microenvironment and immune infiltration. (a) Correlation heat map of differentially expressed *UBR5* gene (TIMER2); (b) The correlation between differentially expressed *UBR5* gene and immune cell infiltration (TIMER2).



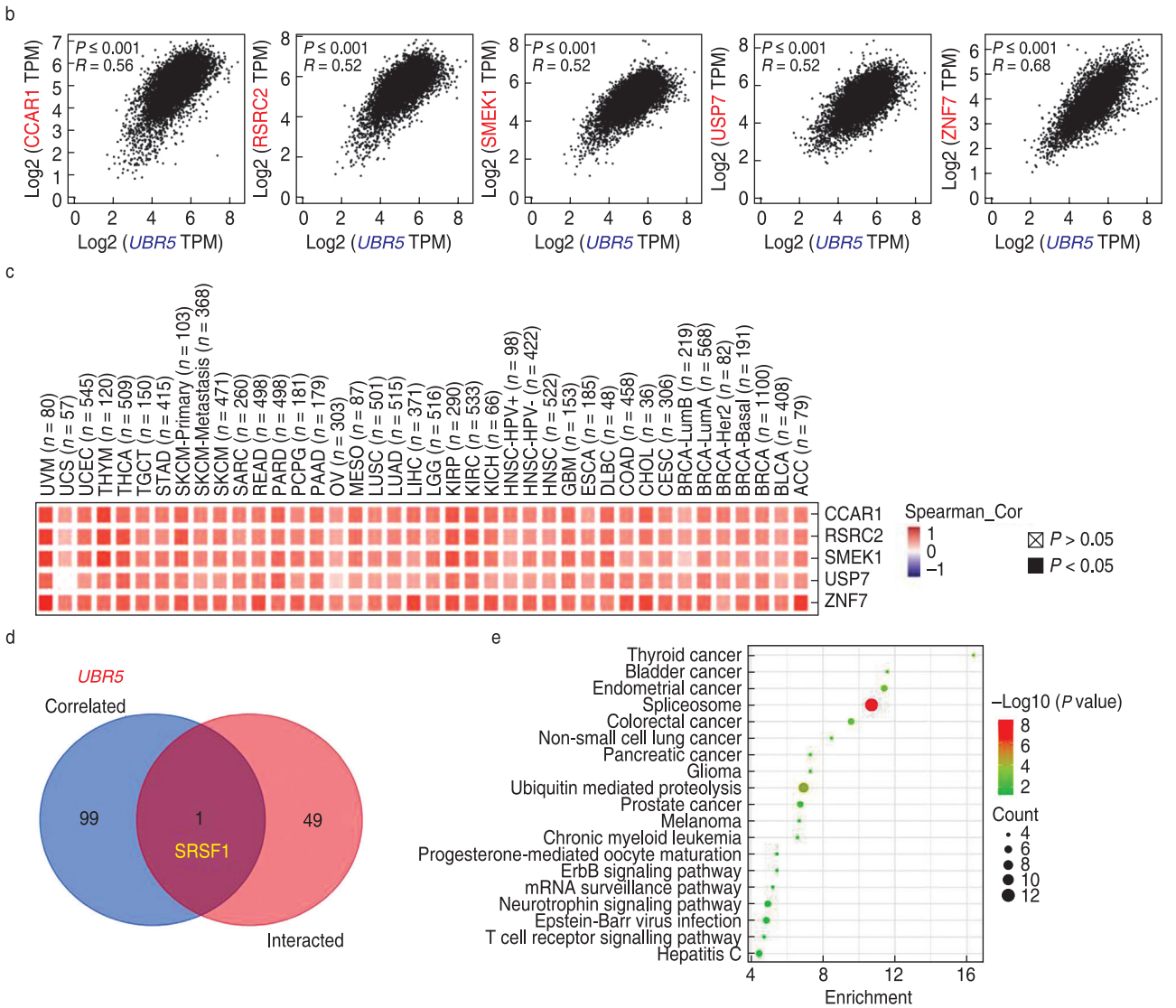


Fig. 6 *UBR5*-related gene enrichment analysis. (a) Interaction network of 50 *UBR5*-binding proteins through STRING tool; (b) *UBR5* expression level was positively correlated with that of CCAR1, RSRC2, SMEK1, USP7 and ZNF7 genes; (c) Correlation heat map of the differentially expressed *UBR5* gene; (d) One common member named SRSF1 was observed through intersection analysis; (e) *UBR5* expression-related genes for enrichment analysis.

(higher expression of BRCA, CHOL, COAD, ESCA, GBM, HNSC, LIHC, LUAD, LUSC, PCPG, PRAD, READ and STAD; lower expression of KICH, KIRC, THCA and UCEC). *UBR5* gene expression in cancerous tissues was further confirmed in studies from GTEx and CPTAC datasets. These data demonstrate that differentially expression of *UBR5* may play a significant role in cancer.

Furthermore, in BRCA, PRAD and READ patients, high expression of *UBR5* were significantly associated with a poor prognosis. Zhang *et al*. found that *UBR5* was overexpressed in gallbladder cancer tumor tissues and was significantly associated with tumor size, histological and tumor differentiation [26]. Yang *et al*. revealed that high expression of *UBR5* was associated with poor overall and disease-free survival in patients with gastric cancer [27]. This analysis demonstrates that *UBR5* may be an important biomarker for predicting the prognosis of patients with cancer.

Since the *UBR5* gene was significantly differentially expressed in cancer tissues, we explored its molecular characteristics. There were frequent genetic alterations in the *UBR5* gene expressed in cancer tissues, with mutation and amplification being the most common. It has been reported that the *UBR5* gene is localized to chromosome 8q22 [28]. Mutation and amplification occur frequently in this region in many types of cancer, including breast cancer, esophageal cancer and mammary ductal carcinoma [29–31]. Tumorigenesis and the progression

of cancer are complex and multi-faceted, and genetic alteration plays an important role in these processes. We found a low to high correlation of prognoses with the differential expression of the *UBR5* gene, suggesting that *UBR5* plays a synergistic role in tumorigenesis and the progression of cancer.

We then focused on protein phosphorylation of *UBR5* in patients with cancer using UALCAN. Phosphorylation is a formidable regulator of many proteins involved in essential intracellular processes. Studies have reported on the possible role of phosphorylation in both protein function and the progression of specific cancers [32]. Phosphorylation may provide key information about the derangements and serve as major targets for therapeutics, which is a rapidly growing area of cancer research [33]. The results showed that S327, S636 and S1549 all exhibited a higher phosphorylation level of *UBR5*. Bethard *et al*. revealed that *UBR5* has 477 potential phosphorylation sites. However, few studies have specifically targeted the identification of these phosphorylation sites [34]. Further laboratory studies to evaluate the potential role of *UBR5* phosphorylation in tumorigenesis are needed.

We also found a negative correlation between *UBR5* expression and immune infiltration of cancer-associated fibroblasts. Evidence indicates that cancer development is a complex process that involves interactions between tumor cells, stromal fibroblasts, and immune cells. Tumor-infiltrating immune cells play a role in the promotion or inhibition of tumor growth [35–37]. This analysis demonstrates the role of *UBR5* in the tumor microenvironment and the promotion or inhibition in different types of cancer.

Additionally, analysis of "protein processing in spliceosome" and "ubiquitin mediated proteolysis" pathways may bring novel insights into the potential association of *UBR5* with etiology or pathogenesis of cancers [38].

In conclusion, we hope these results will be a helpful guide to aid in helping diagnose cancer and to assist in the design of new immunotherapeutic drugs.

Conflicts of interest

The authors indicated no potential conflicts of interest.

Ethics approval and consent to participate

Ethics approval is not applicable because this study did not involve human or animal testing.

References

1. Cui X, Song P, Zhang L. New advances in the treatment for small cell lung cancer. Chin J Lung Cancer (Chinese), 2019, 22: 355–362.
2. Ge J, Yao B, Huang J, *et al*. Molecular genetic characterization reveals linear tumor evolution in a pulmonary sarcomatoid carcinomas patient with a novel PHF20-NTRK1 fusion: a case report. BMC Cancer, 2019, 19: 592.
3. Xu ZH, Huang W, Lin D, *et al*. Advances in Raman spectroscopy for nasopharyngeal carcinoma tissue. Spectrosc Spectr Anal, 2016, 36: 2518–2521.
4. Henderson MJ, Russell AJ, Hird S, *et al*. EDD, the human hyperplastic discs protein, has a role in progesterone receptor coactivation and potential involvement in DNA damage response. J Biol Chem, 2002, 277: 26468–26478.
5. Cipolla L, Bertoletti F, Maffia A, *et al*. UBR5 interacts with the replication fork and protects DNA replication from DNA polymerase η toxicity. Nucleic Acids Res, 2019, 47: 11268–11283.
6. Li CG, Mahon C, Sweeney NM, *et al*. PPARγ interaction with UBR5/ATMIN promotes DNA repair to maintain endothelial homeostasis. Cell Rep, 2019, 26: 1333–1343.
7. de Vivo A, Sanchez A, Yegres J, *et al*. The OTUD5–UBR5 complex regulates FACT-mediated transcription at damaged chromatin. Nucleic Acids Res, 2019, 47: 729–746.
8. Nakagawa T, Mondal K, Swanson PC. VprBP (DCAF1): a promiscuous substrate recognition subunit that incorporates into both RING-family CRL4 and HECT-family EDD/UBR5 E3 ubiquitin ligases. Bmc Mol Biol, 2013, 14: 22.
9. Swenson SA, Gilbreath TJ, Vahle H, *et al*. UBR5 HECT domain mutations identified in mantle cell lymphoma control maturation of B cells. Blood, 2020, 136: 299–312.
10. Qiao X, Liu Y, Prada ML, *et al*. UBR5 is co-amplified with MYC in breast tumors and encodes an ubiquitin ligase that limits MYC-dependent apoptosis. Cancer Res, 2020, 80: 1414–1427.
11. Bradley A, Zheng H, Ziebarth A, *et al*. EDD enhances cell survival and cisplatin resistance and is a therapeutic target for epithelial ovarian cancer. Carcinogenesis, 2014, 35: 1100–1109.
12. Liao L, Song M, Li X, *et al*. E3 Ubiquitin ligase UBR5 drives the growth and metastasis of triple-negative breast cancer. Cancer Res, 2017, 77: 2090–2101.
13. Matsuura K, Huang N, Cocce K, *et al*. Downregulation of the proapoptotic protein MOAP-1 by the UBR5 ubiquitin ligase and its role in ovarian cancer resistance to cisplatin. Oncogene, 2017, 36: 1698–1706.
14. Xie Z, Liang H, Wang J, *et al*. Significance of the E3 ubiquitin protein UBR5 as an oncogene and a prognostic biomarker in colorectal cancer. Oncotarget, 2017, 8: 108079–108092.
15. Meissner B, Kridel R, Lim RS, *et al*. The E3 ubiquitin ligase UBR5 is recurrently mutated in mantle cell lymphoma. Blood, 2013, 121: 3161–3164.
16. Tomczak K, Czerwińska P, Wiznerowicz M. The cancer genome atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol, 2015, 19: 68–77.
17. Blum A, Wang P, Zenklusen JC. SnapShot: TCGA-Analyzed tumors. Cell, 2018, 173: 530.
18. Li T, Fu J, Zeng Z, *et al*. TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res, 2020, 48: 509–514.
19. Tang Z, Li C, Kang B, *et al*. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res, 2017, 45: 98–102.
20. Chandrashekar DS, Bashel B, Balasubramanya SAH, *et al*. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia, 2017, 19: 649–658.
21. Gao J, Aksoy BA, Dogrusoz U, *et al*. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal, 2013, 6: 11.

22. Szklarczyk D, Gable AL, Lyon D, *et al*. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res, 2019, 47: 607–613.
23. Panfil AR, Al-Saleem J, Howard CM, *et al*. Stability of the HTLV-1 Antisense-Derived Protein, HBZ, is Regulated by the E3 Ubiquitin-Protein Ligase, UBR5. Front Microbiol, 2018, 9: 80.
24. Chen L, Yuan R, Wen C, *et al*. E3 ubiquitin ligase UBR5 promotes pancreatic cancer growth and aerobic glycolysis by downregulating FBP1 via destabilization of C/EBPα. Oncogene, 2021, 40: 262–276.
25. Song M, Yeku OO, Rafiq S, *et al*. Tumor derived UBR5 promotes ovarian cancer growth and metastasis through inducing immunosuppressive macrophages. Nat Commun, 2020, 11: 6298.
26. Zhang Z, Zheng X, Li J, *et al*. Overexpression of UBR5 promotes tumor growth in gallbladder cancer via PTEN/PI3K/Akt signal pathway. J Cell Biochem, 2019, 120: 11517–11524.
27. Yang M, Jiang N, Cao Q, *et al*. The E3 ligase UBR5 regulates gastric cancer cell growth by destabilizing the tumor suppressor GKN1. Biochem Bioph Res Co, 2016, 478: 1624–1629.
28. Ding F, Zhu X, Song X, *et al*. UBR5 oncogene as an indicator of poor prognosis in gastric cancer. Exp Ther Med, 2020, 20: 1.
29. Yang Y, Zhao J, Mao Y, *et al*. UBR5 over-expression contributes to poor prognosis and tamoxifen resistance of ERa+ breast cancer by stabilizing β-catenin. Breast Cancer Res Tr, 2020, 184: 699–710.
30. Wang Z, Kang L, Zhang H, *et al*. AKT drives SOX2 overexpression and cancer cell stemness in esophageal cancer by protecting SOX2 from UBR5-mediated degradation. Oncogene, 2019, 38: 5250–5264.
31. Fuja TJ, Lin F, Osann KE, *et al*. Somatic mutations and altered expression of the candidate tumor suppressors CSNK1 epsilon, DLG1, and EDD/hHYD in mammary ductal carcinoma. Cancer Res, 2004, 64: 942–951.
32. Singh V, Ram M, Kumar R, *et al*. Phosphorylation: implications in cancer. Protein J, 2017, 36: 1–6.
33. Ashton TM, McKenna WG, Kunz-Schughart LA, *et al*. Oxidative phosphorylation as an emerging target in cancer therapy. Clin Cancer Res, 2018, 24: 2482–2490.
34. Bethard JR, Zheng H, Roberts L, *et al*. Identification of phosphorylation sites on the E3 ubiquitin ligase UBR5/EDD. J Proteomics, 2011, 75: 603–609.
35. Wang S, Liu W, Ly D, *et al*. Tumor-infiltrating B cells: their role and application in anti-tumor immunity in lung cancer. Cell Mol Immunol, 2019, 16: 6–18.
36. Fridman WH, Galon J, Dieu-Nosjean M, *et al*. Immune infiltration in human cancer: prognostic significance and disease control. Curr Top Microbiol Immunol, 2011, 344: 1–24.
37. Steven A, Seliger B. The role of immune escape and immune cell infiltration in breast cancer. Breast Care, 2018, 13: 16–21.
38. Cui X, Zhang X, Liu M, *et al*. A pan-cancer analysis of the oncogenic role of staphylococcal nuclease domain-containing protein 1 (SND1) in human tumors. Genomics, 2020, 112: 3958–3967.

DOI 10.1007/s10330-021-0515-5
Cite this article as: Wang GY, Yin ST, Justice A, *et al*. Effect of *UBR5* on the tumor microenvironment and its related mechanisms in cancer. Oncol Transl Med, 2021, 7: 294–304.

Oncology and Translational Medicine
DOI 10.1007/s10330-021-0497-7

December 2021, Vol. 7, No. 6, P305–P313

ORIGINAL ARTICLE

Autophagy-related lncRNA and its related mechanism in colon adenocarcinoma

Feifei Tan, Zhongyin Zhou (✉)

Department of Gastroenterology, Renmin Hospital of Wuhan University, Hubei Key Laboratory of Digestive System, Wuhan 430060, China

Abstract

Objective Colon cancer is a type of cancer with high morbidity and mortality, of which adenocarcinoma is the most common type. Numerous studies have found that long noncoding RNAs (lncRNAs) are related to the occurrence and development of colon cancer. Autophagy is a key metabolic process in the human body and has a role in affecting cancer growth. In this study, our aim was to explore the correlation between lncRNAs and colon adenocarcinoma (COAD) from the perspective of autophagy.

Methods A series of bioinformatics methods were used to explore the correlation between lncRNA and COAD from the perspective of autophagy.

Results Four autophagy-related lncRNAs related to the prognosis of COAD were identified: EB1-AS1, LINC02381, AC011462.4, and AC016876.1. These four lncRNAs may act as oncogenes involved in the occurrence and development of COAD. The prognostic model was established, and the accuracy of the model was verified by the receiver operating characteristic curve. The risk score of the model could independently predict the prognosis of patients and was preferable to other clinical indicators, with higher values indicating a worse prognosis of the patients. Gene Set Enrichment Analysis was performed for these four lncRNAs, which showed that the high expression group of these were enriched in the basal cell carcinoma pathway. To make it more convenient for clinicians to use, we constructed a nomogram based on age and risk score, which can be used to evaluate the one-, three-, and five-year survival rates of patients.

Conclusion These results can help us understand the mechanism of action of lncRNA on COAD from the perspective of autophagy and may provide new directions for the diagnosis and treatment of COAD. The EB1-AS1 gene in this study is a potential candidate biological target for COAD treatment in the future.

Received: 25 May 2021
Revised: 30 August 2021
Accepted: 14 October 2021

Key words: colon adenocarcinoma (COAD); prognostic model; long noncoding RNA (lncRNA); EB1-AS1

According to Chen *et al*'s statistical analysis of cancer data from the National Central Cancer Registry of China from 2009 to 2011, colon cancer was ranked in the top five cancers in China regarding new incidence and mortality [1]. Surgery is still the main treatment for colon cancer, though combinations of radiotherapy, chemotherapy, and neoadjuvant chemotherapy may be used in cases where surgery alone is unable to treat the cancer [2]. With developments in biomedical research in recent years, molecularly targeted drugs have become an option for the non-surgical treatment of patients [3]. However, the current median overall survival rate from colon cancer is only approximately two years [4]. At present, the incidence and mortality of colon cancer are increasing in China; particularly concerning is the increased incidence among young people [5]. Therefore, there is an urgent need to find new treatment targets for colon cancer to improve the prognosis of patients.

Unlike messenger RNA (mRNA), long noncoding RNA (lncRNA) does not participate in gene expression as a template for protein translation but can affect various biological activities in the human body by regulating protein synthesis and was found to be strongly correlated with cancer [6]. There are almost 8000 cancer-related lncRNAs in The Cancer Genome Atlas (TCGA) that can affect different stages of cancer, including cell proliferation, apoptosis, and metastasis [7]. Autophagy is an important metabolic pathway for maintaining homeostasis in the human body and is responsible for the degradation of macromolecular substances, such as damaged organelles and long-lived macromolecular proteins [8]. Studies have reported that autophagy can

✉ Correspondence to: Zhongyin Zhou. Email: 13871029766@163.com

promote the occurrence, development, and metastasis of cancer and is related to the invasiveness of cancer cells [9]. In addition to being regulated by the expression of autophagy-related genes, lncRNAs are also involved in autophagy regulation [10]. The most common type of colon cancer is adenocarcinoma (colon adenocarcinoma, hereafter referred to a COAD) originating from the epithelium of the colonic mucosa. Therefore, understanding autophagy-related lncRNAs and their molecular mechanisms in COAD is beneficial for the treatment of colon cancer and may bring improvements to colon cancer therapy. To achieve this, we analyzed the expression level of autophagy-related genes in COAD tissues that were extracted from the TCGA and used a co-expression (Cox) analysis to obtain the related lncRNAs. The independent prognostic genes that were strongly correlated with COAD were then used to establish a clinically relevant prognosis model. To understand the usefulness of this prognosis model to clinicians, it was compared to other clinically relevant indicators. We also investigated the role of these lncRNAs in COAD using Gene Set Enrichment Analysis (GSEA). Finally, we used the risk score and patient age to prepare a nomogram to predict the survival of patients, whose accuracy we then verified using relevant biological methods.

Materials and methods

Data download and preprocessing

Construction of the lncRNA and mRNA matrix: tissues from COAD patients were obtained from TCGA to construct the lncRNA and mRNA matrix. Since the data used in this study were obtained exclusively from the TCGA database and strictly followed the TCGA publication guidelines (http://cancergenome.nih.gov/abouttcga/policies/publicationguidelines), it was not necessary to obtain the approval of the ethics committee. Autophagy-related genes were obtained from the Human Autophagy Database (HADb), and the expression levels of these genes were extracted from the mRNA matrix. The lncRNA matrix and the expression level of autophagy genes were used for the joint analysis. Autophagy-related lncRNAs were found by setting CorFilter > 0.4 and $P < 0.001$, and the expression level of lncRNAs in the matrix was extracted by Cox analysis.

Building a prognostic signature

The obtained autophagy-related lncRNAs were combined with the clinical data obtained from the TCGA database [including age (divided into > and < 65 years old groups), T (the extent of the primary tumor), M (whether there is distant metastasis), and N (the involvement of local lymph nodes) staging, grading, and other related data]. Currently, clinical prognosis is mainly based on the stage and grade of tumor cells and the patient's age. Univariate cox (unicox) analysis was performed to obtain autophagic lncRNAs related to the prognosis of COAD, and then, multivariate cox (multicox) analysis was performed with the obtained genes to obtain independent prognostic lncRNAs that were strongly related to COAD. These lncRNAs were found to be related to specific clinical indicators by the Clinical correlation analysis, and the specific formula was as follows:

$$\text{The risk score} = \sum_{k=1}^{n} \text{EXP}\,\beta$$

where n represents the number of prognostic lncRNAs, the regression coefficient is β, and EXP is the expression value.

Evaluation of the prognostic signature

By calculating the risk score of all samples in this study, the median expression level was obtained. With the median as the boundary, samples with higher expression levels were defined as the high expression level group, and those with lower expression levels were the low expression level group. The receiver operating characteristic (ROC) curve was drawn to evaluate whether the signature was representative of the groups.

An independent prognostic analysis was used to assess whether the prognostic signature could be used as a prognostic factor independent of the other clinical indicators. In order to better understand how these lncRNAs affect autophagy, Cox analysis of these genes and the mRNA was performed. In addition, the "survival" package in R4.0.3 was used to draw a survival curve to evaluate the impact of a single prognostic lncRNA and risk score on patient survival.

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis

In order to better understand the ways in which these lncRNAs acted, KEGG enrichment analysis was performed for both groups using GSEA 4.1.0 (set at $P < 0.05$, $|ES| > 0.6$, gene size ≥ 100, and false discovery rate < 25%) to identify the biological pathways that were enriched in the high and low groups.

Construction and evaluation of the nomogram

In order to better serve clinicians, we included the age and risk score to draw nomograms and evaluated them. First, the ROC curve was used to evaluate their representativeness, and then a c-index was used to evaluate their predictive ability. Finally, a calibration curve for one, three, and five-year survival rates was established.

Results

The entire research process is shown in Fig. 1.

Screening of autophagy-related lncRNA

Four-hundred and thirty-seven tissues were obtained from the TCGA, of which 398 were COAD tissues and 39 were adjacent tissues. A total of 257 autophagy-related genes were obtained from the HADb, and their respective expression levels were obtained. A total of 922 autophagy-related lncRNAs were yielded from the Cox analysis.

Cox regression analysis and clinical correlation analysis

The unicox analysis showed that 18 lncRNAs were correlated with the prognosis of COAD, and anything with $P < 0.05$ was considered meaningful. Finally, four lncRNAs were obtained by the multicox analysis, namely EB1-AS1, LINC02381, AC011462.4, and AC016876.1. The clinical correlation analysis showed that these four lncRNAs were related to the classification and staging of COAD (Fig. 2), and the higher their expression, the higher the COAD grade and staging level, meaning that these four lncRNAs may act as oncogenes in the occurrence and development of COAD.

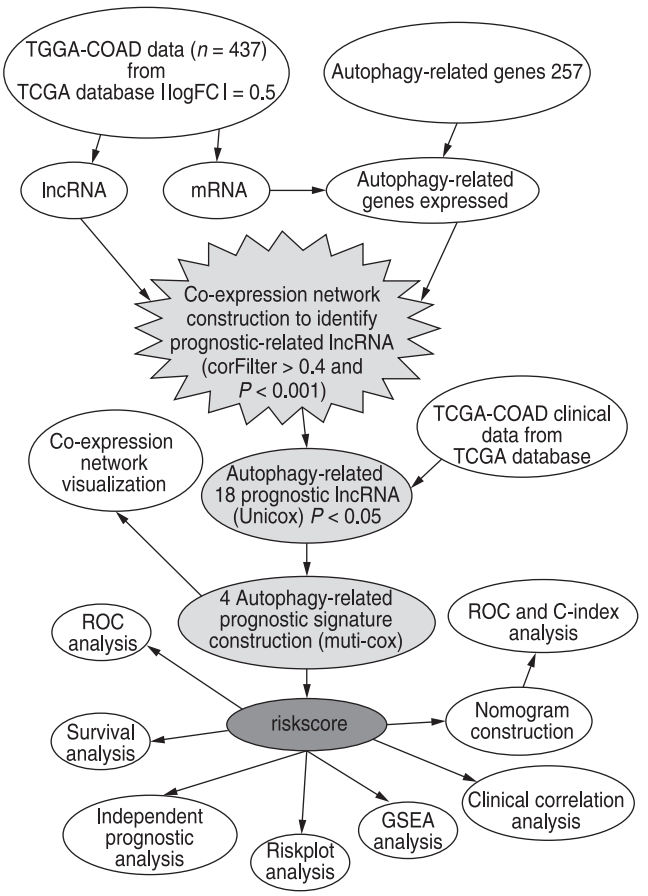


Fig. 1 The entire research process

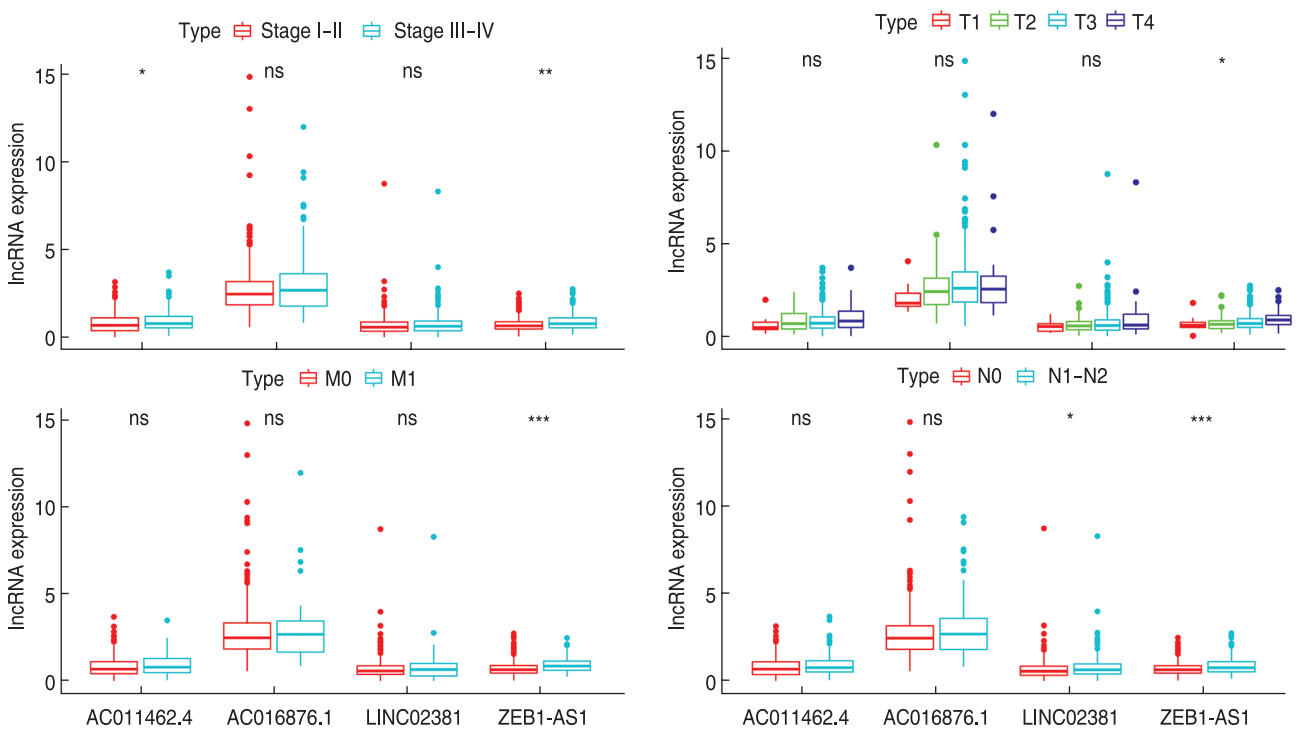


Fig. 2 Clinical correlation analysis, integrating prognostic genes with tumor staging and grading (ns, none significance; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$)

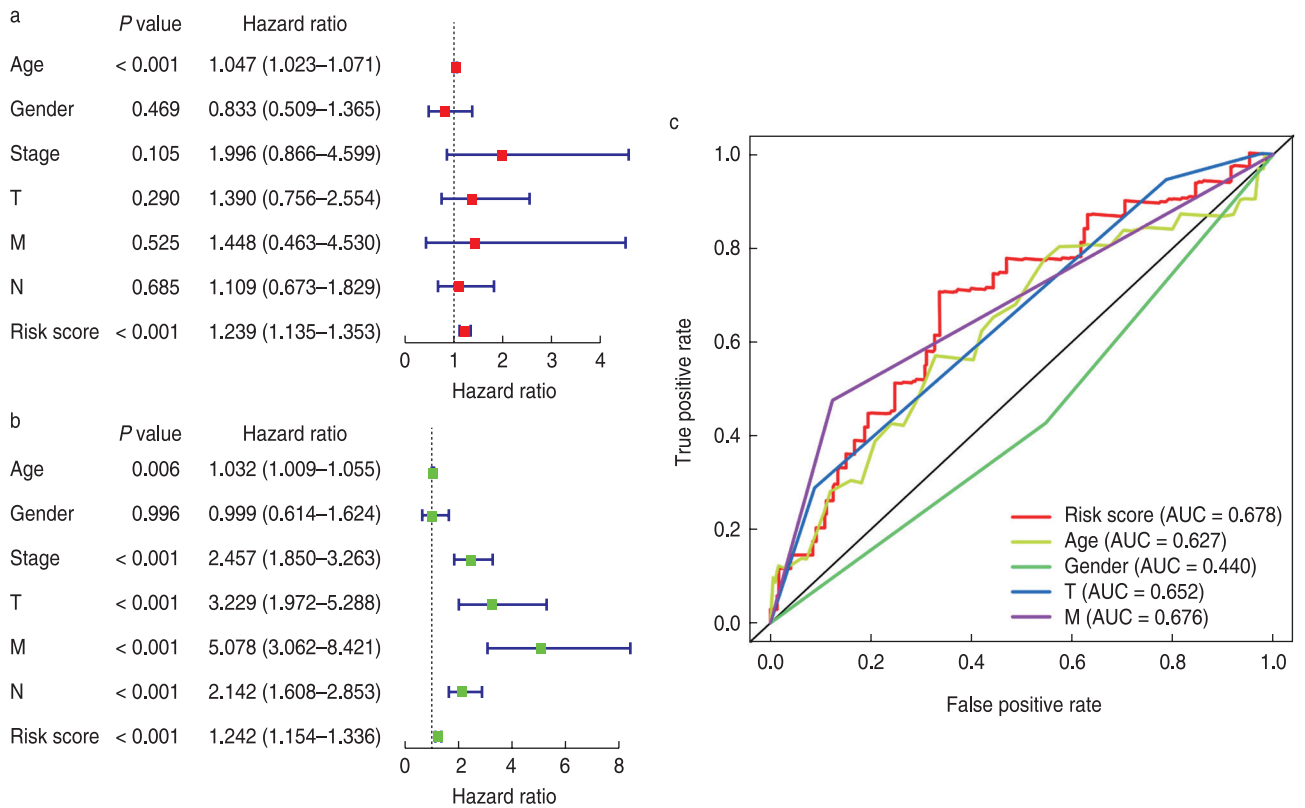


Fig. 3 Receiver operating characteristic (ROC) curve evaluation prognostic model. Independent prognostic analysis of high and low expression groups, which includes age, gender, risk score, tumor stage, and tumor grade

Construction and evaluation of risk signatures

The prognostic model of autophagy-related lncRNAs is given in Fig. 3 (with the left showing the Independent prognostic analysis and significance denoted by $P < 0.05$). Both the high and low expression groups had P-values < than 0.05, indicating that the score could be used to independently judge the prognosis of patients. Age was also an independent prognostic risk factor ($P < 0.05$ in both expression groups). The ROC curve on the right showed that the area under the curve (AUC) of the risk score was 0.678, higher than the tumor cell grade and T/N/M stage, which indicated that the accuracy of the risk score in predicting survival was higher than other clinical features, thus verifying the accuracy of the model.

Construction of the core lncRNA co-expression network

Fig. 4 shows the co-expression network diagram of autophagy-related lncRNAs (left panel), where the diamond block was the prognostic gene and the purple ellipse was the autophagy mRNA. It was evident that the four genes interacted with the autophagy mRNA. The Sankey diagram in the right panel shows the interaction between them more visually. In addition, it showed that these four lncRNAs belonged to the risk group, which was consistent with the data presented in Fig. 2.

Fig. 5 further validates the results displayed in Fig. 4b which includes the survival analysis, risk curve, and heat map, indicating that these genes play a negative role in cancer prognosis.

KEGG enrichment analysis

According to the KEGG analysis of COAD patients (Fig. 6), the high expression group of these lncRNAs was mainly enriched in the basal cell carcinoma pathway, indicating that the basal cell carcinoma pathway may be the carcinogenic target of these lncRNAs. This provides new insights for molecular research of COAD. The pathways for the enrichment of the low expression group were abundant. During the transcription process, it mainly affected DNA replication, pyrimidine metabolism, DNA mismatch repair, nucleotide excision repair, pentose phosphate pathway, RNA metabolism, purine metabolism, and spliceosomes. In the translation process, it mainly affected the cell cycle, proteasomes, protein export and transportation, ribosomes, and AA-tRNA.

Construction and evaluation of a nomogram

To provide a quantitative method for predicting the probability of survival time, we used information from all

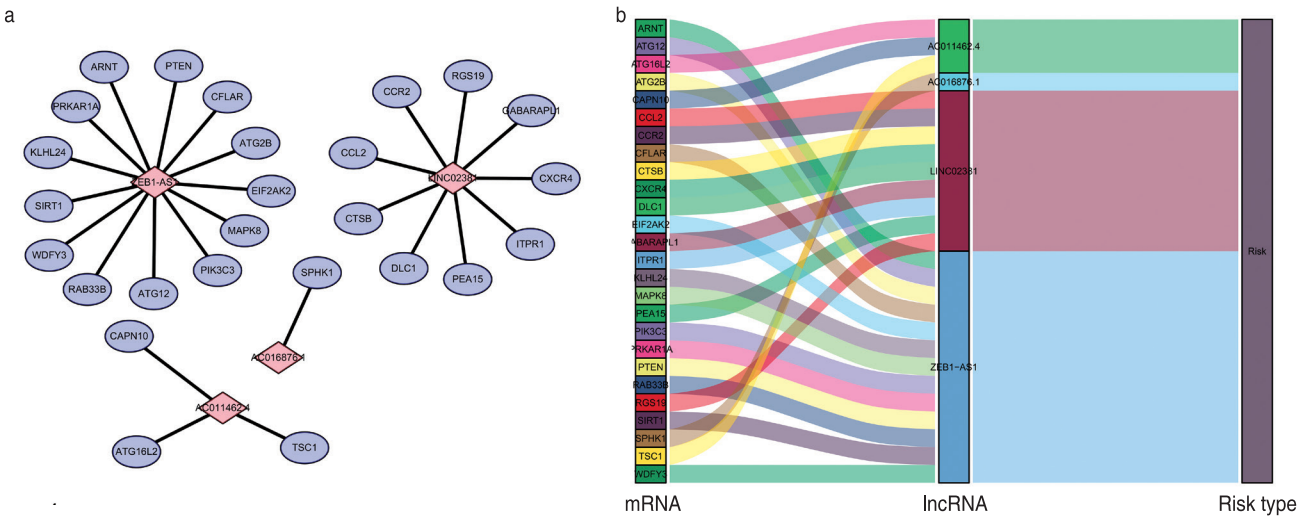


Fig. 4 Co-expression network diagram and Sankey diagram of autophagy-related long noncoding RNAs (lncRNAs). The diamond blocks are prognostic genes, and the purple ellipses are autophagy-related messenger RNAs (mRNAs) showing the autophagy mRNAs with which these four prognostic genes interact. The Sankey diagram (right panel) shows the interaction between these more intuitively. All four genes are in the high-risk group

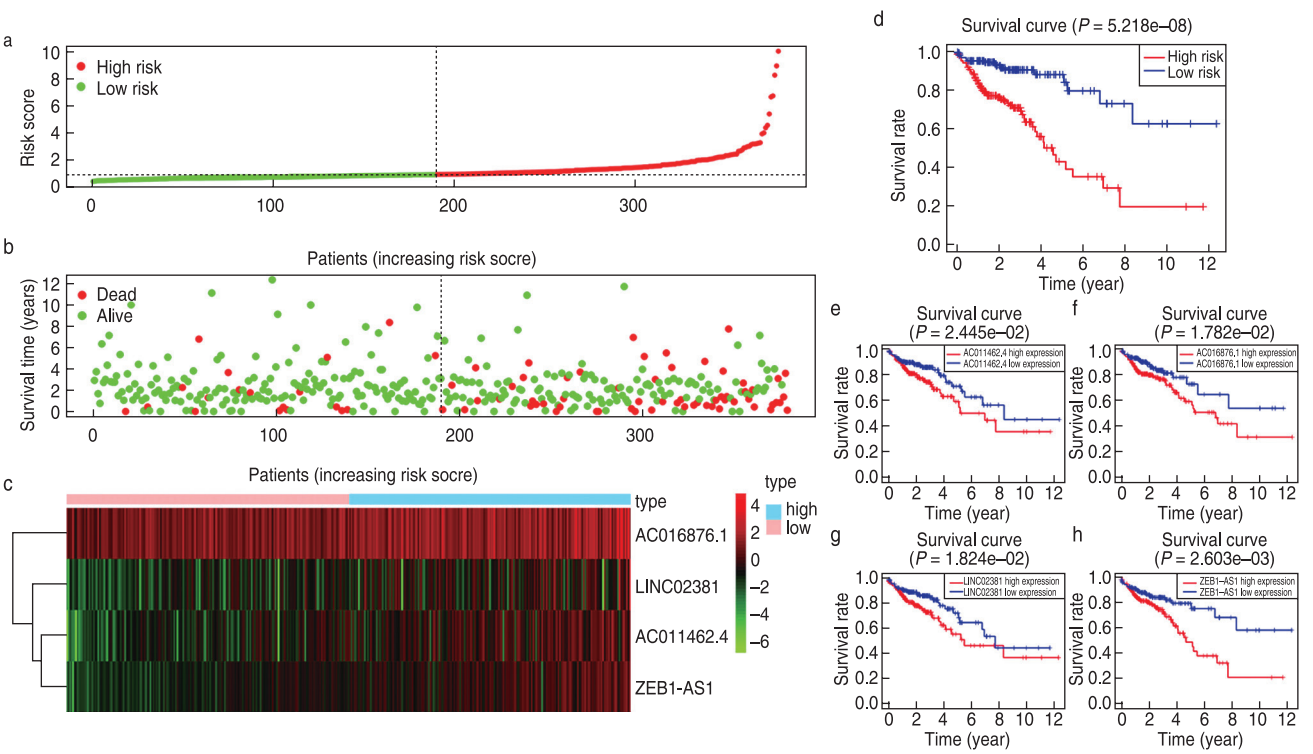


Fig. 5 Correlation between these four prognosis genes and the prognosis of colon adenocarcinoma (COAD) based on survival analysis. The risk curve and heat map further validate the results in Fig. 4b, indicating that these genes play a negative role in cancer prognosis

samples to construct a nomogram that integrated the risk score and patient age (Fig. 7). In the previous steps, we observed that age was related to prognosis, which Aquina *et al*[11] also found in COAD patients. The older the age, the worse the prognosis. Therefore, in the nomogram, we integrated the age and risk score to provide patients with a comprehensive score to better predict the one-year, three-year, and five-year survival rates. The nomogram showed that the risk score made the biggest contribution to the nomogram. The accuracy of the model was verified by the ROC curve (AUC = 0.708) and C-index = 0.691.

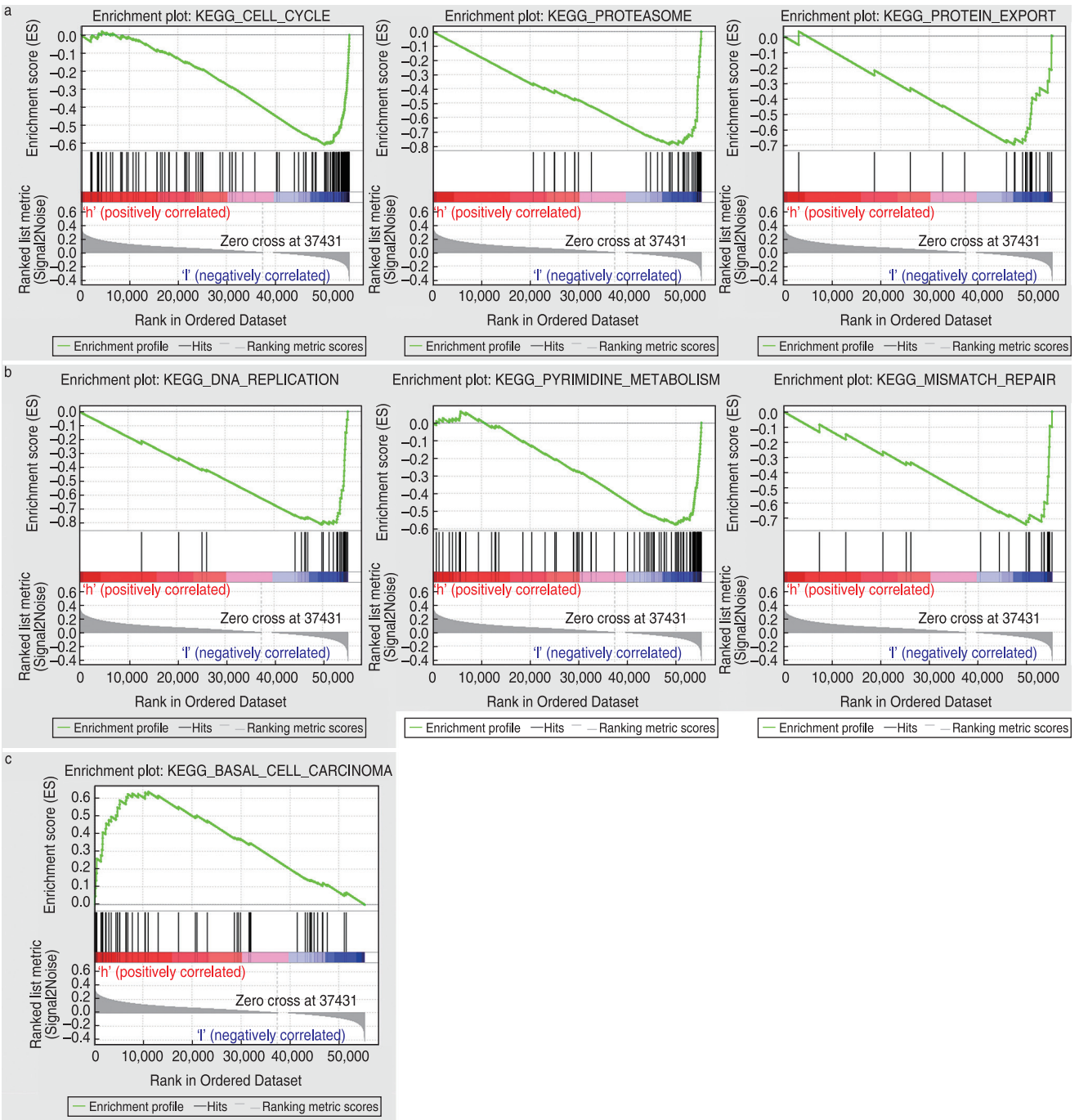


Fig. 6 Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of the four long noncoding RNAs (lncRNAs). High gene expression was related to the basal cell carcinoma pathway, and low expression was enriched in a variety of pathways, including gene transcription and translation

Discussion

Globally, COAD has high morbidity and mortality rates [12]. Current screening methods for colon cancer have shortcomings [13]. The gold standard is colonoscopy, but bowel preparation and contraindications make it unacceptable for many patients. The commonly used intestinal tumor biomarker, carcinoembryonic antigen, lacks sensitivity and specificity. The fecal immunochemical test has a high false positive rate [13]. Therefore, there is an urgent need to develop screening tests that are relevant for prognosis and suitable for all patients to improve the detection rate, prognosis, and five-year survival rate of patients with colon cancer.

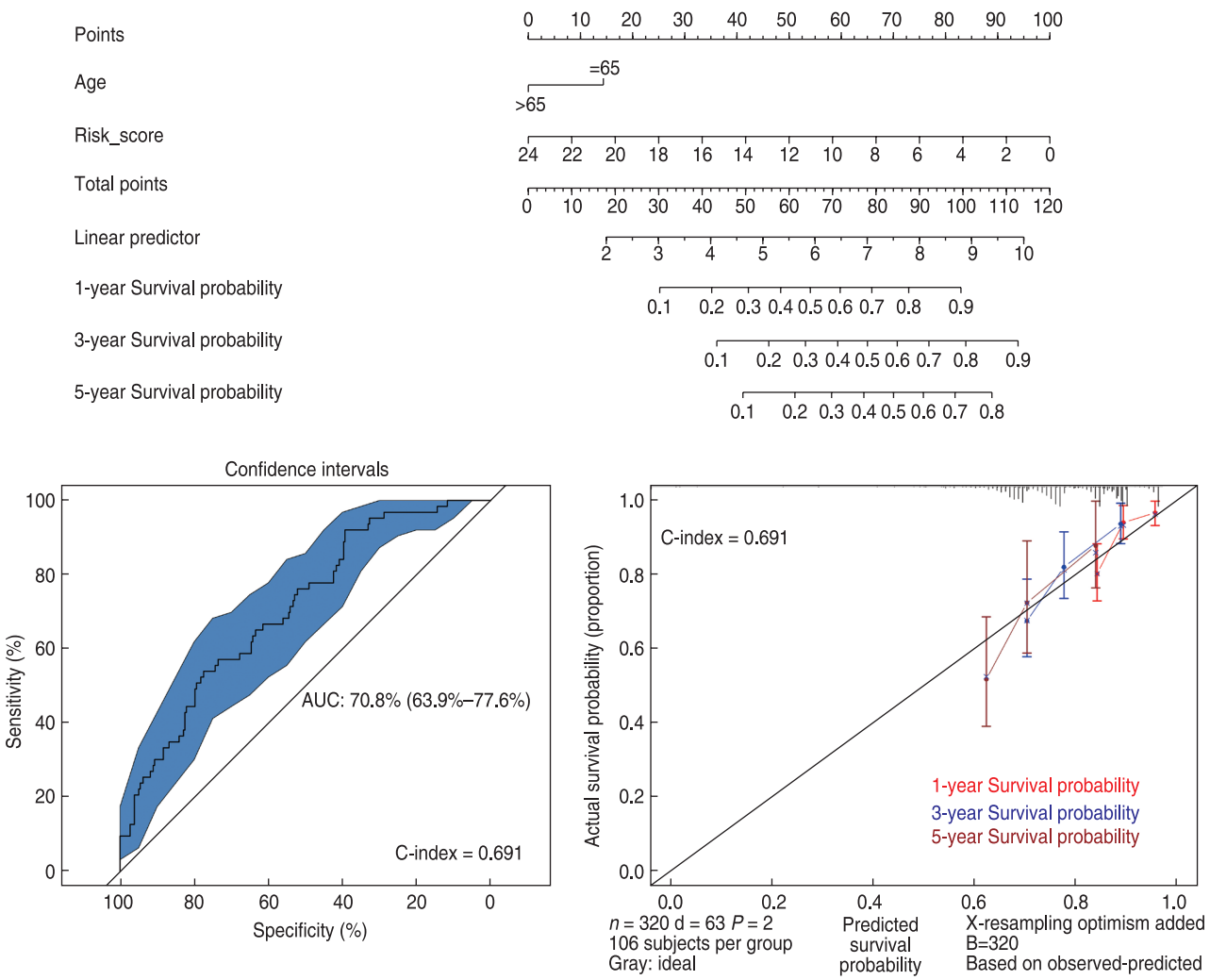


Fig. 7 A nomogram drawn from the risk score and age factors. The patient's value was marked on each axis, and the above variables which include risk score and age were added. The results show that the risk score contributed the most to the nomogram. The receiver operating characteristic (ROC) curve (area under the curve; AUC = 0.708) and C-index (= 0.691) verified the accuracy of the model

Thus, lncRNAs have attracted attention. Firstly, lncRNA detection is performed through taking plasma samples, which are easy to obtain and inexpensive [14]. Secondly, numerous studies have shown that lncRNAs are closely related to the occurrence, development, and metastasis of COAD [15]. The lncRNA activated by TGF-β can promote the epithelial-mesenchymal transition process by inhibiting the expression of E-cad, thereby affecting the occurrence of colon tumors. The lncRNA CASC1 can regulate the miR-4310/LGR5/Wnt/β-catenin signal transduction pathway to promote the proliferation and metastasis of COAD [16]. HOTAIR may be related to distant metastasis and short survival rates from colon cancer by inhibiting the transcription of the tumor suppressor gene miR-34a [17]. Autophagy is a major metabolic pathway in the human body, and many studies have found that it plays an important role in the occurrence and development of colon cancer, although the precise mechanism is not yet clear [18].

Therefore, this study was conducted to explore the relationship between lncRNAs and COAD from the perspective of autophagy. First, relevant genetic information and clinical information in TCGA and HADb were integrated, and four independent prognostic lncRNAs that strongly correlated with COAD were obtained by Cox analysis. These lncRNAs (EB1-AS1, LINC02381, AC011462.4, and AC016876.1) may act like oncogenes. These lncRNAs were used to construct a prognostic signature, and a series of biological processes was used to verify the signatures. Finally, a nomogram was made using risk scores and age, which was converted into specific numbers to predict the one-, three-, and five-year survival of patients.

In this study, EB1-AS1 was the most central among

the four lncRNAs, that is, it had a greater prognostic correlation than the other lncRNAs. There have been many studies in recent years on the relationship between EB1-AS1 and COAD, but these have identified differences in the specific ways of action. For example, ZEB1-AS1 can promote cancer by binding to Mir-181A-5p and inhibiting the microRNA (miRNA)-induced β-catenin inhibitory pathway [19]; ZEB1-AS1 can also inhibit miR-101 to promote the proliferation and metastasis of cancer cells. The expression level was positively correlated with the histological grade and T stage of the cancer, that is, the higher the expression level, the worse the prognosis of the patients [20]. Current studies mainly focus on the interaction between ZEB1-AS1 and miRNA. MicroRNA is a kind of non-coding RNA that generally acts as a tumor suppressor gene, probably mainly by inhibiting transcription or mediating degradation [21]. In this study, it was found that the autophagy effect of ZEB1-AS1, namely the interaction between lncRNA and mRNA, may play a very important role in the occurrence and development of COAD, which is a novel result from this study.

Studies have found that LINC02381 may play an inhibitory role in COAD by regulating the PI3K-Akt signaling pathway [22]. In this study, the expression level of this lncRNA was inversely related to survival, indicating that it may act as an oncogene. In another study of autophagy-related lncRNAs, this gene was also considered to act as an oncogene [23]. This indicates that lncRNA may affect the growth of COAD cells through a variety of ways. Whether it is inhibited or enhanced in cancer cells may be tissue-specific. As far as we are aware, no published studies currently exist for the other two lncRNAs (AC011462.4 and AC016876.1).

Based on the GSEA, we know that high expression of these four genes may play a role in promoting tumor cell recurrence and metastasis through the basal cell carcinoma pathway. A bioinformatics analysis previously showed that the basal cell carcinoma pathway may be the oncogenic target of lncRNAs [24]. However, there is a lack of relevant experiments to confirm this, so this is a direction we can consider in the future. The lncRNAs from our study differ from those in previous studies [23, 25–26], but we believe our findings may be more credible. The reasons for this are: (1) our study only identified four lncRNAs, which was less than in other studies, indicating that our study found more core lncRNAs; (2) our study identified one core lncRNA, the involvement of which in COAD has been confirmed in many experiments and is likely to be a future biological target of COAD treatment; (3) half of the genes we identified (EB1-AS1 and LINC02381) have been experimentally proven to be related to COAD.

Conclusion

In summary, four independent prognostic lncRNAs related to COAD were found in this study, and among these, EB1-AS1 is very likely to be a new biological target for COAD treatment. Moreover, these four lncRNAs were used to construct a prognostic signature that was superior to the prognostic indicators currently used in clinical practice. Finally, the possible carcinogenic pathways of these four lncRNAs were determined through the enrichment analysis. The EB1-AS1 gene and basal cell carcinoma pathway were specifically daintified and will be the focus of the future research direction of our team, and relevant experiments will be carried out to verify their roles in COAD.

Conflicts of interest

The authors indicated no potential conflicts of interest.

References

1. Chen WQ, Zheng RS, Baade PD, *et al*. Cancer statistics in China, 2015. CA Cancer J Clin, 2016, 66: 115–132.
2. Thanikachalam K, Khan G. Colorectal cancer and nutrition. Nutrients, 2019, 11: 164.
3. Boyle KA, Van Wickle J, Hill RB, *et al*. Mitochondria-targeted drugs stimulate mitophagy and abrogate colon cancer cell proliferation. J Biol Chem, 2018, 293: 14891–14904.
4. Zhu CQ, Zhang HB, Li W, *et al*. Suppress orthotopic colon cancer and its metastasis through exact targeting and highly selective drug release by a smart nanomicelle. Biomaterials, 2018, 161: 144–153.
5. Zhang Y, Shi JF, Huang HY, *et al*. Burden of colorectal cancer in China. Chin J Epidemiol (Chinese), 2015, 36: 709–714.
6. Wei YX, Yang H, Liu XQ. Diagnostic value of lncRNAs as potential biomarkers for oral squamous cell carcinoma diagnosis: a meta-analysis. Oncol Transl Med, 2021, 7: 123–129.
7. Evans JR, Feng FY, Chinnaiyan AM. The bright side of dark matter: lncRNAs in cancer. J Clin Invest, 2016, 126: 2775–2782.
8. Li YJ, Lei YH, Yao N, *et al*. Autophagy and multidrug resistance in cancer. Chin J Cancer, 2017, 36: 52.
9. Dower CM, Wills CA, Frisch SM, *et al*. Mechanisms and context underlying the role of autophagy in cancer metastasis. Autophagy, 2018, 14: 1110–1128.
10. Li P, He J, Yang Z, *et al*. *ZNNT1* long noncoding RNA induces autophagy to inhibit tumorigenesis of uveal melanoma by regulating key autophagy gene expression. Autophagy, 2020, 16: 1186–1199.
11. Aquina CT, Mohile SG, Tejani MA, *et al*. The impact of age on complications, survival, and cause of death following colon cancer surgery. Br J Cancer, 2017, 116: 389–397.
12. Ye YQ, Gu BB, Wang Y, *et al*. E2F1-mediated MNX1-AS1-miR-218-5p-SEC61A1 feedback loop contributes to the progression of colon adenocarcinoma. J Cell Biochem, 2019, 120: 6145–6153.
13. Xu W, Zhou G, Wang HZ, *et al*. Circulating lncRNA SNHG11 as a novel biomarker for early diagnosis and prognosis of colorectal cancer. Int J Cancer, 2020, 146: 2901–2912.
14. Rho JH, Ladd JJ, Li CI, *et al*. Protein and glycomic plasma markers for early detection of adenoma and colon cancer. Gut, 2018, 67: 473–484.

15. Xiong W, Qin JY, Cai XY, *et al*. Overexpression LINC01082 suppresses the proliferation, migration and invasion of colon cancer. Mol Cell Biochem, 2019, 462: 33–40.
16. Lv J, Guo Y, Yan LL, *et al*. Development and validation of a five-lncRNA signature with prognostic value in colon cancer. J Cell Biochem, 2019, Nov 3, doi: 10.1002/jcb.29518. Online ahead of print.
17. Peng CL, Zhao XJ, Wei CC, *et al*. LncRNA HOTAIR promotes colon cancer development by down-regulating miRNA-34a. Eur Rev Med Pharmacol Sci, 2019, 23: 5752–5761.
18. Devenport SN, Shah YM. Functions and implications of autophagy in colon cancer. Cells, 2019, 8: 1349.
19. Lv SY, Shan TD, Pan XT, *et al*. The lncRNA ZEB1-AS1 sponges miR-181a-5p to promote colorectal cancer cell proliferation by regulating Wnt/β-catenin signaling. Cell Cycle, 2018, 17: 1245–1254.
20. Xiong WC, Han N, Wu N, *et al*. Interplay between long noncoding RNA ZEB1-AS1 and miR-101/ZEB1 axis regulates proliferation and migration of colorectal cancer cells. Am J Transl Res, 2018, 10: 605–617.
21. Jin Z, Chen B. LncRNA ZEB1-AS1 regulates colorectal cancer cells by miR-205/YAP1 axis. Open Med (Wars), 2020, 15: 175–184.
22. Jafarzadeh M, Soltani BM, Soleimani M, *et al*. Epigenetically silenced LINC02381 functions as a tumor suppressor by regulating PI3K-Akt signaling pathway. Biochimie, 2020, 171–172: 63–71.
23. Zhou WG, Zhang SJ, Li HB, *et al*. Development of prognostic indicator based on autophagy-related lncRNA analysis in colon adenocarcinoma. Biomed Res Int, 2020, 2020: 9807918.
24. Wu MX, Kim KY, Park WC, *et al*. Enhanced expression of GABRD predicts poor prognosis in patients with colon adenocarcinoma. Transl Oncol, 2020, 13: 100861.
25. Huang W, Liu Z, Li Y, *et al*. Identification of long noncoding RNAs biomarkers for diagnosis and prognosis in patients with colon adenocarcinoma. J Cell Biochem, 2019, 120: 4121–4131.
26. Qian K, Huang HY, Jiang J, *et al*. Identifying autophagy gene-associated module biomarkers through construction and analysis of an autophagy-mediated ceRNA-ceRNA interaction network in colorectal cancer. Int J Oncol, 2018, 53: 1083–1093.

DOI 10.1007/s10330-021-0497-7
Cite this article as: Tan FF, Zhou ZY. Autophagy-related lncRNA and its related mechanism in colon adenocarcinoma. Oncol Transl Med, 2021, 7: 305–313.

中国科技核心期刊

（中国科技论文统计源期刊）

收录证书

CERTIFICATE OF SOURCE JOURNAL

FOR CHINESE SCIENTIFIC AND TECHNICAL PAPERS AND CITATIONS

ONCOLOGY AND TRANSLATIONAL MEDICINE

经过多项学术指标综合评定及同行专家评议推荐，贵刊被收录为“中国科技核心期刊”（中国科技论文统计源期刊）。

特颁发此证书。

中国科学技术信息研究所

Institute of Scientific and Technical Information of China

北京复兴路 15 号 100038　　www.istic.ac.cn

2021年12月

证书编号：2020-G616-0122

有效期至：2022年12月

Call For Papers
Oncology and Translational Medicine

(CN 42-1865/R, ISSN 2095-9621)

Dear Authors,

Oncology and Translational Medicine (OTM), a peer-reviewed open-access journal, is very interested in your study. If you have unpublished papers in hand and have the idea of making our journal a vehicle for your research interests, please feel free to submit your manuscripts to us via the Paper Submission System.

Aims & Scope

- Lung Cancer
- Liver Cancer
- Pancreatic Cancer
- Gastrointestinal Tumors
- Breast Cancer
- Thyroid Cancer
- Bone Tumors
- Genitourinary Tumors
- Brain Tumor
- Blood Diseases
- Gynecologic Oncology
- ENT Tumors
- Skin Cancer
- Cancer Translational Medicine
- Cancer Imageology
- Cancer Chemotherapy
- Radiotherapy
- Tumors Psychology
- Other Tumor-related Contents

Contact Us
Editorial office of Oncology and Translational Medicine
Tongji Hospital
Tongji Medical College
Huazhong University of Science and Technology
Jie Fang Da Dao 1095
430030 Wuhan, China
Tel.: 86-27-69378388
Email: dmedizin@tjh.tjmu.edu.cn;
dmedizin@sina.com

Oncology and Translational Medicine (OTM) is sponsored by Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, China (English, bimonthly).

OTM mainly publishes original and review articles on oncology and translational medicine. We are working with the commitment to bring the highest quality research to the widest possible audience and share the research work in a timely fashion.

Manuscripts considered for publication include regular scientific papers, original research, brief reports and case reports. Review articles, commentaries and letters are welcome.

About Us

- Peer-reviewed
- Rapid publication
- Online first
- Open access
- Both print and online versions

For more information about us, please visit:
http://otm.tjh.com.cn