GENETICS AND MOLECULAR MECHANISMS OF ORAL AND ESOPHAGEAL SQUAMOUS CELL CARCINOMA

EDITED BY: Bin Qiao, Di Wu and Die Wang
PUBLISHED IN: Frontiers in Oncology

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-576-0
DOI 10.3389/978-2-88974-576-0

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

GENETICS AND MOLECULAR MECHANISMS OF ORAL AND ESOPHAGEAL SQUAMOUS CELL CARCINOMA

Topic Editors:
Bin Qiao, Zhengzhou University, China
Di Wu, University of North Carolina at Chapel Hill, United States
Die Wang, Hudson Institute of Medical Research, Australia

Citation: Qiao, B., Wu, D., Wang, D., eds. (2022). Genetics and Molecular Mechanisms of Oral and Esophageal Squamous Cell Carcinoma. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-576-0

Table of Contents

frontiers | Frontiers in Oncology

EDITORIAL
published: 06 April 2022
doi: 10.3389/fonc.2022.874353

Editorial: Genetics and Molecular Mechanisms of Oral and Esophageal Squamous Cell Carcinoma

Bin Qiao[1*†], *Shuaize Li*[1†], *Die Wang*[2] *and Di Wu*[3]

[1] Department of Stomatology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [2] Centre for Cancer Research, Hudson Institute of Medical Research, Melbourne, VIC, Australia, [3] Department of Periodontology, School of Dentistry, The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Keywords: oral cancer, precision medicine, genome sequencing, whole exome sequencing, targeted therapy

Editorial on the Research Topic

Genetics and Molecular Mechanisms of Oral and Esophageal Squamous Cell Carcinoma

Oral squamous cell carcinoma (OSCC) is the most common histopathological type of oral cancer, with typical characteristics of low 5-year survival rate and poor prognosis (1, 2). Importantly, there are many factors affecting its occurrence and progression, in which genome alterations are critical indicators of the proper diagnosis and treatment (3). Carcinogenesis is a multi-step process, which involves the accumulation of genetic and epigenetic changes of oncogenes or tumor suppressor genes (4).Therefore, better understanding of the genetic and molecular disorders of the disease is the key to early diagnosis, appropriate treatment and improving the prognosis of patients.

Nowadays, cancer treatment is developed more towards personalized and targeted treatment. The most widely used treatment are targeted immunotherapy, which have significantly improved the 5-year survival rate of many types of cancer (5, 6). However, for OSCC patients, the only approved targeted therapy is a monoclonal antibody against epidermal growth factor receptor (EGFR), with the trade name 'cetuximab' (7). Recently, two immunotherapeutic agents, i.e., pembrolizumab and nivolumab, have been approved for OSCC (8, 9). Nevertheless, patients are widely resistant to the targeted therapy such as cetuximab combined with radiotherapy (10), and only less than 20% of OSCC patients receiving immunotherapy have achieved lasting remission (11). Therefore, it is necessary to implement different treatment schemes for patients based on their different gene mutations. Fortunately, the recent development of high throughput sequencing technologies, including whole genome sequencing and whole exome sequencing, make the detection of gene mutations in tumor tissues more sensitive and comprehensive so that the personalized cancer treatment becomes possible (12). This personalized treatment topic focused on the two genes *EGFR* and *TP53* most commonly mutant in OSCC. In the following, we will demonstrate the rationale and the existing dilemma of targeted therapy of OSCC based on these two genes in the prospect of gene sequencing technology served for precision medicine.

It is reported that *EGFR* is overexpressed in more than 90% of OSCC patients and is involved in tumor cell invasion and metastasis (13). The activation of *EGFR* leads to the phosphorylation and activation of downstream signal transduction mediators and promotes tumor cell proliferation, survival, angiogenesis, invasion and adhesion (14). A variety of strategies to block *EGFR* function have been developed as personalized methods to inhibit tumor growth and metastasis, in which, cetuximab is the only targeted drug approved in OSCC. It has also been widely used and studied in

OPEN ACCESS

Edited and reviewed by:
Heather Cunliffe,
University of Otago, New Zealand

***Correspondence:**
Bin Qiao
qiaobin@zzu.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *12 February 2022*
Accepted: *09 March 2022*
Published: *06 April 2022*

Citation:
Qiao B, Li S, Wang D and Wu D (2022) Editorial: Genetics and Molecular Mechanisms of Oral and Esophageal Squamous Cell Carcinoma. Front. Oncol. 12:874353. doi: 10.3389/fonc.2022.874353

patients with locally advanced OSCC and patients with recurrent and or metastatic OSCC. However, mutations that activate EGFR kinase activity are relatively rare in OSCC (15). In addition, SRC is a nonreceptor tyrosine kinase. It is involved in regulating cell signal transduction downstream of a variety of receptors, including members of EGFR family, and in the regulation of cell proliferation, migration, adhesion and apoptosis (16). SRC kinase activity also enhances EGFR signal transduction (17). Therefore, SRC activity may promote resistance to *EGFR* targeted personalized therapy through independent activation or association with other receptors (17). Therefore, the determination of SRC kinase activity may be the key to predict the possible positive clinical response of targeted therapy.

Another important gene is TP53, which regulates cell cycle and apoptosis induced by DNA damage (18). Studies have shown that TP53 regulates the expression of forkhead box M1 (FOXM1) transcription factor and can directly bind and inactivate Aurora kinase A (AURKA) (19). FOXM1, an important cell cycle mediator, is a transcription factor downstream of EGFR/PI3K/AKT cascade and controls cell survival, apoptosis, migration and angiogenesis (20). In addition, AURKA and AURKB are two cell cycle regulators controlled by FOXM1 (21). AURKA and AURKB both control the structure and function of cytoskeleton and chromosome and contribute to tumor progression, metastasis and diffusion (22). EGFR signaling pathway can improve the translation and transcription efficiency of AURKA and induce the overexpression of AURKA (23). As an important cell cycle regulator, Aurora kinase is a reliable target in a variety of malignant tumors. At present, several Aurora kinase inhibitors have been developed (24, 25). In preclinical evaluation studies, it was found that AURKA and AURKB inhibitors ENMD2076 and AZD1152, as well as pan Aurora agents such as AMG900, can induce growth arrest and apoptosis (26, 27). In a phase I/II study, laser kinase inhibitors were evaluated as a single drug for a variety of solid tumors including OSCC. However, only 3 of 20 OSCC patients receiving AMG900 had partial remission. This low success rate suggests that the overall the personalized intervention needs to be improved at the patient population, and meanwhile those patients who are more sensitive to Aurora kinase inhibitors also need to have further investigator e.g., identify the biomarkers they may share to indicate the efficacy of these compounds, so as to achieve better clinical results from this target therapy.

As a summary, one of the ultimate goals of cancer research is to better understand the disease-related biological process to identify the predictive biomarkers, which runs through the whole process of patient diagnosis, prognosis and treatment. The effect of clinical treatment often depends on the existence of specific cell targets. Despite the complexity of cancer genetics, tumor heterogeneity and drug resistance are still the difficulties of targeted therapy. However, the development of genomics related technologies, including whole genome sequencing and whole exome sequencing, has had a far-reaching impact on the personalized diagnosis and treatment of cancer patients (28).

AUTHOR CONTRIBUTIONS

BQ and SL wrote the manuscript. DWa and DWu made the draft revision. All authors contributed to the article and approved the submitted version.

REFERENCES

1. Hedberg ML, Goh G, Chiosea SI. Genetic Landscape of Metastatic and Recurrent Head and Neck Squamous Cell Carcinoma. *J Clin Invest* (2016) 126(4):1606. doi: 10.1172/JCI86862
2. Lee YS, Johnson DE, Grandis JR. An Update: Emerging Drugs to Treat Squamous Cell Carcinomas of the Head and Neck. *Expert Opin Emerg Drugs* (2018) 23(4):283–99. doi: 10.1080/14728214.2018.1543400
3. Wang Y, Ow TJ, Myers JN. Pathways for Cervical Metastasis in Malignant Neoplasms of the Head and Neck Region. *Clin Anat* (2012) 25(1):54–71. doi: 10.1002/ca.21249
4. Chow LQM. Head and Neck Cancer. *N Engl J Med* (2020) 382(1):60–72. doi: 10.1056/NEJMra1715715
5. Ferris RL. Immunology and Immunotherapy of Head and Neck Cancer. *J Clin Oncol* (2015) 33(29):3293–304. doi: 10.1200/JCO.2015.61.1509
6. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N Engl J Med* (2017) 377(25):2500–1. doi: 10.1056/NEJMc1713444
7. Vermorken JB, Trigo J, Hitt R, Koralewski P, Diaz-Rubio E, Rolland F, et al. Open-Label, Uncontrolled, Multicenter Phase II Study to Evaluate the Efficacy and Toxicity of Cetuximab as a Single Agent in Patients With Recurrent and/or Metastatic Squamous Cell Carcinoma of the Head and Neck Who Failed to Respond to Platinum-Based Therapy. *J Clin Oncol* (2007) 25(16):2171–7. doi: 10.1200/JCO.2006.06.7447
8. De Felice F, Musio D, Tombolini V. Immune Check-Point Inhibitors and Standard Chemoradiotherapy in Definitive Head and Neck Cancer Treatment. *J Pers Med* (2021) 11(5):393. doi: 10.3390/jpm11050393
9. Chow LQM, Haddad R, Gupta S, Mahipal A, Mehra R, Tahara M, et al. Antitumor Activity of Pembrolizumab in Biomarker-Unselected Patients With Recurrent and/or Metastatic Head and Neck Squamous Cell Carcinoma: Results From the Phase Ib KEYNOTE-012 Expansion Cohort. *J Clin Oncol* (2016) 34(32):3838–45. doi: 10.1200/JCO.2016.68.1478
10. Machiels JP, Schmitz S. Molecular-Targeted Therapy of Head and Neck Squamous Cell Carcinoma: Beyond Cetuximab-Based Therapy. *Curr Opin Oncol* (2011) 23(3):241–8. doi: 10.1097/CCO.0b013e328344f581
11. Wang Z, Goto Y, Allevato MM, Wu VH, Saddawi-Konefka R, Gilardi M, et al. Disruption of the HER3-PI3K-mTOR Oncogenic Signaling Axis and PD-1 Blockade as a Multimodal Precision Immunotherapy in Head and Neck Cancer. *Nat Commun* (2021) 12(1):2383. doi: 10.1038/s41467-021-22619-w
12. Kim DH, Kim YS, Son NI, Kang CK, Kim AR. Recent Omics Technologies and Their Emerging Applications for Personalised Medicine. *IET Syst Biol* (2017) 11(3):87–98. doi: 10.1049/iet-syb.2016.0016
13. Psyrri A, Yu Z, Weinberger PM, Sasaki C, Haffty B, Camp R, et al. Quantitative Determination of Nuclear and Cytoplasmic Epidermal Growth Factor Receptor Expression in Oropharyngeal Squamous Cell Cancer by Using Automated Quantitative Analysis. *Clin Cancer Res* (2005) 11(16):5856–62. doi: 10.1158/1078-0432.CCR-05-0420
14. Rubin Grandis J, Zeng Q, Drenning SD. Epidermal Growth Factor Receptor–Mediated Stat3 Signaling Blocks Apoptosis in Head and Neck Cancer. *Laryngoscope* (2000) 110(5):868–74. doi: 10.1097/00005537-200005000-00016
15. Lee JW, Soung YH, Kim SY. Somatic Mutations of EGFR Gene in Squamous Cell Carcinoma of the Head and Neck. *Clin Cancer Res* (2005) 11(8):2879–82. doi: 10.1158/1078-0432.CCR-04-2029
16. Frame MC. Newest Findings on the Oldest Oncogene; How Activated Src Does it. *J Cell Sci* (2004) 117(7):989–98. doi: 10.1242/jcs.01111
17. Maa MC, Leu TH, McCarley DJ, Schatzman RC, Parsons SJ. Potentiation of Epidermal Growth Factor Receptor-Mediated Oncogenesis by C-Src:

Implications for the Etiology of Multiple Human Cancers. *Proc Natl Acad Sci USA* (1995) 92(15):6981–5. doi: 10.1073/pnas.92.15.6981
18. Wang S, Zhang Y, Huang J, Wong CC, Zhai J, Li C, et al. TRIM67 Activates P53 to Suppress Colorectal Cancer Initiation and Progression. *Cancer Res* (2019) 79(16):4086–98. doi: 10.1158/0008-5472.CAN-18-3614
19. Chen SS, Chang PC, Cheng YW, Tang FM, Lin YS. Suppression of the STK15 Oncogenic Activity Requires a Transactivation-Independent P53 Function. *EMBO J* (2002) 21(17):4491–9. doi: 10.1093/emboj/cdf409
20. Halasi M, Gartel AL. FOX(M1) News–It Is Cancer. *Mol Cancer Ther* (2013) 12 (3):245–54. doi: 10.1158/1535-7163.MCT-12-0712
21. Yang N, Wang C, Wang Z, Zona S, Lin S-X, Wang X, et al. FOXM1 Recruits Nuclear Aurora Kinase A to Participate in a Positive Feedback Loop Essential for the Self-Renewal of Breast Cancer Stem Cells. *Oncogene* (2017) 36 (24):3428–40. doi: 10.1038/onc.2016.490
22. Wang IC, Chen YJ, Hughes D, Petrovic V, Major ML, Park HJ, et al. Forkhead Box M1 Regulates the Transcriptional Network of Genes Essential for Mitotic Progression and Genes Encoding the SCF (Skp2-Cks1) Ubiquitin Ligase. *Mol Cell Biol* (2005) 25(24):10875–94. doi: 10.1128/MCB.25.24.10875-10894.2005
23. Hung LY, Tseng JT, Lee YC, Xia W, Wang Y-N, Wu M-L, et al. Nuclear Epidermal Growth Factor Receptor (EGFR) Interacts With Signal Transducer and Activator of Transcription 5 (STAT5) in Activating Aurora-A Gene Expression. *Nucleic Acids Res* (2008) 36(13):4337–51. doi: 10.1093/nar/gkn417
24. Tatsuka M, Yu Z, Weinberger PM, Sasaki C, Haffty B, Camp R, et al. Overexpression of Aurora-A Potentiates HRAS-Mediated Oncogenic Transformation and is Implicated in Oral Carcinogenesis. *Oncogene* (2005) 24(6):1122–7. doi: 10.1038/sj.onc.1208293
25. Pannone G, Hindi SA, Santoro A, Sanguedolce F, Rubini C, Cincione RI, et al. Aurora B Expression as a Prognostic Indicator and Possible Therapeutic Target in Oral Squamous Cell Carcinoma. *Int J Immunopathol Pharmacol* (2011) 24(1):79–88. doi: 10.1177/039463201102400110
26. Fletcher GC, Brokx RD, Denny TA, Hembrough TA, Plum SM, Fogler WE, et al. ENMD-2076 Is An Orally Active Kinase Inhibitor With Antiangiogenic and Antiproliferative Mechanisms of Action. *Mol Cancer Ther* (2011) 10 (1):126–37. doi: 10.1158/1535-7163.MCT-10-0574
27. Payton M, Cheung HK, Ninniri MSS, Marinaccio C, Wayne WC, Hanestad K, et al. Dual Targeting of Aurora Kinases With AMG 900 Exhibits Potent Preclinical Activity Against Acute Myeloid Leukemia With Distinct Post-Mitotic Outcomes. *Mol Cancer Ther* (2018) 17(12):2575–85. doi: 10.1158/1535-7163.MCT-18-0186
28. Biankin AV, Hudson TJ. Somatic Variation and Cancer: Therapies Lost in the Mix. *Hum Genet* (2011) 130(1):79–91. doi: 10.1007/s00439-011-1010-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 23 March 2021
doi: 10.3389/fonc.2021.629916

PTIP Inhibits Cell Invasion in Esophageal Squamous Cell Carcinoma *via* Modulation of EphA2 Expression

Xiao Han[1†], Yaning Zhu[2†], Li Shen[3], Yu Zhou[4], Liqun Pang[5], Wubi Zhou[2], Hao Gu[1], Kairong Han[1], Yijun Yang[6], Chao Jiang[4], Jun Xie[7], Chengwan Zhang[1] and Lianshu Ding[8*]*

[1] Department of Central Laboratory, The Affiliated Huaian No. 1 People's Hospital of Nanjing Medical University, Huai'an, China, [2] Department of Pathology, The Affiliated Huaian No. 1 People's Hospital of Nanjing Medical University, Huai'an, China, [3] Department of Hematology, The Affiliated Huaian No. 1 People's Hospital of Nanjing Medical University, Huai'an, China, [4] Department of Medical Oncology, The Affiliated Huaian No. 1 People's Hospital of Nanjing Medical University, Huai'an, Jiangsu, China, [5] Department of General Surgery, The Affiliated Huaian No. 1 People's Hospital of Nanjing Medical University, Huai'an, China, [6] Department of Obstetrics and Gynecology, The Affiliated Huaian No. 1 People's Hospital of Nanjing Medical University, Huai'an, China, [7] Department of Clinical Laboratory, Xuyi People's Hospital, Huai'an, China, [8] Department of Neurosurgery, The Affiliated Huaian No. 1 People's Hospital of Nanjing Medical University, Huai'an, China

OPEN ACCESS

Edited by:
Bin Qiao, Zhengzhou University, China

Reviewed by:
Deanna Edwards, Vanderbilt University Medical Center, United States
Arsheed A. Ganaie, University of Minnesota Twin Cities, United States

***Correspondence:**
Chengwan Zhang
hayyzhchw@njmu.edu.cn
Lianshu Ding
dlshu@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *16 November 2020*
Accepted: *05 March 2021*
Published: *23 March 2021*

Citation:
Han X, Zhu Y, Shen L, Zhou Y, Pang L, Zhou W, Gu H, Han K, Yang Y, Jiang C, Xie J, Zhang C and Ding L (2021) PTIP Inhibits Cell Invasion in Esophageal Squamous Cell Carcinoma via Modulation of EphA2 Expression. Front. Oncol. 11:629916. doi: 10.3389/fonc.2021.629916

Esophageal squamous cell carcinoma (ESCC) is a highly aggressive malignancy and treatment failure is largely due to metastasis and invasion. Aberrant tumor cell adhesion is often associated with tumor progression and metastasis. However, the exact details of cell adhesion in ESCC progression have yet to be determined. In our study, the clinical relevance of Pax2 transactivation domain-interacting protein (PTIP/PAXIP1) was analyzed by immunohistochemistry of ESCC tissues. We found that low expression of PTIP was associated with lymph node metastasis in ESCC, and loss-of-function approaches showed that depletion of PTIP promoted ESCC cell migration and invasion both *in vitro* and *in vivo*. Analysis integrating RNA-seq and ChIP-seq data revealed that PTIP directly regulated ephrin type-A receptor 2 (EphA2) expression in ESCC cells. Moreover, PTIP inhibited EphA2 expression by competing with Fosl2, which attenuated the invasion ability of ESCC cells. These results collectively suggest that PTIP regulates ESCC invasion through modulation of EphA2 expression and hence presents a potential therapeutic target for its treatment.

Keywords: esophageal squamous cell carcinoma, PTIP, EphA2, Fosl2, invasion

INTRODUCTION

Esophageal cancer, a highly aggressive malignancy, globally ranks seventh in the incidence of cancer cases and sixth as the leading cause of cancer-related deaths. Moreover, in 2018, the number of new esophageal cancer cases and deaths worldwide was about 572000 and 509000, respectively (1). According to histopathological analysis, esophageal carcinoma can be classified as two main types: esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EAC) (2). ESCC is

the predominant subtype of esophageal cancer in China, accounting for 90% of esophageal carcinomas (3). Approximately 50% of patients when diagnosed with ESCC already have distant metastases and more than one-third develop distant metastases following surgery or radiotherapy (4, 5). In addition, most distant metastases of ESCC involve the distant lymph nodes (5). Therefore, dissecting the mechanisms underlying esophageal cancer invasion and metastasis is fundamental for the development of effective therapeutic strategies to improve patients' outcomes.

Tumor invasion and metastasis are complex processes based on angiogenesis and the weakening of tumor cell adhesion, that involve multiple stages, genes, and the accumulation of different factors (6). Receptor tyrosine kinases (RTKs) function as key regulators of signal transduction pathways that control cell proliferation, survival and migration in the progression of malignant solid tumors. EphA2 belongs to the Eph family, the largest family of membrane-bound receptor tyrosine kinases (7). Previous works show that EphA2 has dual roles in both promoting and inhibiting cancer cell metastatic progression (8, 9). In breast cancer, ligand-dependent EphA2 signaling inhibits proliferation and invasiveness, whereas ligand-independent manner promotes tumor malignancy through EphA2 phosphorylation at serine-897 (9, 10). Recent study has shown that EphA2 is highly expressed and associated with poor degree of tumor differentiation and lymph node metastasis in ESCC (11). However, how EphA2 is modulated in ESCC remains unclear.

PTIP (Pax2 transactivation domain-interacting protein) is a nuclear protein containing six BRCT domains and is an essential component of histone H3K4 methyltransferase complexes that are associated with gene activation (12), DNA repair (13, 14), embryonic vascular development (15), and embryonic stem cell pluripotency (16). In addition, PTIP can interact with the Pax family of transcription regulators and inhibit the transactivation of the glucagon promoter in pancreatic cells (17). Recent studies have shown that low PTIP is associate with more aggressive tumor phenotypes in breast cancer (18). Results from a meta-analysis analysis also found that downregulation of PTIP was associated with poor prognosis in ovarian cancer (19). These data suggest that PTIP can inhibit tumor progression.

In this study, we examined the role of PTIP in ESCC by analyzing the expression characteristics of PTIP in esophageal tumor tissues. The biological function and underlying molecular mechanisms of PTIP in ESCC invasion were also investigated. Here we found that low expression of PTIP was positively associated with ESCC tissue lymph node metastasis. Moreover, we demonstrated that PTIP participates in ESCC invasion and metastasis *via* suppressing the expression of EphA2, a crucial factor involved in tumor cell adhesion.

MATERIALS AND METHODS

Patient Information and Tissue Samples

Eighty-seven patients with ESCC underwent surgery with curative intent at the Affiliated Huaian No.1 People's Hospital of Nanjing Medical University (Huai'an, China). Those with confirmed or suspected lymph node metastasis received regional lymph node dissection. All patients provided written informed consent. The study protocol was approved by the Huaian No.1 People's Hospitals' Ethics Committee (No. YX-2020-162-01). The resected specimens were fixed in 10% formaldehyde solution and embedded in paraffin. The tumor stage was classified according to the 5th edition of the TNM classification of the International Union against Cancer (UICC).

Cell Culture

HEK293T cells and human ESCC cell lines, including TE1 and KYSE-150 were obtained from the Chinese Type Culture Collection (Shanghai, China) and cultured in Dulbecco's modified Eagle's medium (DMEM, Gibico, C11995500BT, Beijing, China) supplemented with 10% fetal bovine serum (FBS, AusgeneX, FBSSA500-S, QLD, Australia) and 1% penicillin/streptomycin (Gibico, 15140-122, USA) at 37°C in a humidified incubator with 5% CO_2. All cells were tested negative for mycoplasma.

Mouse Experiments

Female BALB/c nude mice (4–5 weeks old) were purchased from Nanjing Medical University and housed in a specific-pathogen-free barrier facility with free access to food and water. All animal experiments were approved by the Animal Experimentation Ethics Committee of Huai'an First People's Hospital. KYSE-150 cells (1×10^6) were infected with scrambled shRNA (shCtrl) or shPTIP#1 lentivirus containing a constitutively expressed luciferase reporter for 72 hours. Each group of cells were injected into the lateral tail vein of nude mice (n=8). The transplanted animals were monitored *in vivo* every week by bioluminescent imaging. Anesthetized mice were injected intraperitoneally with D-Luciferin (150 mg/kg, D1007, US Everbright, China) and were imaged with 2 min acquisition time using an In-Vivo FX PRO (BRUKER, NY, USA) imaging system, 10 min after injection. The bioluminescence intensity of the captured images were quantitated using Bruker MI SE acquisition and analysis software (BRUKER). Four weeks later, mice were euthanized, and lung tissue sections were stained with Hematoxylin-Eosin (H&E) for histopathological analysis.

Vector Construction

Lentiviral vectors harboring short hairpin RNA (shRNA) targeting PTIP (shPTIP), Fosl2 (shFosl2) and EphA2 (shEphA2) were synthesized by Genscript (Nanjing, China). cDNAs encoding PTIP were cloned into the expression vector pLenti-EFs-BSD with a blasticidin-resistant gene for N-terminal tagging of MYC epitope (PTIP-MYC) or FLAG epitope (PTIP-FLAG). Lentiviral luciferase-expressing vectors pLVshRNA-Luci (2A)-puro-shScrambled and pLVshRNA-Luci(2A)-puro-shPTIP were constructed by cloning shScrambled and shPTIP into pLVshRNA-Luci(2A)-puro vector (Inovogen, Beijing, China), respectively. The primer sequences used for plasmid construction are listed in **Table 1**.

TABLE 1 | Primers used for plasmid construction.

Primer	Sequence (5′-3′)
shPTIP-1-forward	GATCGGGGCAGGAAGACAGATATAATAACGAATTATTATATCTGTCTTCCTGCTTTTTC
shPTIP-1-reverse	AATTGAAAAAGCAGGAAGACAGATATAATAATTCGTTATTATATCTGTCTTCCTGCCCC
shPTIP-2- forward	GATCGGGGCAGCAACACAGTCCTCATCTCGAAAGATGAGGACTGTGTTGCTGCTTTTTC
shPTIP-2- reverse	AATTGAAAAAGCAGCAACACAGTCCTCATCTTTCGAGATGAGGACTGTGTTGCTGCCCC
shFosl2-1- forward	GATCGGGGGATTATCCCGGGAACTTTGACGAATCAAAGTTCCCGGGATAATCCTTTTTC
shFosl2-1- reverse	AATTGAAAAAGGATTATCCCGGGAACTTTGATTCGTCAAAGTTCCCGGGATAATCCCCC
shFosl2-2- forward	GATCGGGGATCATGTACCAGGATTATCCCGAAGGATAATCCTGGTACATGATCTTTTTC
shFosl2-2- reverse	AATTGAAAAAGATCATGTACCAGGATTATCCTTCGGGATAATCCTGGTACATGATCCCC
PTIP- forward	GGATCTATTTCCGGTGAATTCCATGTCGGACCAGGCGC
PTIP- reverse	GGAGGGAGAGGGGCGGGATCCCCAGATCCTCTTCTGAGATGAGTTTCTG
shScrambled- forward	GATCGGGTTCTCCGAACGTGTCACGTTTCCGAAGAAACGTGACACGTTCGGAGAATTTTTC
shScrambled- reverse	AATTGAAAAATTCTCCGAACGTGTCACGTTTCTTCGGAAACGTGACACGTTCGGAGAACCC
shEphA2-1-forward	GATCGGGGCATCTTCCTAGTGCCCTACTCGAAAGTAGGGCACTAGGAAGATGCTTTTTC
shEphA2-1- reverse	AATTGAAAAAGCATCTTCCTAGTGCCCTACTTTCGAGTAGGGCACTAGGAAGATGCCCC
shEphA2-2-forward	GATCGGGGCTCCTCTTTATACCTCTAGACGAATCTAGAGGTATAAAGAGGAGCTTTTTC
shEphA2-2- reverse	AATTGAAAAAGCTCCTCTTTATACCTCTAGATTCGTCTAGAGGTATAAAGAGGAGCCCC

Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR)

Total RNA in ESCC cells was extracted using RNeasy Kits (Qiagen, 74104, Hilden, Germany) according to the manufacturer's instructions. Primescript RT-reagent kit (Takara, RR047A, Shiga, Japan) was used to synthesize cDNA. SYBR Premix Ex Taq (Takara, Dalian, China) was used for qPCR. qPCR analysis was performed using LightCycler 480 system (Roche, Basel, Switzerland) and the data were analyzed by the ΔΔCt method with normalization to β-actin. The primer sequences used for real-time PCR are listed in **Table 2**.

Western Blot

Proteins were isolated from human ESCC cells using lysis buffer (2% SDS, 62.5 mM Tris pH 6.8, 10% glycerol, 5% β-mercaptoethanol) and supplemented with an EDTA-free protease inhibitor cocktail (Roche, 4693132001, MO, USA), phosphatase inhibitor cocktail (MCE, HY-K0021, NJ, USA), and PMSF (MCE, HY-B0496, NJ, USA). SDS-PAGE was used to separate the proteins with different molecular weights (45 min, 200 V) (Bio-Rad Laboratories) and then transferred onto PVDF membranes (LC2002, Invitrogen, CA, USA) using a wet blotting system (70 min, 250 mA) (BioRad Laboratories). Membranes were blocked with 5% non-fat milk in TBST and then incubated with primary antibodies, followed by the appropriate secondary antibody.

The antibodies used in this study are listed as follows: PTIP (1:2000, 130kDa, A300-370A, BETHYL, USA), EphA2 (1:1000, 108 kDa, AF5238, Affinity Biosciences, USA), pS897-EphA2 (1:1000, 125 kDa, 6347S, CST, USA), pY588-EphA2(1:1000, 125 kDa, 12677S, CST, USA), YY1 (1:1000, 70kDa, 22156-1-AP, Proteintech, USA), MYC (1:1000, 16286-1-AP, Proteintech, IL, USA), FLAG (1μg/μL, T0003, Proteintech, IL, USA), Fosl2 (1:1000, 35 kDa, A2729, ABclonal, USA), GAPDH (1:2000, 36 kDa, AP0063, Bioworld, Nanjing, China), β-actin (1:2000, 42 kDa, 20536-1-AP, Proteintech, IL, USA).

Proteins were detected using a chemiluminescence reagent (WesternBright peroxide, Advansta, CA, USA) in a ChemiDoc XRS+ System (Bio-Rad, CA, USA).

TABLE 2 | Primers used for Real-time PCR.

Primer	Sequence (5′-3′)
PTIP-RT-forward	CCAGCTGTACGGACACTGAGG
PTIP-RT-reverse	TTGTATGTCCCTGCTGGCTGT
EphA2- RT-forward	TGGCTCACACACCCGTATG
EphA2- RT- reverse	GTCGCCAGACATCACGTTG
TACSTD2- RT-forward	CGGCAGAACACGTCTCAGAAG
TACSTD2- RT-reverse	CCTTGATGTCCCTCTCGAAGTAG
GPRC5A - RT- forward	ATGGCTACAACAGTCCCTGAT
GPRC5A - RT- reverse	CCACCGTTTCTAGGACGATGC
β-actin- RT-forward	AAGACCTGTACACCAACACAG
β-actin- RT-reverse	AGGGCAGTGATCTCCTTCT
BML-RT-forward	CAGACTCCGAAGGAAGTTGTATG
BML-RT-reverse	TTTGGGGTGGTGTAACAAATGAT
P53-RT-forward	CAGCACATGACGGAGGTTGT
P53-RT-reverse	TCATCCAAATACTCCACACGC

Immunohistochemistry

Paraffin-embedded blocks from human specimens or mouse xenografts were sectioned into 4-μm thick slices and positioned on pre-coated slides. The unstained slides were deparaffinized, rehydrated, and incubated with fresh 0.3% H_2O_2 in methanol for 30 min at about 25°C. Thereafter, antigen retrieval was performed in 10 mM citrate buffer by heating the samples to 121°C. Tissue sections were then blocked for 30 min with 5% horse serum in PBST and incubated with the primary antibody at 4°C overnight, washed in PBS, and incubated with secondary antibody for 2 h at room temperature. Immunoreactivity was visualized by DAB chromogen followed by hematoxylin counterstain. Two pathologists scored the staining of PTIP and EphA2 independently, then a multiplicative Quick-score (Q-score) was calculated by multiplying the percentage of positive cells by the intensity of the staining. Ten sections were collected randomly, and the average Q-score was calculated for each section.

RNA-Seq Analysis

Transcriptome analyses were performed to compare the control TE1 with PTIP-depleted cells. This same methodology was applied to compare non-invasive with invasive tumor biopsies from ESCC patients. cDNA libraries were generated using NEB Next Ultra Directional RNA Library Prep Kit (NEB) following the manufacturer's instructions, and then sequenced using the Illumina sequencing technology on an Illumina Novaseq 6000 at LC Bio (Zhejiang, China) according to the manufacturer's instructions. Differential expression analysis of the two groups was performed using the ballgown R package (20). The resulting *P*-values were adjusted using Benjamini and Hochberg's approaches to control the false discovery rate. Genes with an adjusted *P*-value <0.05 were considered to be differentially expressed. The clusterProfiler R package was used to detect the statistical enrichment of differentially expressed genes in Gene Ontology (GO).

Chromatin Immunoprecipitation Sequencing (ChIP-Seq) Analysis

ChIP-seq experiments were performed on TE1 cells overexpressed with MYC-tagged PTIP. For histone ChIP–seq, cells were fixed in 1% formaldehyde (FA, Thermo Fisher Scientific, 28906) for 10 min at room temperature. The reaction was stopped by the addition of glycine (0.125 M, Sigma), and the cells were washed in ice-cold PBS. Cells were sonicated using a Bioruptor sonicator (Diagenode) until 200–500 bp DNA fragments were obtained. Further procedures were performed according to Richard A Young's protocol (21). DNA libraries were sequenced on an Illumina HiSeq platform (Novogene, Beijing, China). The sequencing data were mapped to the hg38 genome, and peak calling was performed using Model-based analysis of ChIP-Seq (MACS) version 2.1.1 with default parameters to obtain primary binding regions. The pie plot and heatmap of ChIP binding the TSS regions were generated using the ChIPseeker R package (22). Motif discovery was performed using HOMER. ChIP-Seq datasets were subsequently visualized using the IGV software.

Chromatin Immunoprecipitation Quantitative Real-Time PCR (ChIP-qPCR)

ChIP assays were performed with a Millipore ChIP kit according to the manufacturer's protocol (21). Briefly, 3×10^7 cells were collected, fixed, and sonicated with a Bioruptor sonicator (Diagenode) to generate DNA fragments of approximately 500 bp in length. Chromatin immunoprecipitates for proteins were amplified by quantitative PCR, normalized to input, and calculated as percentages of inputs. Fold enrichment levels indicate fold changes over the negative control immunoglobulin G (IgG). The PCR primer sequences are listed in **Table 3**.

TABLE 3 | Primers used for ChIP-qPCR.

Primer	Sequence (5′-3′)
EphA2-F forward	GCCCTTATCGTGACGCAAGT
EphA2-F reverse	CCCTAGGTGAATTGCCACCA
EphA2-NC-forward	CAGCAGGCAGTGGGATGAG
EphA2-NC- reverse	TCCCACAGCTAGGAGGTGACA

Transwell Assay

Cells (4×10^4/well) resuspended in serum-free DMEM were added to the upper transwell chambers coated with matrigel. DMEM with 10% FBS was added to the lower chambers. After 48 h of incubation at 37°C, cells remaining in the upper chamber were wiped with cotton buds, whereas cells that had invaded the lower chamber were fixed with 10% methanol and stained with 0.1% crystal violet. Images were captured under a microscope. Each transwell assay was conducted in duplicate and repeated 3 times. The rate of invasion was measured using an ImageJ analysis system (version 1.8.0; National Institutes of Health, Bethesda, MD, USA).

Wound Healing Assay

Cells (6×10^5/well) were seeded into a 6-well culture plate. The next day, when the cells were almost 80% confluent, a 10 μL pipette tip was used to scratch artificial linear wounds in the monolayer. Cells were washed 3 times with PBS and cultured in serum-free DMEM at 37°C. The width of the scratch gap was monitored by microscopy and photographed at 0 h and 12 h. The cell migration rate was quantified according to the original width of the wound and the width after cell migration. The ImageJ analysis system was used to analyze the rate.

Apoptosis Assay

Analyses were performed using YF647A-AnnexinV/PI apoptosis detection kit (Y6026, US EVERBRIGHT, Suzhou, China) according to the manufacturer's protocol. Briefly, ESCC cells were seeded into 6-well plates with 5×10^5. After the cell density reached 80∼90%, cells were then harvested and washed with PBS for three times before addition of 500 μL binding buffer, 5 μL YF647A-AnnexinV and 5 μL PI solution to the cell pellet for 15 min at room temperature in the dark before flow cytometric analysis to detect early apoptotic cells (Annexin-V positive and PI-negative) and late apoptotic cells (Annexin-V and PI double positive)(Beckman Coulter, USA).

Cell-Matrix Adhesion Assay

ESCC cells were seeded at a density of 2×10^5 per ml (100 μl per well) in 96-well plates coated with Collagen I (40 μg/ml; Corning, NY, USA). Following 1 h incubation at 37°C in an incubator containing 5% CO2, the cells were washed with PBS to remove non-adherent cells. A total of 10 μl CCK8 (Dojindo, CK04, Japan) was added to each well. Following 1 h of additional incubation, absorbance values were determined using Hidex Sense instrument (Hidex, Finland) at a wavelength of 450 nm. The percentage of adhesive cells was calculated according to the following formula: Percentage of adhesion= [(optical density (OD) 450 of PTIP knockdown cells-Blank)/(OD450 of Ctrl cells-Blank)]x100%. Three independent experiments were performed in triplicate.

Immunofluorescence by Confocal Microscopy

ESCC cells were cultured on glass coverslips, fixed with 4% paraformaldehyde in PBS for 20 min, and subsequently

permeabilized with 0.5% Triton X-100 for 10 min. Cell were incubated with phalloidin (YP0063, US EVERBRIGHT, Suzhou, China) for 20 min at room temperature. After three washes with PBS, cells were mounted with Aqueous Mounting Medium containing 4',6-diamidino-2-phenylindole (DAPI, Beyotime, C1005, Shanghai, China) and visualized with a confocal microscope.

Statistical Analysis

SPSS software version 19.0 (IBM Corp., Armonk, New York, US) and GraphPad Prism v6 (GraphPad Software, Inc., San Diego, California, US) were used for all statistical analyses. Data were first evaluated for normal distribution using the Shapiro-Wilk method and homogeneity of variance with the Levene method. Pairwise comparisons of normally distributed data were analyzed using Student's t test or for multigroup comparisons, one-way analysis of variance (ANOVA) with *post hoc* Tukey's test. Data not meeting normal distribution/homogeneity of variance were compared using Kruskal-Wallis and Mann-Whitney non-parametric tests. Data are presented as the mean ± SEM (error bars). $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***) denote statistically significant changes.

RESULTS

PTIP Expression Negatively Correlates With Lymph Node Metastasis in ESCC

To identify metastatic genes during ESCC progression, we used an RNA-seq-based approach to compare the transcription profile differences between non-invasive and invasive tumor biopsies from ESCC patients. Overall, 776 genes were differentially expressed. Of these, 435 genes were upregulated in invasive tumor samples while 341 genes were downregulated (**Figure 1A**). We conducted functional enrichment analysis using online tool Metascape and obtained significant enrichment CORUM (the comprehensive resource of mammalian protein complexes) gene set and Gene Ontology (GO) terms (three terms: Biological Process: BP, Cellular Component: CC and Molecular Function: MF). We had 2 significant CORUM gene set, 330 significant GO: BP terms, 28 significant GO : CC terms and 58 significant GO : MF terms. The top CORUM gene set from the functional enrichment analysis of DEGs is "PTIP-DNA damage response complex" (**Figures 1B, C**). Of this set of genes, three were downregulated in patients with lymph node metastasis, including BLM, TP53 and PTIP. These genes were further investigated by RT-qPCR. The results shown that the most significantly changed gene was PTIP (**Figure 1D**).

We next analyzed the correlation between PTIP expression and the clinicopathological parameters in tissues from 79 ESCC patients. The expression of PTIP varies greatly in ESCC tissues (**Figure 1E**). Based on the PTIP immunostaining intensity scores mentioned in the Methods section, the cohort of 79 ESCC patients were divided into low PTIP and high PTIP groups. We found a statistically significant negative correlation between PTIP nuclear expression in ESCC and lymph node status ($P = 0.0399$; **Tables 4**, **S1** and **Figure 1F**). PTIP staining intensity was higher in lymph node metastasis–negative samples when compared with metastasis-positive samples. When taken in context with our PTIP RNAi sequencing data, our findings indicate that PTIP was reduced in more aggressive ESCC phenotypes.

Knockdown of PTIP Promotes ESCC Cell Invasion and Migration

To investigate the function of PTIP in ESCC, we silenced PTIP in TE1 and KYSE-150 cell lines using shRNA lentivirus. PTIP knockdown substantially promoted the invasion abilities of TE1 and KYSE-150 cells (**Figures 2A–F**). Furthermore, the wound healing assay demonstrated that suppression of PTIP resulted in a higher scratch closure rate compared with control groups (**Figures 2G–J**). To exclude the effect of cell apoptosis on cell migration, we conducted an apoptosis assay in ESCC cells using Annexin V and PI staining. Flow cytometry revealed that there were no remarkable differences in apoptosis rate between shPTIP and shCtrl cells (**Figures S1a, b**). Moreover, the EC-matrix adhesion ability and cell morphology of ESCC cell lines were determined by cell-matrix adhesion assay and phalloidin staining, respectively. No significant differences were observed between shPTIP and shCtrl cells (**Figures S1c, d**). Furthermore, we checked the effect of PTIP overexpression in ESCC cells, observing that transformed cells did not significantly inhibit cell invasion (**Figures S2a-f**). This may be duo to endogenously-expressed basal level of PTIP is high enough to suppress ESCC cell invasion. Collectively, these findings indicate that PTIP is important for the invasion and migration of ESCC cells.

Further, to determine the role of PTIP in ESCC invasion *in vivo*, we treated KYSE-150 cells with shScrambled-Luc (control) or shPTIP#1-Luc-containing viruses, and subsequently delivered these 2 groups of cells into mice by tail vein injection. *In vivo* imaging analysis demonstrated that mice injected with shPTIP-Luc-treated KYSE-150 cells had stronger luciferase signals than those injected with shScrambled-Luc treated KYSE-150 cells (**Figure 3A**). In particular, there was a significant difference in the luciferase signal between shScrambled-Luc group and shPTIP group at 4th week (**Figure 3B**). Metastatic tumors in bioluminescence-positive tissues were further confirmed by H&E staining (**Figure 3C**). These results suggest that depletion of PTIP promotes invasiveness of ESCC cells *in vivo*.

PTIP Attenuates EphA2 Gene Expression in ESCC Cells

A comparative analysis of the transcriptomes between the two groups of ESCC cells (shScrambled vs. shPTIP) was performed to understand the mechanism by which PTIP attenuates ESCC invasiveness. A total of 6005 differentially expressed genes (DEGs) were identified, of which 3076 were upregulated and 2929 downregulated (**Figure 4A**). GO enrichment analysis was performed for the DEGs to investigate the biological functions of these genes. The top 20 GO biological processes based on *P*-value are listed in particular (**Figure 4B**), the cell-cell adhesion processes were significantly changed after depletion of PTIP compared to control cells (**Table S2**).

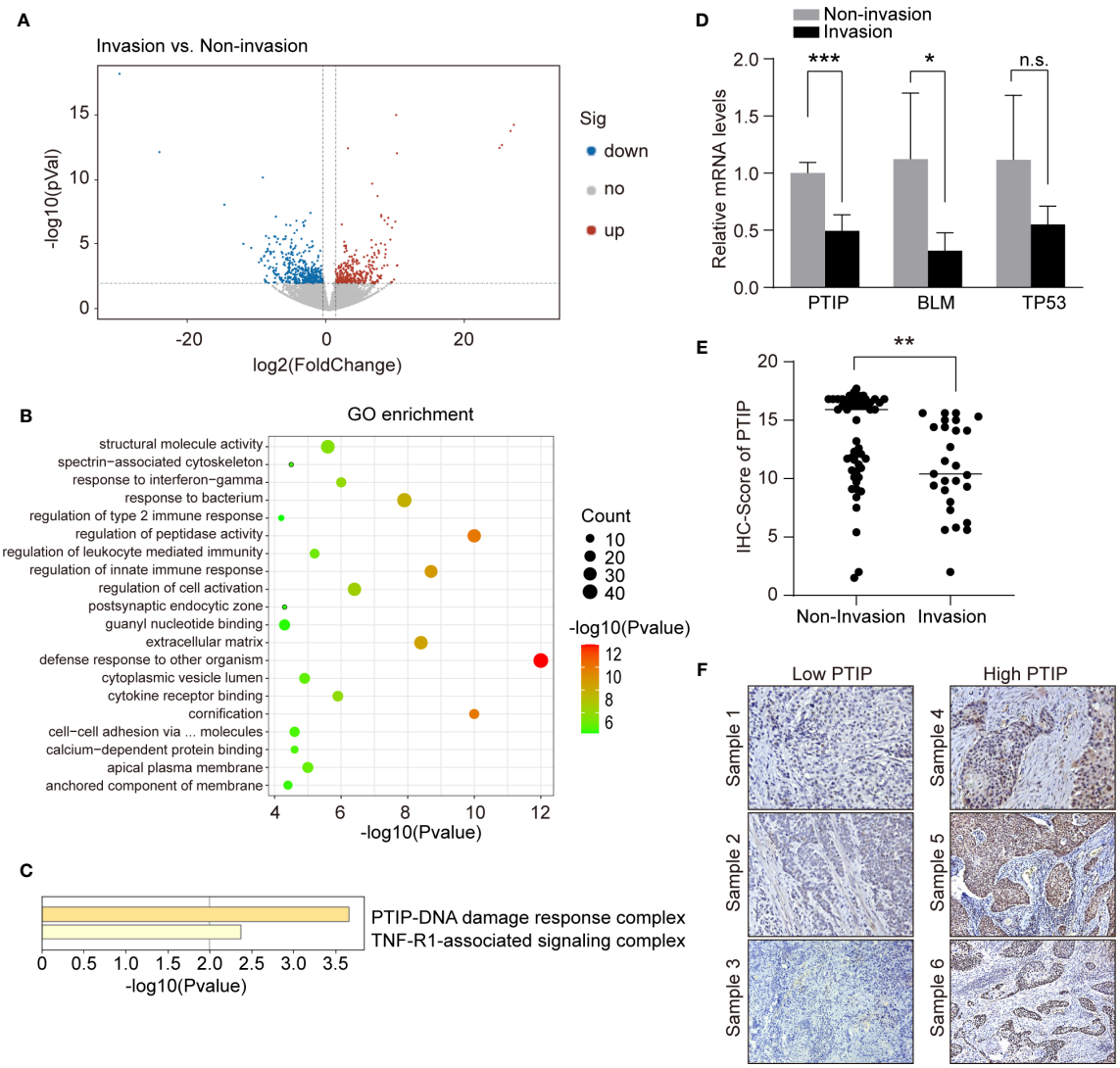


FIGURE 1 | PTIP expression negatively correlates with lymph node metastasis in ESCC. **(A)** Volcano plot showing differentially expressed genes between tumor biopsies with and without lymph node metastasis using RNA-seq. Volcano plot shows the log2 fold change (x-axis) and significance (–log10*adjusted *P*-value; y-axis) with significantly downregulated and upregulated genes shown in blue and red, respectively (adjust $P < 0.01$ and |log2Foldchange| > 0.9). **(B)** GO enrichment of differentially expressed genes from **(A)**. **(C)** Functional enrichment analysis of differentially expressed genes from **(B)** using CORUM gene set. **(D)** quantitative RT-PCR analysis of genes that downregulated in lymph node metastasis positive ESCC samples. Unpaired, two-tailed Student's t-test; $^{*}P < 0.05$; $^{***}P < 0.001$; n.s., not significant. **(E)** IHC score for PTIP in ESCC tumor sections. Unpaired, two-tailed Student's t-test; $^{**}P < 0.01$. **(F)** Representative IHC images for PTIP in ESCC tumor sections. Low PTIP and high PTIP groups were divided based on the PTIP immunostaining intensity scores mentioned in method. Cut off for high and low PTIP expression in ESCC was defined < or > 11. Percentage of invasion and non-invasion in ESCC groups is depicted in **Table 4**.

To further investigate the direct target of PTIP in ESCC cells, we used ChIP-Seq with a Myc tag antibody in ESCC cells overexpressing Myc-tagged PTIP (**Figure 4C**). We documented 8439 PTIP-binding sites, corresponding to 126 genes. Further analysis revealed that more than 57% of the binding sites were found at the promoter and transcription start site (TSS) regions of annotated genes (**Figure 4D**), with the peak regions located at −3,000 to +3,000 bp from the TSS (**Figure 4E**), supporting the hypothesis that PTIP may function as a transcriptional co-factor. Among the PTIP target genes, genes assigned to the terms "Transcription, DNA-templated", "Positive regulation of transcription from RNA polymerase II promoter", "Cell proliferation", and "Cell migration" in the BP category were highly enriched (**Figure 4F**). Intersection of the ChIP-Seq target genes with the above RNA-Seq DEGs revealed 38 overlapping genes, including 3 cell-cell adhesion genes (**Figures 4G, H**). These 3 genes were further analyzed by RT-qPCR (**Figure 4I**). The results showed that EphA2 is directly regulated by PTIP.

TABLE 4 | Summary of clinical and histopathological characteristics of the 79 esophageal squamous cell carcinoma patients.

Variables	All cases (N=79; %)	PTIP		*P*-values
		Low (n = 25; %)	High (n = 54; %)	
Gender				0.5995/0.4599
Male	55(69.6)	16(20.3)	39(49.4)	
Female	24(30.4)	9(11.4)	15(19.0)	
Age(year)				0.3323/0.2452
<65	36(45.6)	9(11.4)	27(34.2)	
>=65	43(54.4)	16(20.3)	27(34.2)	
Size(cm)				0.3845/0.3402
<5	62(78.5)	18(22.8)	44(55.7)	
>=5	17(21.5)	7(8.9)	10(12.7)	
Grade				0.973
Poor	27(34.2)	7(8.9)	20(25.3)	
Moderate	34(43.0)	14(17.7)	20(25.3)	
Well	18(22.8)	4(5.1)	14(17.7)	
TNM Stage				0.4338/0.4015
I-II	24(30.4)	6(7.6)	18(22.8)	
III-IV	55(69.6)	19(24.1)	36(45.5)	
Lymph node status				0.0399/0.0231*
0	52(65.8)	12(15.2)	40(50.6)	
>=1	27(34.2)	13(16.5)	14(17.7)	
Distant metastasis				
M0	79(100)	25(31.6)	54(68.4)	
M1	0(0)	0(0)	0(0)	

The numbers in parentheses indicate the percentages of tumors with a special clinical or pathologic feature for a given PTIP subtype.
The differences between rates were tested by $\chi 2$ or Fisher exact tests, if appropriate.
**Statistically significant.*

PTIP Inhibits ESCC Cell Invasion and Migration Through Suppression of EphA2

Since PTIP was shown to directly regulate EphA2 expression in ESCC cells, we hypothesized that PTIP inhibits ESCC cell invasion and migration through transcriptional regulation of EphA2. Western blotting was performed to evaluate the expression of EphA2 in PTIP knockdown cells and further understand how PTIP regulates EphA2 expression (**Figure 5A**). Next, we assessed the correlation of EphA2 expression level and lymph node status in ESCC patients. Compare to non-invasive ESCC patients, EphA2 is highly expression in invasive patients (**Figures 5B–D**). In addition, pearson correlation coefficient indicates significant negative correlation between PTIP and EphA2 expression levels (**Figure 5E**). Rescue experiments show that knockdown of EphA2 could attenuate PTIP depletion promoted ESCC invasion (**Figures 5F–H**). These results showed that EphA2 was negatively correlated with the expression of PTIP, which is consistent with the sequencing results. Depletion of PTIP by shRNAs increased S897phosphorylation and decreased Y588 phosphorylation of EphA2 (**Figure 5A**), indicating that EphA2 enhance ESCC cell invasion through ligand-independent manner. In agreement with previous reports that EFNA1 promotes rapid turnover of phosphorylated EphA2. Overexpression of EFNA1 reduced the protein level of EphA2 in ESCC cells and inhibited the invasion of KYSE-150 cells. Although not significant, there was a similar trend in TE1 cells (**Figures S3a-f**).

We next thoroughly analyzed the above ChIP-seq data and observed that a specific peak of PTIP overlapped in the intragenic regions of the EphA2 gene (**Figure 5I**). This result was verified by ChIP-qPCR using specific primers targeting the intron region of EphA2 (**Figures 5J, K**). Taken together, these data demonstrate that EphA2 is one of the downstream target genes of PTIP, indicating that PTIP plays a substantial role in ESCC through EphA2.

PTIP Inhibits EphA2 Expression by Competing With Fosl2 for Binding to EphA2

To clarify the mechanism through which PTIP inhibits transcription of EphA2, we identified a series of PTIP highly enriched binding motifs using HOMER. The top 2 motifs of PTIP association were closely similar to the Fosl2(AP-1) and YY1 binding motifs (**Figure 6A**). Loss of function analysis shown that knockdown of Fosl2, but not YY1, increased EphA2 expression (**Figures 6B, C** and **Figure S4**). Next, cell invasion capability was robustly inhibited by depletion of Fosl2 in ESCC cells (**Figures 6D–G**). Indeed, ectopic expression of EphA2 partially rescued the hypo-invasive phenotype caused by knockdown of Fosl2 (**Figures 6H–M**).

Since PTIP and Fosl2 play opposing roles in regulating EphA2 and ESCC invasion, we hypothesized that PTIP and Fosl2 competitively occupy the cis-acting element of EphA2 to regulate EphA2 expression. To test this hypothesis, we next examined the relationship between Fosl2 and PTIP in the transcriptional regulation of EphA2. Indeed, when PTIP was knocked-down, the recruitment of Fosl2 to the EphA2 gene was dramatically increased compared to the control, and vice versa (**Figures 6N, O**). In a word, PTIP inhibited EphA2 expression by competing with Fosl2.

Our results provide evidence that PTIP competitively regulates EphA2 gene expression with Fosl2 to contribute to the invasiveness of ESCC cells. Our findings may offer potential therapeutic agents for the treatment of ESCC, such as small molecule modulators of PTIP or Fosl2.

DISCUSSION

A better understanding of the adhesion mechanisms responsible for tumor cell invasiveness is critical, as tumor cells with low adhesion may separate from each other. In this study, the mRNA profiles of lymph node metastasis–positive ESCC samples were compared to those without lymph node metastasis. Our results showed that PTIP was downregulated in ESCC samples positive for lymph node metastasis. Moreover, a significantly negative correlation was detected between PTIP expression and the lymph node metastasis status of ESCC patients. Western blotting and RNA-seq analyses showed that PTIP inhibited the expression of EphA2 in ESCC cells. More importantly, we found that PTIP knockdown could promote ESCC metastasis *in vivo* using a nude mouse xenograft model. In xenograft mouse tumor sections, the PTIP knockdown tumors showed increased lung metastasis compared with the control group. Taken together, these results indicate that PTIP inhibits lymph node metastasis in ESCC *in*

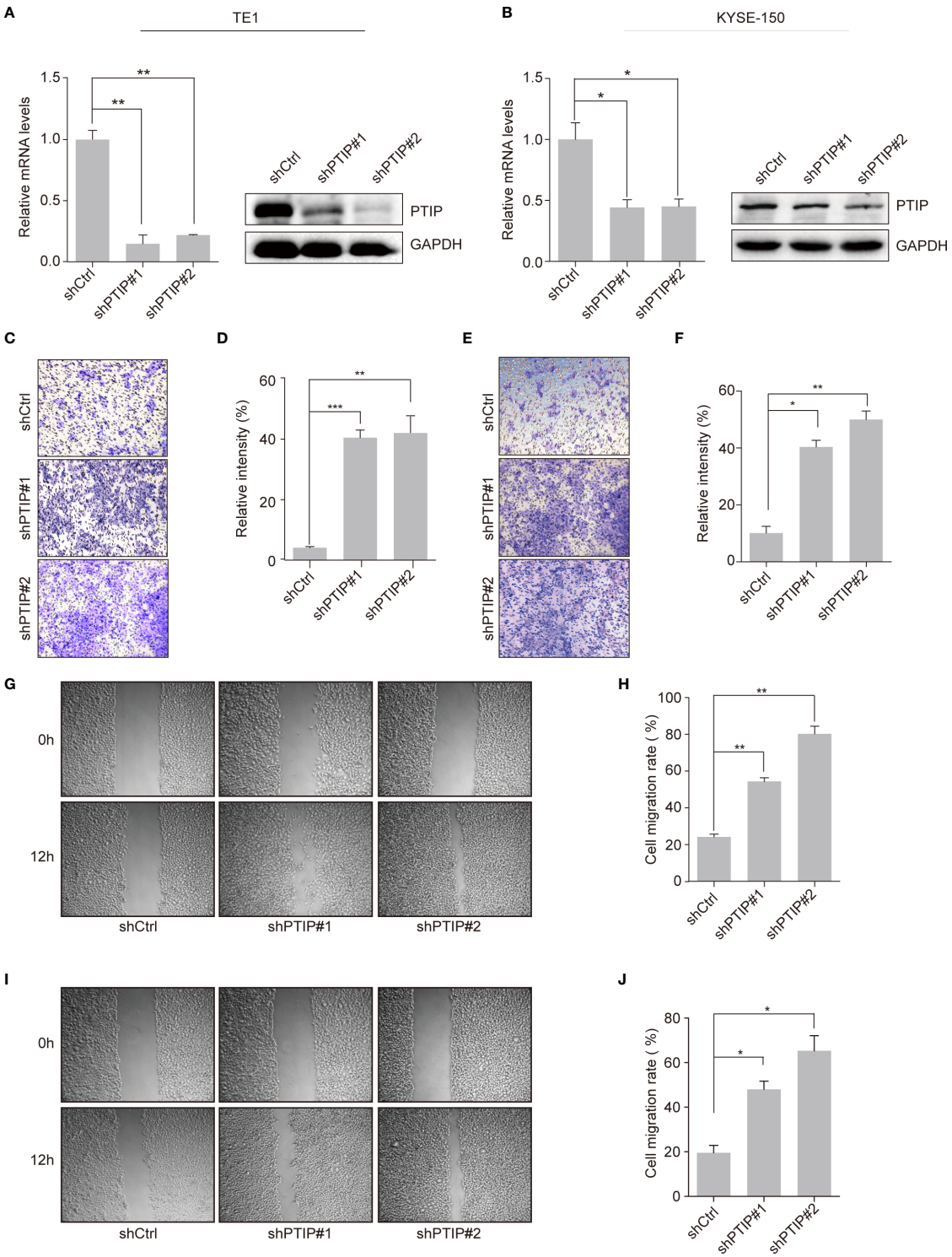


FIGURE 2 | Effects of PTIP on ESCC cell invasion and migration. **(A, B)** Knockdown efficiency of two independent shRNAs targeting PTIP in TE1 **(A)** and KYSE-150 **(B)** cells as determined by qRT-PCR and Western blot. **(C–F)** The effect of PTIP knockdown on the invasiveness of TE1 **(C, D)** and KYSE-150**(E, F)** cells. For invasion assay, six different microscopic fields (magnification, ×10) from at least three independent experiments were examined; Relative intensities of the fields were measured (n ≥3). Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. One-way ANOVA; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. **(G–J)** Wound healing assay was performed to determine the cell migration in TE1 **(G, H)** and KYSE-150 **(I, J)** cells. Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. One-way ANOVA; $^{*}P < 0.05$,$^{**}P < 0.01$.

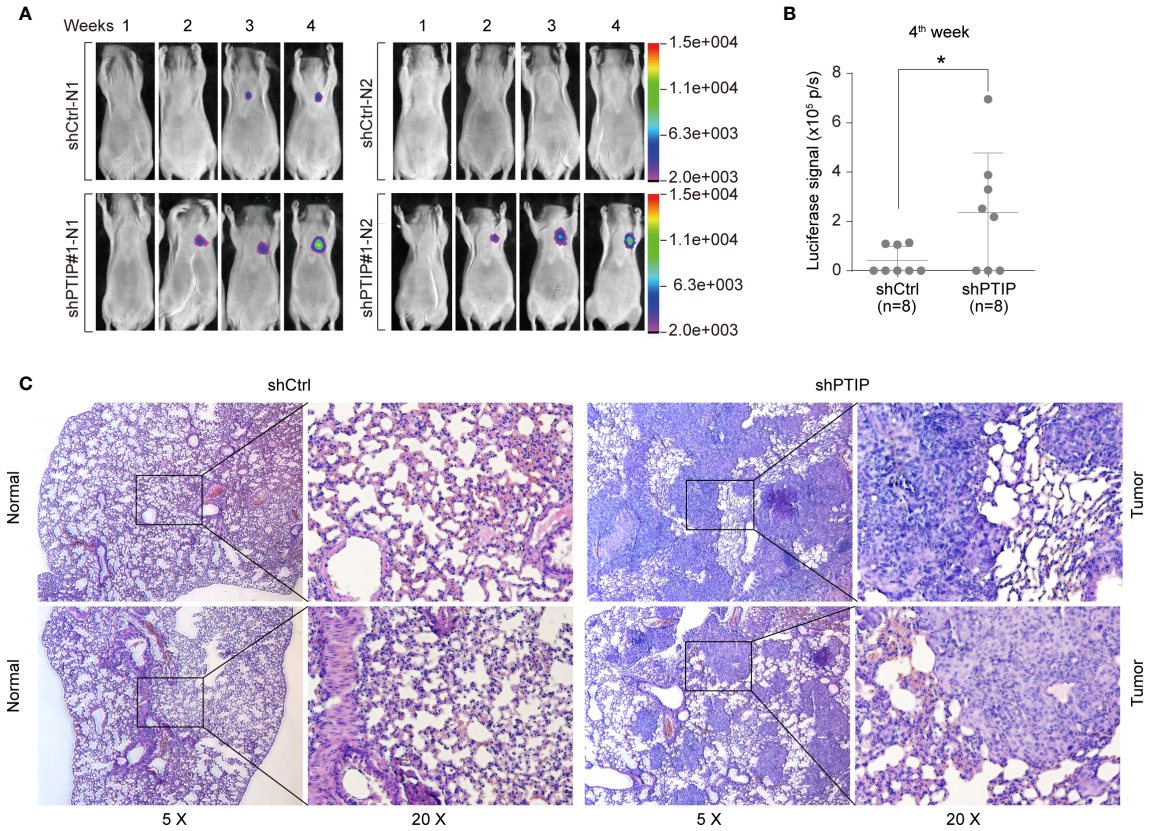


FIGURE 3 | PTIP inhibits ESCC cell invasion and migration *in vivo*. **(A, B)** The effect of PTIP knockdown on tumor growth and metastasis in a mouse model based on tail vein injection. KYSE-150 cells were infected with Scrambled shRNA (shCtrl) or shPTIP#1 lentivirus containing a constitutively expressed luciferase reporter. Then, shScrambled-treated or shPTIP cells were injected into mice *via* tail veins(n=8), and lung metastases in the two groups were evaluated by an *in vivo* imaging system(In-Vivo FX PRO). The luciferase signals were compared between the shCtrl group and the shPTIP group. Unpaired, two-tailed Student's t-test; $^{*}P < 0.05$. **(C)** Metastatic lesions in lungs from mice at the 4th week were analyzed by H&E staining. Overall staining patterns were shown at low magnification (original magnification ×5) as a composite figure and zoomed views (original magnification ×20) of the indicated areas (box).

vivo. Therefore, our findings reveal a novel PTIP function in which PTIP negatively regulates ESCC cell invasiveness by transcriptionally suppressing EphA2 gene expression.

During tumor progression, elevated levels of matrix metalloproteinases (MMPs) and decreased epithelial-mesenchymal transition (EMT) facilitate tumor cell invasion and metastasis (23, 24). It has been shown that mixed-lineage leukemia (MLL)-like complexes, including UTX, MLL4 and PTIP, play a significant role in cancer invasion and metastasis (25–27). In breast cancer, UTX (also known as KDM6A) and MLL4 enhance the invasion ability of breast cancer cells by promoting the expression of MMP family proteins (27). In our study, GO analysis showed that PTIP does not regulate the expression of MMP family proteins in ESCC cells, but regulates the expression of EMT-related genes. Further ChIP-seq results indicated that PTIP may indirectly regulate the expression of EMT-related genes.

In general, it is believed that PTIP acts as a transcription activator, through interaction with other MLL-like components such as UTX and MLL3 (16, 28, 29). Contrary to the prevailing perception that PTIP is solely a transcriptional activator, our study identified that it also functions as a repressor. These results are in agreement with findings by Fang et al. which showed that 52 genes were upregulated after PTIP knockdown in *Drosophila* Kc cells (29). In our study, we found that PTIP inhibited EphA2 expression by competing with Fosl2. Fosl2, a transcription factor of the activator protein-1 family, has been linked to cell adhesion, movement, invasion, metastasis, and cell growth (30). Overexpression of Fosl2 is associated with higher invasiveness in breast cancer (31). Indeed, our results showed that knockdown of Fosl2 significantly downregulated EphA2 expression and reduced the invasion ability of ESCC cells. Further investigation showed that PTIP and Fosl2 compete for binding to EphA2 cis-acting elements, thereby regulating gene expression and ultimately affecting the outcome of ESCC. It will be interesting to determine the relation between PTIP expression levels and overall survival in patients with ESCC.

Eph and ephrin are identified as contributors to tumor progression, hence they are considered to be attractive tumor markers and they are attractive targets for therapy. But so far, there are no drugs targeting Eph/ephrin family for medical use, because the interactions between Eph and ephrin are not specific and promiscuous (8). Therefore, targeting the upstream regulators of EphA2 may provide additional targets for cancers.

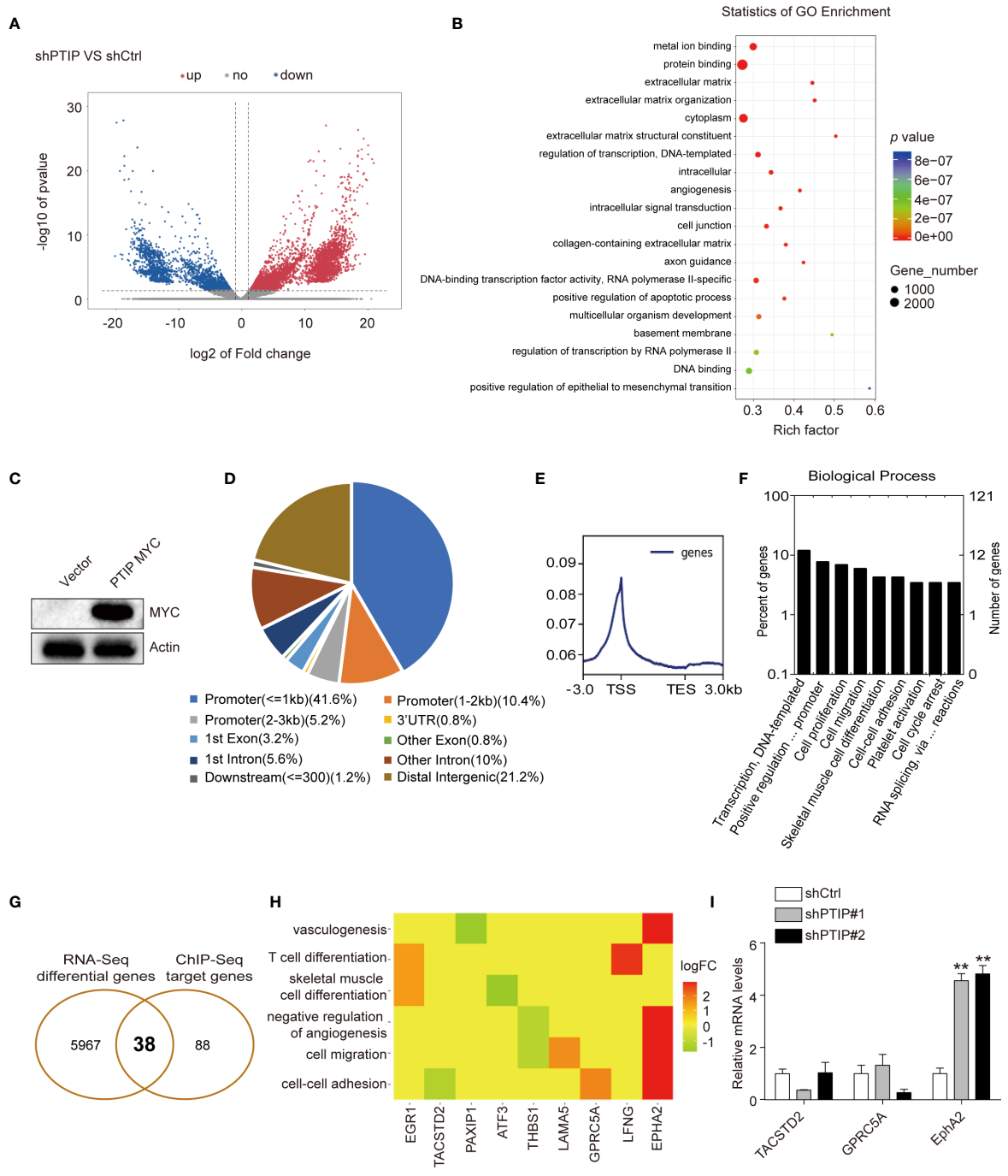


FIGURE 4 | Analysis of the differentially expressed genes(DEGs) and PTIP binding sites. **(A)** The differentially expressed genes between PTIP knockdown(shPTIP) and control (shScrambled) group were identified by RNA-seq analysis. Volcano plot shows the log2 fold change (x-axis) and significance (–log10 * adjusted p-value; y-axis) with significantly downregulated and upregulated genes shown in blue and red, respectively (adjust $P < 0.01$ and |log2Foldchange| < 1). **(B)** Gene ontology (GO) analysis for enrichment of the DEGs based on the results from **(A)** RNA sequencing. **(C)** Western blot verification of MYC tagged PTIP over-expression efficiency in TE1 cells. **(D–F)** PTIP binding sites and target genes were identificated by chromatin immunoprecipitation (ChIP)-Seq using MYC tag antibody in MYC tagged PTIP ESCC cell line. Pie charts showing the distribution of PTIP-binding sites in the genome. The percentage of binding sites are indicated in parentheses. The graphs were generated using the ChIPseeker, Deeptools and GOplot package in R. Binding profile around the transcription start site (TSS) of PTIP **(E)**. Mean read coverage is plotted (y-axis) against a sliding window around the TSS (x-axis). TES, transcription end site; TSS, transcription start site. Enrichment of biological process terms among PTIP target genes**(F)**. **(G)** Venn diagram showing the overlapping genes identified by RNA-Seq data and ChIP-Seq data. **(H)** Enrichment of biological process terms among the 38 overlapping genes. **(I)** qRT-PCR analysis of three potentially PTIP directly regulated cell-cell adhesion genes. Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. One-way ANOVA; $^{**}P < 0.01$.

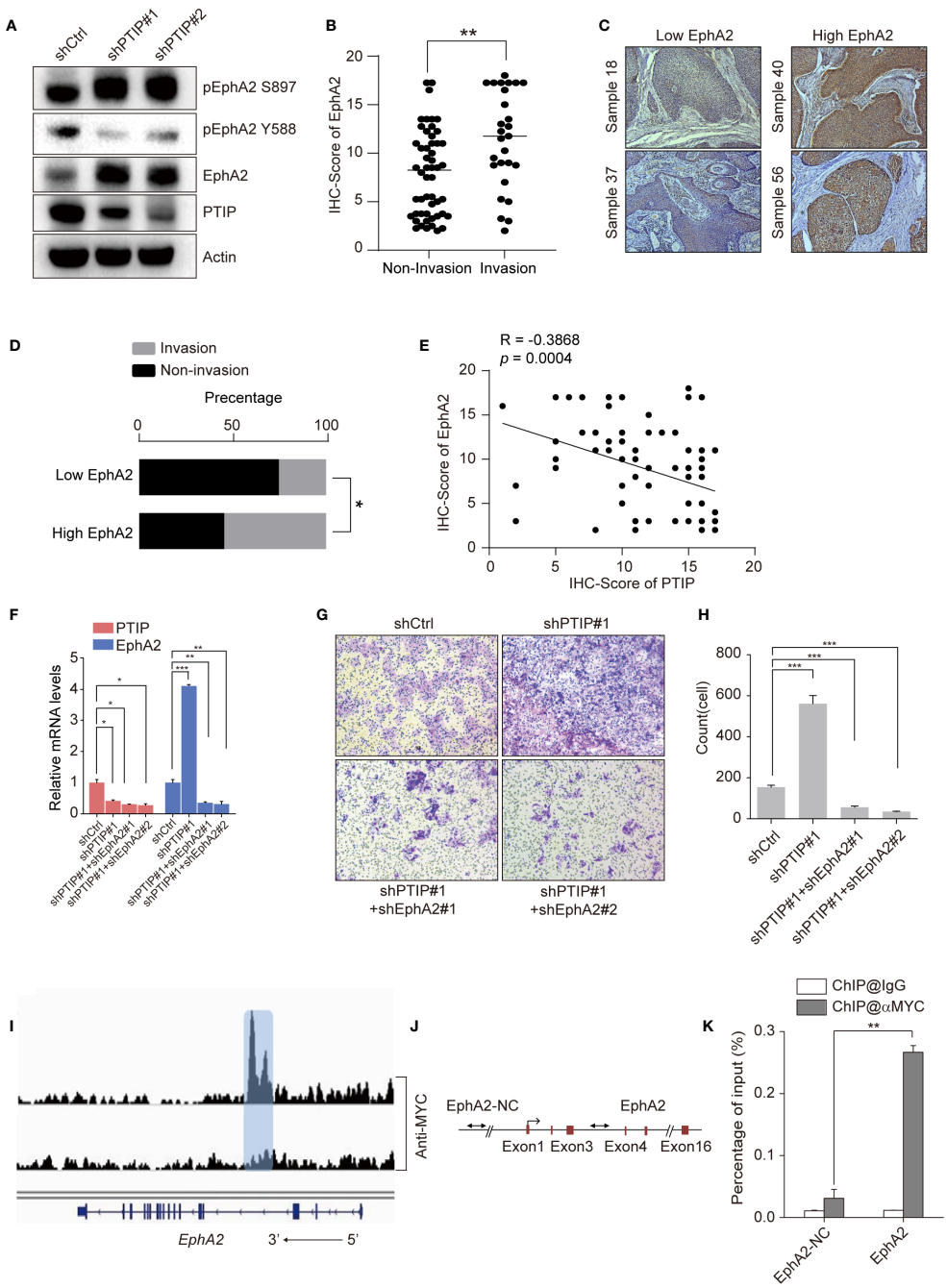


FIGURE 5 | PTIP inhibits ESCC cell invasion and migration through EphA2. **(A)** Western blotting analysis against PTIP, EphA2, pS897-EphA2, pY588-EphA2 in PTIP knockdown(shPTIP#1, shPTIP#2) and control (shCtrl) TE1 cells. **(B)** IHC score for EphA2 in ESCC tumor sections. Unpaired, two-tailed Student's t-test; $^{**}P < 0.01$. **(C)** Representative IHC images for EphA2 in ESCC tumor sections. Low EphA2 and high EphA2 groups were divided based on the EphA2 immunostaining intensity scores mentioned in method. Cut off for high and low EphA2 expression in ESCC was defined < or > 11. **(D)** Percentage of invasion and non-invasion in ESCC groups. The differences between rates were tested by $\chi 2$; $^{*}P < 0.05$. **(E)** Comparative expression between PTIP and EphA2 in ESCC samples from **(B)** analyzed by Pearson correlation. **(F)** Knockdown efficiency of shRNAs targeting PTIP and EphA2 in TE1 cells as determined by qRT-PCR. One-way ANOVA; $^{*}P < 0.05$,$^{**}P < 0.01$, $^{***}P < 0.001$. **(G, H)** The effect of PTIP and EphA2 double knockdown on the invasiveness of TE1 cells. For invasion assay, six different microscopic fields (magnification, ×10) from at least three independent experiments were examined; Relative intensities of the fields were measured (n ≥3). Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. One-way ANOVA; $^{***}P$ <0.001. **(I)** ChIP-seq density profiles for PTIP in TE1 cells. Gene models are shown below the density profiles. **(J, K)** ChIP-qPCR primer sets marked with arrows were designed to cover regions present within (EphA2) or outside (EphA2-NC) of the EphA2 gene **(C)**. ChIP-qPCR analyses of EphA2 binding **(D)**. Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. Unpaired, two-tailed Student's t-test; $^{**}P < 0.01$.

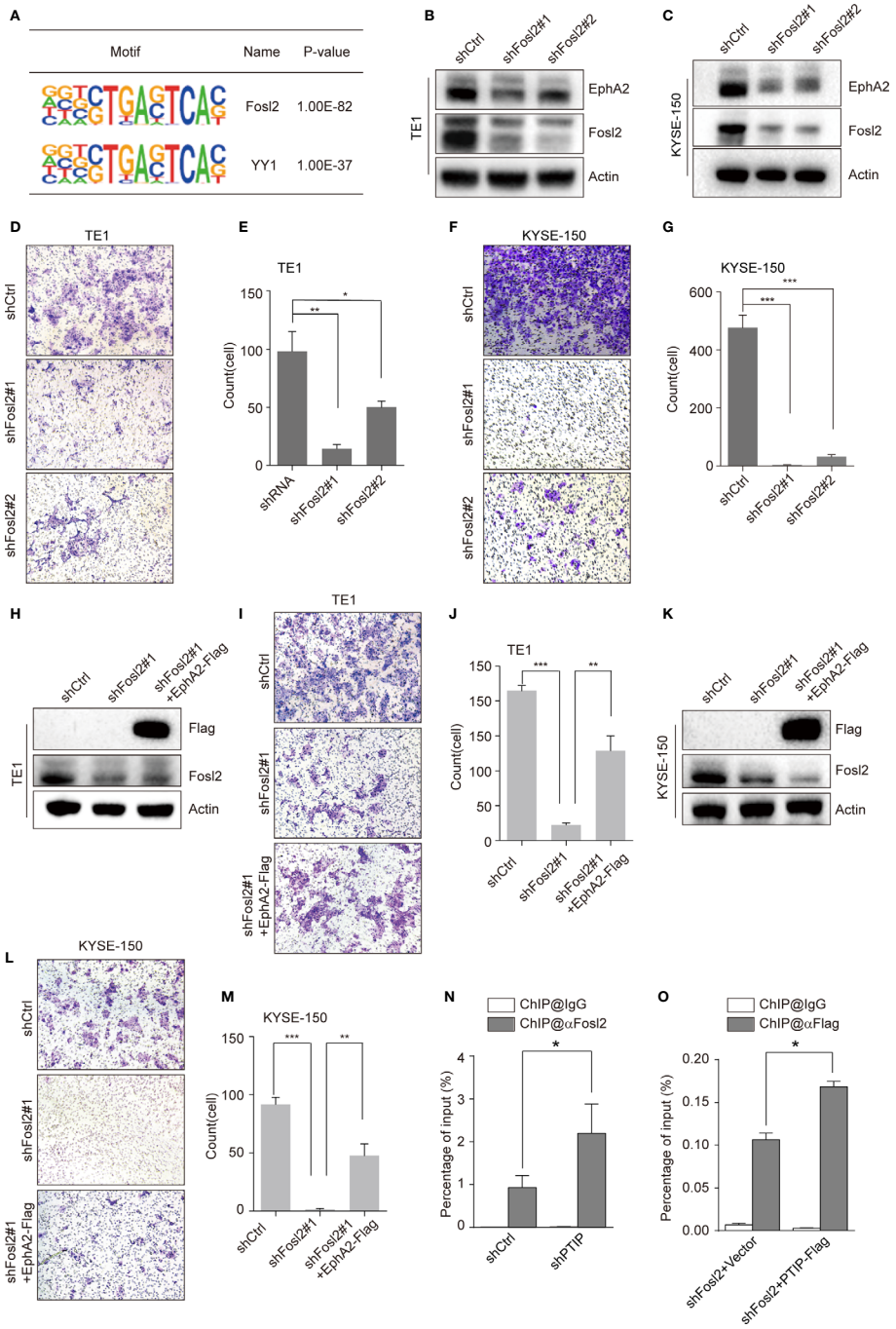


FIGURE 6 | PTIP inhibits EphA2 expression by competing with Fosl2 for binding to EphA2. **(A)** PTIP highly enriched binding motifs in bioPTIP ChIPseq were identified by HOMER. *P*-values are indicated within the boxes. **(B, C)** Western blot analysis of Fosl2 and EphA2 expression in Fosl2 knockdown (shFosl2#1, shFosl2#2) and control (shCtrl) TE1 **(B)** and KYSE-150 **(C)** cells. **(D–G)** The effect of Fosl2 knockdown on the invasiveness of TE1 **(D, E)** and KYSE-150 **(F, G)** cells. For invasion assay, six different microscopic fields (magnification, ×10) from at least three independent experiments were examined; Relative intensities of the fields were measured (n≥3). Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. One-way ANOVA; $^{*}P < 0.05$; $^{*}P < 0.01$; $^{***}P < 0.001$. **(H–M)** Western blot showing stable expression of Flag-EphA2 in shFosl2#1 TE1 cells **(H)** and KYSE-150 cells **(K)**. Overexpression EphA2 partially restored TE1 cells **(I, J)** and KYSE-150 cells **(L, M)** invasive ability, which attenuated by depletion of Fosl2. For invasion assay, six different microscopic fields (magnification, ×10) from at least three independent experiments were examined; Relative intensities of the fields were measured (n≥3). Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. One-way ANOVA; $^{*}P < 0.05$; $^{***}P < 0.001$. **(N, O)** ChIP-qPCR analyses of the relationship between Fosl2 and PTIP in transcriptional regulation for EphA2. Mean ± s.d. are given for three independent experiments. Unpaired, two-tailed Student's t-test; $^{*}P < 0.05$.

Previous studies have shown that the expression of EphA2 is regulated by multiple elements in human cancers (32–35). In HCC cells, testicular nuclear receptor 4 (TR_4) directly binds to the TR_4-response element located on the 5' promoter of EphA2, suppressing its transcription, and subsequently inhibiting HCC cell migration/invasion (32). Another study showed that the KRAS-driven MAPK and RalGDS-RalA signaling pathways promote EphA2 expression in colorectal (36). Moreover, EphA2 was also downregulated by miR-302b in gastric cancer. In this study, we revealed that EphA2 is regulated by PTIP and Fosl2 in ESCC. Targeting PTIP or Fosl2 may offer a route to circumvent the limitation of Eph/ephrin in drug development. Further preclinical evaluation of targeting PTIP and Fosl2 as a strategy to block tumor metastasis is therefore warranted. Similarly, further details of how PTIP is upregulated in ESCC cells need to be determined.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The name of the repository and accession numbers can be found below: National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA), https://www.ncbi.nlm.nih.gov/sra/, SRR13089664 to SRR13089679.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Huaian No.1 People's Hospitals' Ethics Committee. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by the Animal Experimentation Ethics Committee of Huai'an First People's Hospital.

AUTHOR CONTRIBUTIONS

XH performed experiments, analyzed, and interpreted data. YZhu performed experiments and analyzed data. LS, YZhou, and HG helped perform the experiments. WZ assisted with the Immunohistochemistry. LP, KH, YY, CJ, and JX collected and performed experiments using human CML samples. CZ designed the research, wrote the paper and provided the funding. LD designed the research and supervised the experiments. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (81500128), the Entrepreneurship and Innovation Project in Jiangsu Province and the Innovative Capacity-Building Plan in Huai'an (HAP201702).

ACKNOWLEDGMENTS

The authors would like to acknowledge the helpful suggestions concerning this study received from their colleagues.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.629916/full#supplementary-material

Supplementary Figure 1 | Depletion of PTIP do not alter ESCC cell survival, EC-matrix adhesion and morphology. **(a)** Representative FACS dotplots of YF647A-AnnexinV/PI staining used to assess apoptosis after knockdown of PTIP using two independent shRNAs targeting PTIP in the TE1 and KYSE-150 cell lines. **(b)** The rates of apoptosis in shCtrl versus shPTIP ESCC cells were determined by FACS. Mean ± s.d. are given for three in-dependent experiments. **(c)** Knockdown of PTIP did not significantly alter ESCC cell adhesion to collagen I. One-way ANOVA; $^{*}P < 0.05$. **(d)** Representative images of phalloidin staining in ESCC cells.

Supplementary Figure 2 | PTIP overexpression in ESCC cells did not significantly inhibit cell invasion. **(a-f)** Western blot showing stable expression of Flag-PTIP in TE1 cells **(a)** and KYSE-150 cells **(d)**. Overexpression PTIP did not inhibit TE1 cells **(b,c)** and KYSE-150 cells **(e,f)** invasion. For invasion assay, six different microscopic fields (magnification, ×10) from at least three independent experiments were examined; Relative intensities of the fields were measured (n≥3). Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. One-way ANOVA; n.s., not significant.

Supplementary Figure 3 | EFNA1 overexpression in ESCC cells decreased EphA2 expression level and inhibit cell invasion in KYSE-150. **(a-f)** Western blot showing stable expression of Flag-EFNA1 in TE1 cells **(a)** and KYSE-150 cells **(d)** decreased EphA2 expression. Overexpression EFNA1 did not inhibit TE1 cells invasion **(b,c)**, but not in KYSE-150 cells **(e,f)**. For invasion assay, six different microscopic fields (magnification, ×10) from at least three independent experiments were examined; Relative intensities of the fields were measured (n≥3). Representative images and statistical plots are shown; Mean ± s.d. are given for three independent experiments. One-way ANOVA; n.s., not significant; $^{**}P < 0.01$.

Supplementary Figure 4 | Depletion of YY1 did not increase EphA2 expression in ESCC cells. Western blotting analysis against EphA2, YY1 and GAPDH in YY1 knockdown(shYY1#1, shYY1#2) and control (shCtrl) TE1 cells.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
2. Zhao Y, Ma K, Yang S, Zhang X, Wang F, Zhang X, et al. MicroRNA-125a-5p enhances the sensitivity of esophageal squamous cell carcinoma cells to cisplatin by suppressing the activation of the STAT3 signaling pathway. *Int J Oncol* (2018) 53(2):644–58. doi: 10.3892/ijo.2018.4409
3. Hao JJ, Lin DC, Dinh HQ, Mayakonda A, Jiang YY, Chang C, et al. Spatial intratumoral heterogeneity and temporal clonal evolution in esophageal squamous cell carcinoma. *Nat Genet* (2016) 48(12):1500–7. doi: 10.1038/ng.3683
4. Gao J, Wang Y, Yang J, Zhang W, Meng K, Sun Y, et al. RNF128 Promotes Invasion and Metastasis Via the EGFR/MAPK/MMP-2 Pathway in Esophageal Squamous Cell Carcinoma. *Cancers* (2019) 11(6):840. doi: 10.3390/cancers11060840
5. Nieman DR, Peters JH. Treatment Strategies for Esophageal Cancer. *Gastroenterol Clinics North America* (2013) 42(1):187–97. doi: 10.1016/j.gtc.2012.11.007

6. Geiger TR, Peeper DS. Metastasis mechanisms. *Biochim Biophys Acta* (2009) 1796(2):293–308. doi: 10.1016/j.bbcan.2009.07.006
7. Pasquale EB. Eph receptors and ephrins in cancer: bidirectional signalling and beyond. *Nat Rev Cancer* (2010) 10(3):165–80. doi: 10.1038/nrc2806
8. Ieguchi K, Maru Y. Roles of EphA1/A2 and ephrin-A1 in cancer. *Cancer Sci* (2019) 110(3):841–8. doi: 10.1111/cas.13942
9. Miao H, Li DQ, Mukherjee A, Guo H, Petty A, Cutter J, et al. EphA2 mediates ligand-dependent inhibition and ligand-independent promotion of cell migration and invasion *via a* reciprocal regulatory loop with Akt. *Cancer Cell* (2009) 16(1):9–20. doi: 10.1016/j.ccr.2009.04.009
10. Hiramoto-Yamaki N, Takeuchi S, Ueda S, Harada K, Fujimoto S, Negishi M, et al. Ephexin4 and EphA2 mediate cell migration through a RhoG-dependent mechanism. *J Cell Biol* (2010) 190(3):461–77. doi: 10.1083/jcb.201005141
11. Miyazaki T, Kato H, Fukuchi M, Nakajima M, Kuwano H. EphA2 overexpression correlates with poor prognosis in esophageal squamous cell carcinoma. *Int J Cancer* (2003) 103(5):657–63. doi: 10.1002/ijc.10860
12. Das P, Veazey KJ, Van HT, Kaushik S, Lin K, Lu Y, et al. Histone methylation regulator PTIP is required to maintain normal and leukemic bone marrow niches. *Proc Natl Acad Sci USA* (2018) 115(43):E10137–E46. doi: 10.1073/pnas.1806019115
13. Lechner MS, Levitan I, Dressler GR. PTIP, a novel BRCT domain-containing protein interacts with Pax2 and is associated with active chromatin. *Nucleic Acids Res* (2000) 28(14):2741–51.
14. Wang X, Takenaka K, Takeda S. PTIP promotes DNA double-strand break repair through homologous recombination. *Genes Cells: Devoted Mol Cell Mech* (2010) 15(3):243–54. doi: 10.1111/j.1365-2443.2009.01379.x
15. Mu W, Wang W, Schimenti JC. An allelic series uncovers novel roles of the BRCT domain-containing protein PTIP in mouse embryonic vascular development. *Mol Cell Biol* (2008) 28(20):6439–51.
16. Kim D, Patel SR, Xiao H, Dressler GR. The role of PTIP in maintaining embryonic stem cell pluripotency. *Stem Cells* (2009) 27(7):1516–23. doi: 10.1002/stem.79
17. Hoffmeister A, Ropolo A, Vasseur S, Mallo GV, Bodeker H, Ritz-Laser B, et al. The HMG-I/Y-related protein p8 binds to p300 and Pax2 trans-activation domain-interacting protein to regulate the trans-activation activity of the Pax2A and Pax2B transcription factors on the glucagon gene promoter. *J Biol Chem* (2002) 277(25):22314–9. doi: 10.1074/jbc.M201657200
18. De Gregoriis G, Ramos JA, Fernandes PV, Vignal GM, Brianese RC, Carraro DM, et al. DNA repair genes PAXIP1 and TP53BP1 expression is associated with breast cancer prognosis. *Cancer Biol Ther* (2017) 18(6):439–49. doi: 10.1080/15384047.2017.1323590
19. Willis S, Villalobos VM, Gevaert O, Abramovitz M, Williams C, Sikic BI, et al. Single Gene Prognostic Biomarkers in Ovarian Cancer: A Meta-Analysis. *PloS One* (2016) 11(2):e0149183. doi: 10.1371/journal.pone.0149183
20. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* (2016) 11(9):1650–67. doi: 10.1038/nprot.2016.095
21. Lee TI, Johnstone SE, Young RA. Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat Protoc* (2006) 1(2):729–48.
22. Yu G, Wang L-G, He Q-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinf (Oxford England)* (2015) 31(14):2382–3. doi: 10.1093/bioinformatics/btv145
23. Jiang H, Shi Q-Q, Ge L-Y, Zhuang Q-F, Xue D, Xu H-Y, et al. Selenoprotein M stimulates the proliferative and metastatic capacities of renal cell carcinoma through activating the PI3K/AKT/mTOR pathway. *Cancer Med* (2019) 8 (10):4836–44. doi: 10.1002/cam4.2403
24. Pala A, Karpel-Massler G, Kast RE, Wirtz CR, Halatsch M-E. Epidermal to Mesenchymal Transition and Failure of EGFR-Targeted Therapy in Glioblastoma. *Cancers* (2012) 4(2):523–30. doi: 10.3390/cancers4020523
25. Li X, Zhang Y, Zheng L, Liu M, Chen CD, Jiang H. UTX is an escape from X-inactivation tumor-suppressor in B cell lymphoma. *Nat Commun* (2018) 9 (1):2720. doi: 10.1038/s41467-018-05084-w
26. Mihaylova Y, Abnave P, Kao D, Hughes S, Lai A, Jaber-Hijazi F, et al. Conservation of epigenetic regulation by the MLL3/4 tumour suppressor in planarian pluripotent stem cells. *Nat Commun* (2018) 9(1):3633. doi: 10.1038/s41467-018-06092-6
27. Kim JH, Sharma A, Dhar SS, Lee SH, Gu B, Chan CH, et al. UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells. *Cancer Res* (2014) 74(6):1705–17.
28. Cho YW, Hong T, Hong S, Guo H, Yu H, Kim D, et al. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem* (2007) 282(28):20395–406.
29. Fang M, Ren H, Liu J, Cadigan KM, Patel SR, Dressler GR. Drosophila ptip is essential for anterior/posterior patterning in development and interacts with the PcG and trxG pathways. *Development* (2009) 136(11):1929–38.
30. Wang J, Sun D, Wang Y, Ren F, Pang S, Wang D, et al. FOSL2 positively regulates TGF-β1 signalling in non-small cell lung cancer. *PloS One* (2014) 9 (11):e112150. doi: 10.1371/journal.pone.0112150
31. He J, Mai J, Li Y, Chen L, Xu H, Zhu X, et al. miR-597 inhibits breast cancer cell proliferation, migration and invasion through FOSL2. *Oncol Rep* (2017) 37(5):2672–8. doi: 10.3892/or.2017.5558
32. Jin RA, Lin H, Li G, Xu J, Shi L, Chang C, et al. TR nuclear receptor suppresses HCC cell invasion *via* downregulating the EphA2 expression. *Cell Death Dis* (2018) 9(3):283. doi: 10.1038/s41419-018-0287-5
33. Zhou Y, Sakurai H. Emerging and Diverse Functions of the EphA2 Noncanonical Pathway in Cancer Progression. *Biol Pharm Bull* (2017) 40 (10):1616–24. doi: 10.1248/bpb.b17-00446
34. Garcia-Monclús S, López-Alemany R, Almacellas-Rabaiget O, Herrero-Martín D, Huertas-Martinez J, Lagares-Tena L, et al. EphA2 receptor is a key player in the metastatic onset of Ewing sarcoma. *Int J Cancer* (2018) 143 (5):1188–201. doi: 10.1002/ijc.31405
35. Hong HN, Won YJ, Shim JH, Kim HJ, Han SH, Kim BS, et al. Cancer-associated fibroblasts promote gastric tumorigenesis through EphA2 activation in a ligand-independent manner. *J Cancer Res Clin Oncol* (2018) 144(9):1649–63. doi: 10.1007/s00432-018-2683-8
36. Dunne PD, Dasgupta S, Blayney JK, McArt DG, Redmond KL, Weir J-A, et al. EphA2 Expression Is a Key Driver of Migration and Invasion and a Poor Prognostic Marker in Colorectal Cancer. *Clin Cancer Res: Off J Am Assoc Cancer Res* (2016) 22(1):230–42. doi: 10.1158/1078-0432.ccr-15-0603

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 25 March 2021
doi: 10.3389/fonc.2021.660955

FLVCR1 Predicts Poor Prognosis and Promotes Malignant Phenotype in Esophageal Squamous Cell Carcinoma *via* Upregulating CSE1L

Suna Zhou[1,2,3†], *Mingxin Zhang*[4†], *Chao Zhou*[1,2,3], *Yinnan Meng*[1,2,3], *Haihua Yang*[1,2,3] *and Wenguang Ye*[5*]

[1] Laboratory of Cellular and Molecular Radiation Oncology, The Affiliated Taizhou Hospital, Wenzhou Medical University, Taizhou, China, [2] Department of Radiation Oncology, The Affiliated Taizhou Hospital, Wenzhou Medical University, Taizhou, China, [3] Key Laboratory of Minimally Invasive Techniques & Rapid Rehabilitation of Digestive System Tumor of Zhejiang Province, Taizhou, China, [4] Department of Gastroenterology, The First Affiliated Hospital of Xi'an Medical University, Xi'an, China, [5] Department of Gastroenterology, The Affiliated Taizhou Hospital, Wenzhou Medical University, Taizhou, China

OPEN ACCESS

Edited by:
Die Wang,
Hudson Institute of Medical Research, Australia

Reviewed by:
Shutao Zheng,
Xinjiang Medical University, China
Victor Wong,
OncoSeek limited, Hong Kong
Fan Shi,
Sixth Medical Center of PLA General Hospital, China

***Correspondence:**
Wenguang Ye
yewenguang_1982@163.com

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *30 January 2021*
Accepted: *10 March 2021*
Published: *25 March 2021*

Citation:
Zhou S, Zhang M, Zhou C, Meng Y, Yang H and Ye W (2021) FLVCR1 Predicts Poor Prognosis and Promotes Malignant Phenotype in Esophageal Squamous Cell Carcinoma via Upregulating CSE1L. Front. Oncol. 11:660955. doi: 10.3389/fonc.2021.660955

Objective: Dysregulation of feline leukemia virus subgroup C receptor 1(FLVCR1) expression has been investigated in several tumors. However, the expression and role of FLVCR1 in esophageal squamous cell carcinoma (ESCC) remain largely unknown.

Methods: FLVCR1 expression in tissues was measured by immunohistochemical staining (IHC). Celigo assay, MTT assay, colony formation, caspase 3/7 activity analysis, wound healing assay, Transwell migration, and invasion assay were applied to assess the effects of FLVCR1 on ESCC tumorigenesis. Coimmunoprecipitation (Co-IP) and liquid chromatography-mass spectrometry (LC-MS) were used to identify protein interactions with FLVCR1. An *in vivo* imaging system (IVIS) was used to investigate the functions of FLVCR1 on the growth and metastatic capability of ESCC cells in a xenograft model and a tail vein metastasis model.

Results: Elevated expression of FLVCR1 was detected in ESCC tissues and predicted poor survival. Upregulated FLVCR1 was positively correlated with lymph node metastasis (N stage) and late tumor-node-metastasis (TNM) stage. FLVCR1 knockdown inhibited cell proliferation and colony formation ability, induced cell apoptosis, and repressed cell migration and invasion of ESCC *in vitro*. Inhibition of FLVCR1 markedly repressed tumorigenicity and metastasis of ESCC cells *in vivo*. Mechanistically, chromosome segregation 1–like (CSE1L) was identified to interact with FLVCR1 using a Co-IP assay. Moreover, the inhibitory effect of FLVCR1 knockdown on proliferation and migration was counteracted by the exogenous expression of CSE1L.

Conclusion: FLVCR1 plays a pivotal role in ESCC cell survival, growth, and migration. These functions may be partially dependent upon the protein interaction between FLVCR1 and CSE1L. In addition, FLVCR1 can be applied as a clinical prognostic marker for patients with ESCC.

Keywords: FLVCR1, ESCC, CSE1L, migration, growth

INTRODUCTION

Esophageal squamous cell carcinoma (ESCC), the predominant type of esophageal carcinoma in Asia, is one of the most aggressive tumors with high incidence and mortality (1, 2). China is one of the countries with the highest incidences of esophageal carcinoma worldwide (3). Although the morbidity and mortality rates of ESCC are decreasing in China, the number of new cases and deaths accounts for more than 50% of the world's data (3). Despite improvements in both diagnostic and therapeutic technology, the 5-year survival rate of patients with ESCC remains proximately 20.9% in China (4). High metastatic ability and tumor recurrence are the leading causes of mortality in ESCC patients with or without surgery (5). The occurrence and development of ESCC is a complex process that involves multiple steps and genes. Thus, a better understanding of the role and the molecular mechanisms involved in the carcinogenesis of ESCC is expected to improve early diagnosis and effective targeting treatment.

FLVCR1, containing 12 hydrophobic transmembrane domains, is a member of the major facilitator superfamily and can transfer small solute molecules (6). FLVCR1 encodes two heme exporters, FLVCR1α localizing on the plasma membrane and FLVCR1β restricted to the mitochondria (6). The abnormal level of heme is essential for tumor progression and metastasis (7, 8). Chiabrando et al. reported that the heme exporter FLVCR1 is crucial for the survival of neuroblastoma cells by regulating heme metabolism (9). Peng et al. found that FLVCR1 plays a critical role in promoting the tumorigenicity and proliferation of synovial sarcoma *via* impeding apoptosis and autophagy (10). The overexpression of FLVCR1 in hepatocellular carcinoma is associated with higher disease staging, vascular invasion, histologic grade, and poorer outcomes (11). In summary, these results suggest that FLVCR1 may function as an oncoprotein during the process of tumor development. However, the role and molecular mechanisms of FLVCR1 involved in the malignant transformation of ESCC are still unknown.

In the present study, high expression of FLVCR1 was also detected in ESCC and played essential roles in promoting proliferation, invasion, and metastasis of ESCC and served as a poor prognostic marker. The results of co-immunoprecipitation (Co-IP) combined with western blotting suggested that CSE1L bound to the FLVCR1 protein. Further investigation showed that CSE1L was indeed involved in FLVCR1-mediated modulation of proliferation and migration of ESCC. Our study reveals that FLVCR1 may be a potential therapeutic target for ESCC treatment.

MATERIALS AND METHODS

Clinical Samples

Two sets of ESCC samples were included in this study. The first set containing 31 matched pairs of ESCC tumor tissues and normal tissues was purchased from Shanghai Outdo Biotech Inc (OD-CT-DgEso03-002) and used to analyze FLVCR1 expression at the protein level. FLVCR1 expression and the correlation between FLVCR1 expression and clinicopathological features were further analyzed in the second set containing 103 ESCC samples with clinical follow-up information. The second set of ESCC samples was taken from patients who underwent surgery at The Affiliated Taizhou Hospital of Wenzhou Medical University from January 2006 to December 2008. The follow-up term is from 6.6 to 9.5 years. The last follow-up time was July 2015. The Institutional Review Board of Taizhou Hospital approved the study, and informed consent was obtained from each patient.

Cell Lines

The human ESCC cell lines TE-1, KYSE-150, Eca-109, and Ec-9706 and the human normal esophageal epithelial cell line HEEC used in this study were purchased from Shanghai Cell Biology Institute of Chinese Academy of Sciences (Shanghai, China). Cells were maintained in RPMI 1640 supplemented with 10% FBS (Gibco, Grand Island, NY) and penicillin (100 IU/ml)/ streptomycin (100 μg/mL) (HyClone, Logan, UT, USA) at 37°C with 5% CO2.

Immunohistochemical Staining

Briefly, immunohistochemical staining (IHC) was performed as follows. The tissue sections were deparaffinized in xylene, rehydrated in graded ethanol and soaked in distilled water. After heat-mediated citric acid antigen retrieval, the tissue sections were incubated with the primary antibody for FLVCR1(Abcam, ab70838, rabbit polyclonal antibody) at a dilution of 1/500 at 4°C overnight and incubated with horseradish peroxidase (HRP)-conjugated antibody (Thermo Scientific, USA) for 60 min at room temperature. Next, the solutions were replaced with DAB solutions (Gene Tech, Shanghai, China). Finally, these processed tissue sections were observed under a microscope and then counterstained with hematoxylin. Known immunostaining positive slides served as a positive control. Slides incubated with an irrelevant rabbit antiserum were used as a negative control. The expression of FLVCR1 was scored according to the percentage and intensity of positively stained cells. As shown in **Supplementary 1**, the percentage of positive cells indicated the following scores: negative scored 0, 1-25% scored 1, 26-50% scored 2, 51-75% scored 3, and 76-100% scored 4; the intensity of positive cells indicated the following scores: negative scored 0; I grade scored 1; II grade scored 2; and III grade scored 3. The product of these two scores was defined as the total score. In this study, patients with a score≤ 6 were designated as the low-expression group, while patients with a score > 6 were designated as the high-expression group. The expression of FLVCR1 was assessed in a blinded manner by two observers.

Generation of Stable Overexpression or Downregulation Cell Lines

To downregulate the expression of FLVCR1, the most effective siRNA sequences were subcloned into the GV115(hU6-MCS-CMV-EGFP) vector to generate an shFLVCR1 recombinant lentivirus (GeneChem Corporation, Shanghai, China). TE-1 and KYSE-150 cells were infected with the recombinant lentivirus. To upregulate the expression of FLVCR1 or CSE1L,

the cDNA of FLVCR1 or CSE1L was amplified and subcloned into GV610(Ubi-3FLAG (Sigma)-MCS-CMV-EGFP-SV40-puro) to generate a FLVCR1-OE or CSE1L-OE recombinant lentivirus, and then KYSE-150 cells were infected with the recombinant lentivirus. Subsequently, cells were selected with puromycin (2 μg/ml) for 2 weeks.

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

By using Trizol (Invitrogen (Invitrogen, Carlsbad, CA), total RNA was extracted from cells followed by reverse transcription of RNA to cDNA using the PrimeScript RT reagent Kit (Takara, Shiga, Japan). qRT-PCR was performed with SYBR Green Real-time PCR Master Mixture (Takara, Shiga, Japan) using an Mx3000Ps qRT-PCR system (Applied, Foster City, CA, USA). The relative gene expression was normalized to the internal control GAPDH using the formula $2^{-\Delta\Delta Ct}$. The PCR conditions were as follows: pre-denaturing at 95° C for 15s and 45 cycles at 95°C for 5s and 60°C for 30s. The primers used in study were as follows: FLVCR1: 5′- GTGAGATTGGAGGGACAA GTAT-3′(upstream), 5′- TTTCATGGATGAGGAAAACG-3′ (downstream); CSE1L: 5′- CAGAAC ACGCTGACAAGTATCT-3′(upstream), 5′-AGCCCTGCGTCTAGTATCAATA-3′ (downstream); GAPDH:5′-TGACTTCAACAGCGACACCCA-3′ (upstream), 5′- CACCCTGTT GCTGTAGCCA AA-3′(downstream).

Western Blot Assay

Briefly, total cell protein was extracted using RIPA buffer (Beyotime, Shanghai, China) and then quantified by a BCA Protein Assay Kit (Beyotime). Protein samples were separated by 10% SDS-PAGE followed by transfer onto PVDF membranes, which were then blocked with 5% skim milk for 1 h. Subsequently, cells were incubated with anti-rabbit FLVCR1 (Abcam, ab70838, which is a rabbit polyclonal antibody, and can be used to analyze FLVCR1 protein expression in IHC-P and WB. The immunogenicity: Synthetic peptide conjugated to KLH derived from within residues 500 to the C-terminus of Human FLVCR.) or CSE1L (Abcam, ab96755) or anti-mouse Flag (Sigma, F1804) at 1/300,1/500, and 1/2000 dilutions for 12h at 4°. The anti-mouse GAPDH antibody (Santa-Cruz, sc-32233) was used as a loading control at a dilution of 1:2000. HRP-conjugated goat anti-rabbit or goat anti-mouse IgG antibody (CST, #7076, #7074) was applied as a secondary antibody at a dilution of 1:2000. The gray value of the protein band was analyzed by ImageJ software.

Celigo Assay

After transfection, the cells were seeded into 96-well plates at a density of $2x10^3$ cells/well. The number of cells with green fluorescence was scanned daily for 5 days by a Celigo® Image Cytometer (Nexcelom, Lawrence, MA, USA) from the second day after plating.

MTT Assay

Cells were seeded in 96-well plates at a density of $2x10^3$ cells/well after transfection. 20 μl 5 mg/ml MTT (Genview, Beijing, China) was added for the last 4 h before incubation termination at 37°C. Then, the medium was removed and 100 μl DMSO (Shanghai Shiyi Chemical Reagent, Shanghai, China) was added to dissolve the formazan crystals. After 2-3 min of oscillation, the viable cells were measured at 490 nm by a microplate reader (Tecan, Mnnedorf, Switzerland).

Colony Formation Assay

After infection, TE-1 and KYSE-150 cells were seeded into six-well plates at 600 and 800 cells/well followed by further incubation for 8 days, respectively. Plates were washed once with PBS, fixed with 4% paraformaldehyde (Sinopharm Chemical Reagent, Shanghai, China) for 30–60 min and washed again. Then, they were stained with crystal violet (Shanghai Biotechnology, Shanghai, China) for 10–20 min and washed with ddH2O several times. Finally, the colony number was counted and photographed under an IX71 fluorescence microscope (Olympus, Tokyo, Japan).

Caspase 3/7 Activity Analysis

After infection, TE-1 and KYSE-150 cells were seeded at a density of $2x10^4$ cells/well into 96-well plates. Ten milliliters of Caspase-Glo 3/7 buffer and substrate (Promega, Madison, WI, USA) were blended down to Caspase–Glo reaction solution. After the preparation of a reaction solution, each well was added with 100 μl Caspase-Glo reaction solution followed by 2 h incubation at room temperature. Next, the luminescence intensity was detected by a microplate reader at 570 nm.

Wound Healing Assay

ESCC cells were seeded into 96-well plates at a density of $5x10^4$ cells/well. FBS-free medium was replaced with 10% FBS-containing RPMI-1640 medium until the monolayer adherent cells reached 90% confluence. A wound was scraped across the cell monolayer. Photomicrographs were taken with an IX71 fluorescence microscope at zero time points and after 4 h and 8 h. A Celigo® Image Cytometer was used to quantify the wound area and to calculate the ratio of migration.

Transwell Migration and Invasion Assay

After infection, TE-1 and KYSE-150 cells were respectively seeded onto the upper chamber with or without Matrigel (Corning Costar, Cambridge, MA, USA) at 8×10^4 and 1×10^5 cells/well. RPMI-1640 medium containing 30% FBS was added to the lower chamber. A cotton swab was used to wipe the non-migrated/non-invasive cells after 20-48 h incubation. The migrated/invasive cells were fixed and stained with Giemsa, followed by cellular count and photograph under a microscope.

In Vivo Growth and Metastasis Assay

For the *in vivo* tumor growth assays, 4×10^6 KYSE-150 cells stably transfected with shFLVCR1 (KD) or the negative control (NC) were separately transplanted subcutaneously into the right-back flank of 4-week-old female BALB/c nude mice (Shanghai SLAC Laboratory Animal, Shanghai, China). They were divided into KD and NC groups with 10 nude mice per group. Tumor sizes

were measured every other 4 days until the mice were humanely sacrificed and calculated using the formula: (length × width × width)/2. The mice were intraperitoneally injected with D-luciferin and photographed with the In Vivo Imaging System (IVIS)-Lumina LT (Perkin Elmer, Waltham, USA) at the end of the fourth week after transplantation. Then, the tumors were harvested and weighed when the mice were euthanized by CO2 asphyxia. For *in vivo* metastasis assays, 2×10^6 KYSE-150 cells stably transfected with shFLVCR1(KD) or the negative control (NC) (10 per group) were separately injected into the lateral tail veins of 5-week-old female BALB/c nude mice. After 8 weeks, the mice were also photographed under the In Vivo Imaging System (IVIS)-Lumina LT and euthanized to harvest their lungs, livers, and tumors.

Co-IP and LC-MS

KYSE-150 cells stably transfected with 3×Flag-tagged-FLVCR1 (OE) or control (NC) were washed two times with PBS for further study when the cells reached more than 80% confluence. Total cell protein was extracted using immunoprecipitation lysis buffer (Beyotime, Shanghai, China) and then quantified by a BCA Protein Assay Kit. Equal amounts of protein from the NC and OE groups were co-immunoprecipitated with anti-FLAG beads (Sigma, A2220). The products of Co-IP were exchanged and concentrated with 3×FLAG peptide (Sigma, F4799). The concentrated samples were analyzed by SDS-PAGE utilizing Coomassie Blue staining. Protein bands were excised from the gel and digested into peptides by trypsin. The sample of each peptide was identified by the LC-MS method, and then PD/MASCOT software was used to search the protein database to identify FLVCR1-interacting proteins. Finally, these potential FLVCR1-interacting proteins were summarized in the form of a gene network diagram by using GO analysis and KEGG analysis.

Statistical Analysis

All experiments were repeated in triplicate. The data are represented as the mean ± SD in our study. Statistical analyses were performed using SPSS software (version 23.0) and GraphPad Prism software (version 8.0). The correlation between FLVCR1 expression and clinicopathological factors of ESCC was analyzed by the Chi-square test and Fisher's exact probability test. The prognosis of FLVCR1 was identified by using univariate and multivariate Cox regression analyses. Survival analysis was performed by the Kaplan–Meier method and the log-rank test. The comparison between the groups was analyzed *via* Student's t-test and Mann–Whitney U test. $P<0.05$ was defined as statistically significant.

RESULTS

FLVCR1 Is Overexpressed in ESCC and Associated With Poor Prognosis

Based on TCGA data (http://ualcan.path.uab.edu), higher expression of FLVCR1 was detected in most carcinomas, including esophageal cancer (**Figure 1A**). In the present paper, a tissue microarray containing 31 paired ESCC and non-tumorous samples was stained with IHC to reveal the phenotype of FLVCR1 expression in ESCC. The results showed that the appearance of FLVCR1 was significantly higher in ESCC samples than in adjacent non-tumorous samples (**Figures 1B, C**). Moreover, the second set, including 77 matched ESCC and normal tissues, was stained with the IHC method to analyze the expression and prognosis of FLVCR1 in ESCC. As a result, the upregulated expression of FLVCR1 was noticed in 27 out of 77 ESCC cases (**Table 1**), but rarely detected in corresponding normal esophageal tissues. As shown in **Table 2** and **Figure 1D**, the results from 103 ESCC patients indicated that high FLVCR1 expression was significantly associated with lymph node-metastasis (N stage) and late tumor-node-metastasis (TNM stage), while there was no significant association between FLVCR1 expression and age, gender, differentiation grade, P53 positivity, Ki67 positivity, PDL-1 positivity or CD8 positivity. Univariate Cox regression analyses showed that the variables, including high FLVCR1 expression, and male sex, advanced T, N, and TNM stage, were associated with shorter overall survival (**Figure 1E**). However, multivariate Cox regression analysis revealed that FLVCR1 was not an independent prognostic factor for overall patient survival in patients with ESCC (**Figure 1F**). The representative staining results of ESCC patients with each T grade, N grade, and TNM stage were shown in **Figure 1G**. Furthermore, Kaplan-Meier analyses were used to analyze the correlation between FLVCR1 expression and overall survival in ESCC patients. ESCC patients with higher FLVCR1 expression exhibited a remarkably decreased overall survival (**Figure 1H**). Overall, these data indicated that the upregulation of FLVCR1 acts as a poor prognostic indicator of ESCC and might contribute to ESCC advancement.

Knockdown of FLVCR1 Inhibits ESCC Cell Proliferation, and Colony Formation and Promotes Apoptosis

The RT-PCR results showed that FLVCR1 mRNA levels were overexpressed in KYSE-150, TE-1, Eca-109, and EC9706 cells. In contrast, FLVCR1 expression was low in HEEC cells (**Supplementary 2**). TE-1 and KYSE-150 cell lines were used to discover the function of FLVCR1 in the malignant characteristics of ESCC. ESCC cells with FLVCR1 knockdown (TE-1-shFLVCR1 and KYSE-150- shFLVCR1) were established, and the knockdown efficiency was confirmed by western blot and RT-PCR (**Figures 2A, B**). The Celigo and MTT assay data confirmed the inhibitory effect of FLVCR knockdown on the proliferation of ESCC cells (**Figure 2C**). As shown in **Figure 2D**, enhanced caspase-3/7 activity was detected in ESCC cells with FLVCR1 knockdown, indicating that FLVCR1 knockdown could induce apoptosis. A colony formation assay was conducted to further verify that cell colony formation could be reduced by FLVCR1 knockdown in ESCC cells (**Figure 2E**).

Knockdown of FLVCR1 Impedes ESCC Cell Migration and Invasion *In Vitro*

Initially, we conducted a wound-healing assay to elucidate the role of FLVCR1 in the migration ability of ESCC cells. As shown in **Figure 3A**, inhibition of FLVCR1 did not have any effect on the migration of KYSE-150 and TE-1 cells compared to that of shCtrl groups after 8 h. Subsequently, we examined the role of

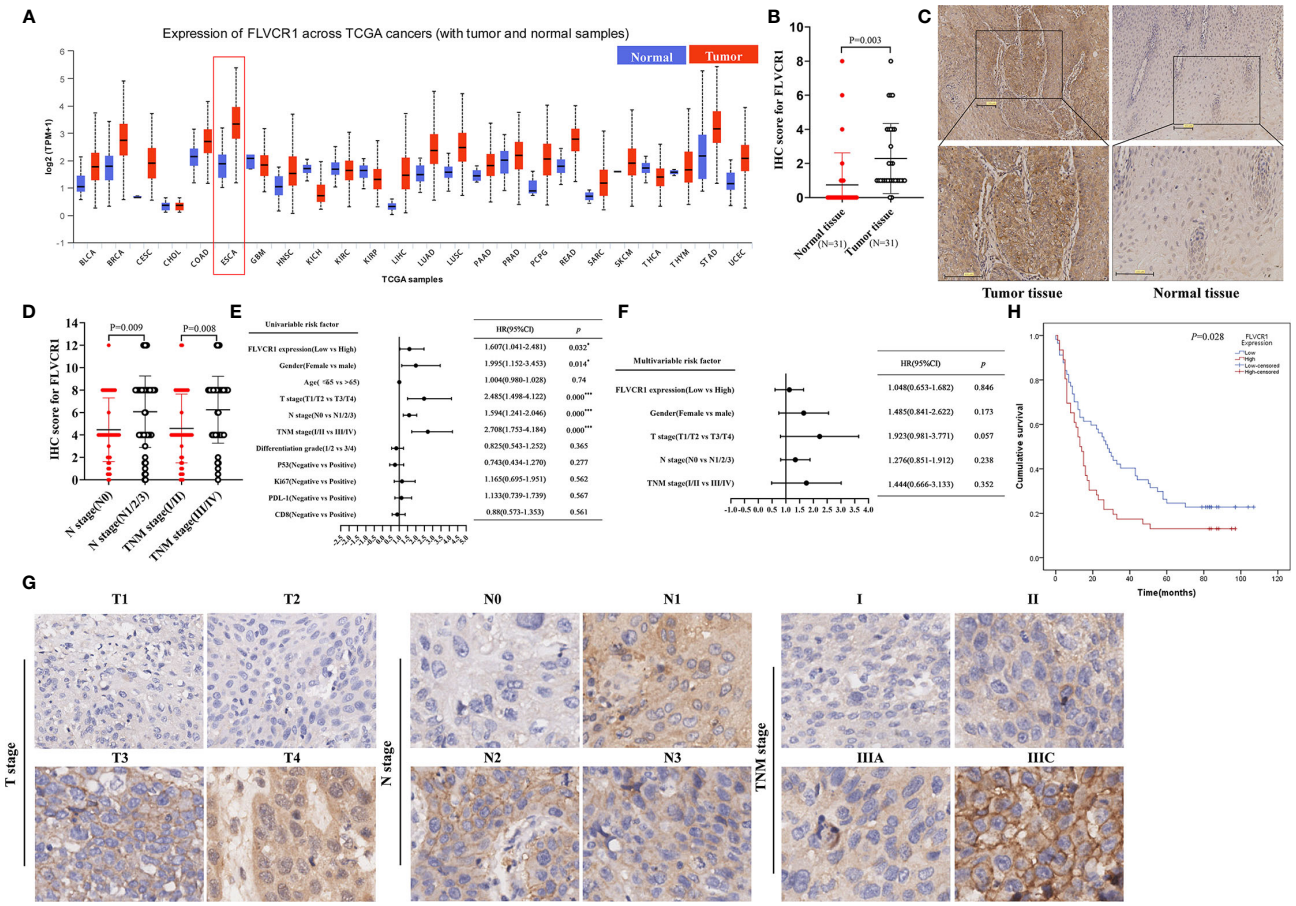


FIGURE 1 | FLVCR1 expression is upregulated in ESCC tissues and predicts a poor clinical outcome. **(A)** Pan-cancer analysis of FLVCR1 based on TCGA data. **(B)** The IHC scores of FLVCR1 in 31 matched normal tissues and ESCC tissues. **(C)** Representative photographs of FLVCR1 IHC staining in normal and tumor tissues. (Scale bar: 100μm; Magnification×200.) **(D)** The IHC scores of FLVCR1 in ESCC patients with different N stages and TNM stages. **(E, F)**. Forest plot showing the association between FLVCR1 expression and ESCC survival using univariate **(E)** and multivariate **(F)** analyses. (HR, hazard ratio; CI, confidence interval) **(G)** Representative images of FLVCR1 staining according to the T stage, N stage, and TNM stage of the ESCC patients (magnification, 200×). **(H)** Comparison of overall survival of ESCC patients with different FLVCR1 expression.

TABLE 1 | Differential expression of FLVCR1 in 77 paired of ESCC and adjacent tissues.

	FLVCR1 expression		*p*
	High	Low	
			0.000***
Tumor tissues	27	50	
Adjacent tissues	2	75	

****P < 0.001.*

FLVCR1 in the migratory ability of ESCC cells by using a Transwell migration assay. In contrast to the results of the wound healing assay, the negative effect of FLVCR1 knockdown on the migration ability of ESCC cells was confirmed in a Transwell migration assay (**Figure 3B**). Finally, a Transwell invasion assay was performed to assess the role of FLVCR1 in invasion. Similarly, the results of the Transwell invasion assay demonstrated that knockdown of FLVCR1 markedly attenuated the invasion of KYSE-150 and TE-1 cells compared with that of cells transfected with shCtrl (**Figure 3B**). In summary, these data revealed the negative effect of FLVCR1 knockdown on the migration and invasion of ESCC cells *in vitro*.

Knockdown of FLVCR1 Suppresses ESCC Tumor Growth and Metastasis *In Vivo*

The role of FLVCR1 in tumor growth and metastasis was further investigated in a tumor-bearing mouse model. As a result, inhibition of FLVCR1 slowed the tumor growth of KYSE-150 cells *in vivo* (**Figure 4A**). Furthermore, KYSE-150-shFLVCR1 cell-generated tumors had a smaller volume and lighter weight than KYSE-150-shCtrl cell-generated tumors (**Figures 4B, C**). The tumor-bearing mice were also photographed and analyzed by a non-invasive *in vivo* imaging system before sacrifice. The fluorescent radiant efficiency for the region of interests was calculated and demonstrated significant attenuation over

TABLE 2 | The relationship between FLVCR1 expression status and clinicpathologic features of ESCC.

Clinicpathologic features		Total	FLVCR1 expression		*p*
			Low	High	
Age (years)					0.116
	≤65	50	24	26	
	>65	52	33	19	
Gender					0.233
	Female	26	17	9	
	Male	77	40	37	
T stage					0.063
	T1/T2	19	14	5	
	T3/T4	78	39	39	
N stage					0.009**
	N0	46	32	14	
	N1/N2/N3	57	25	32	
TNM stage					0.008**
	I/II	47	32	15	
	III/IV	51	21	30	
Differentiation grade					0.411
	1/2	80	46	34	
	3/4	23	11	12	
P53					0.234
	Negative	24	11	13	
	Positive	53	32	21	
Ki67					0.723
	Negative	30	16	14	
	Positive	47	27	20	
PDL-1					0.896
	Negative	51	28	22	
	Positive	52	28	24	
CD8					0.896
	Negative	50	28	22	
	Positive	53	29	24	

*T, tumor-metastasis; N, node-metastasis; TNM, tumor-node-metastasis; **P < 0.01.*

control mice that were transplanted with KYSE-150-shCtrl cells (**Figures 4D, E**). We further investigated whether inhibition of FLVCR1 impeded metastasis *in vivo*. Seven weeks after tail vein injection, the transplanted mice were studied under an *in vivo* imaging system and carefully dissected. The lungs, livers, and tumors were removed for further analysis. Metastases were found in the lungs but not in the livers. Although the florescent intensity measurement is not significantly different between KYSE-150-shFLVCR1 and -shCtrl groups (as shown in **Supplementary Materials**), the frequency of metastasis to the lungs was decreased in mice injected with KYSE-150-shFLVCR1 cells as compared to KYSE-150-shCtrl cells (**Figures 5A–C**), suggesting a trend of functional impact of FLVCR1 on metastasis. Additionally, 5 out of 10(50%) mice implanted with shCtrl cells generated metastases in the groin, but none was found in the groin of the shFLVCR1 group. According to the above results, we confirmed the promoting function of FLVCR1 on tumor growth and metastasis *in vivo*.

FLVCR1 Physically Interacted With CSE1L in ESCC Cells

To explore the molecular mechanism by which FLVCR1 promotes proliferation and metastasis in ESCC cells, we sought to identify FLVCR1-interacting proteins in ESCC cells. We established stable KYSE-150 cell lines that constitutively overexpressed FLVCR1 containing a FLAG-tag (KYSE-150-FLVCR1-OE). The overexpression efficiency was confirmed by using a western blot assay with an anti-FLAG monoclonal antibody (**Figure 6A**). MS analysis identified 28 specific FLVCR1 interacting proteins based on the criteria of proteins that could be detected in the OE group (**Figures 6B, C**). Based on the results of MS and bioinformatics analysis, 21 genes (PLCD4, HK1, CSE1L, TNPO1, MDK, DSC1, CEBPZ, SEC61A1, DDX47, RPS15A, SFPQ, EIF4A2, UTP6, RPS26, CAPZB, MDH2, IMMT, DLST, CYFIP2, DLD, PKLR) were found. To investigate the interaction between FLVCR1 and these genes, bioinformatic analysis was applied to draw a diagram of the gene interaction network surrounding the regulation of FLVCR1 (**Figure 6D**). The top-ten highly expressed genes (HK1, CSE1L, TNPO1, SEC61A1, RPS15A, SFPQ, EIF4A2, UTP6, CAPZB, and MDH2) in esophageal cancer were further selected to analyze the differential expression in the esophageal cancer patient cohort in TCGA database using UCSC Xena (**Figure 6E**). Subsequently, UALCAN websites were employed to conduct expression profile analysis based on tumor histology (**Figure 6F**). The expression levels of CSE1L, RPS15A, SFPQ, and CAPZB in ESCC were significcantly higher than those in normal tissue and esophageal adenocarcinoma (EAC). Furthermore, we analyzed the survival prognosis of these genes in esophageal cancer patients *via* UCSC Xena. As shown in **Figure 6G**, lower expression of CSE1L was associated with longer overall survival in esophageal cancer patients. Therefore, CSE1L was selected as a potential downstream target gene of FLVCR1. The Co-IP assay illustrated that FLVCR1 and CSE1L could coprecipitate with each other in KYSE-150 cells (**Figure 6H**). Although anti-FLVCR1 antibody could not detect the target protein in the input sample, gray values of WB bands showed that the FLVCR1 level of the Flag-tagged transfected cell is higher than the control in the input sample after fusing the Flag tag (**Supplementary 3**).

Restoration of CSE1L Expression Counteracts the Effects of FLVCR1 Knockdown

Next, rescue experiments were performed to verify the role of CSE1L in the impacts of FLVCR1 on the proliferation and migration of ESCC cells. The visualization of green and red fluorescence was detected to evaluate the infective efficiency under a fluorescence microscope, and more than 80% of KYSE-150 cells were successfully transfected with the recombinant plasmids (**Figure 7A**). The results of western blot analysis suggested that FLVCR1 knockdown substantially downregulated the expression of CSE1L in KYSE-150 cells (**Figure 7B**). However, CSE1L was highly re-expressed in KYSE-150 cells cotransfected with FLVCR1 knockdown and ectopic-expression of CSE1L (**Figure 7B**). For re-expressed CSE1L, the growth inhibitory effect of FLVCR1 knockdown on KYSE-150 cells was attenuated (**Figures 7C–E**), and the reduced

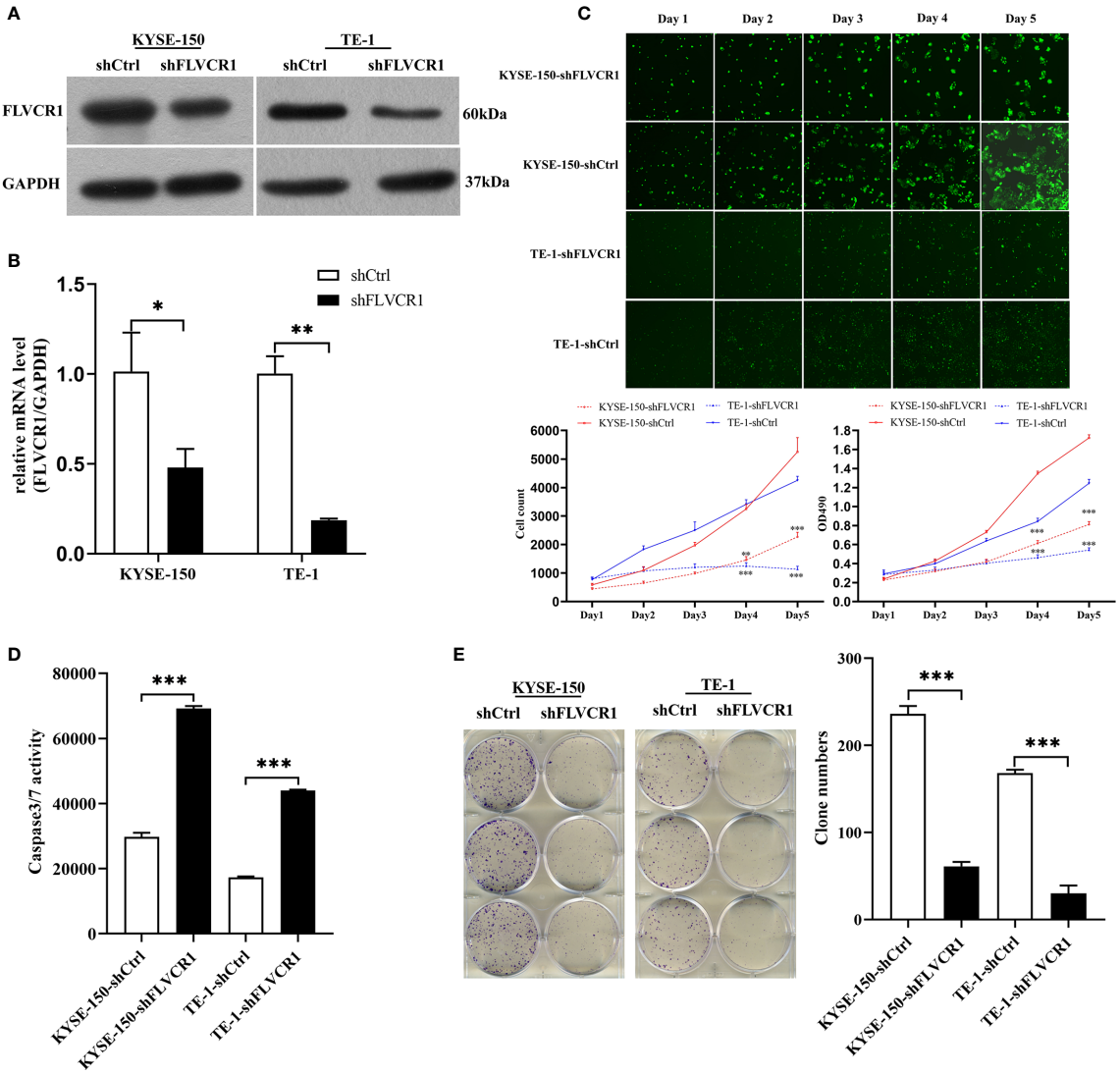


FIGURE 2 | The effect of FLVCR1 on ESCC cell proliferation and apoptosis in vitro. **(A, B)** Protein and mRNA expression of FLVCR1 was significantly downregulated in shFLVCR1 groups compared with shCtrl groups. **(C)** The proliferation rate of TE-1 and KYSE-150 cells was reduced by FLVCR1 knockdown, which was confirmed by the Celigo assay and MTT assay. **(D)** Knockdown of FLVCR1 remarkably enhanced caspase-3/7 activity in the shFLVCR1 groups. **(E)** FLVCR1 knockdown decreased the colon numbers in cells of shFLVCR1 groups compared to cells of shCtrl groups, respectively. Data were presented as the mean±SD of three separate experiments. *P < 0.05, **P < 0.01, ***P < 0.001.

migratory capability of FLVCR1 knockdown on ESCC cells was also abolished (**Figures 7F, G**).

DISCUSSION

Recent studies have focused on the crucial role of FLVCR1 in tumor promotion. For instance, upregulated FLVCR1 promoted the cellular proliferation of neuroblastoma and synovial sarcoma (9, 10). Aberrant expression of FLVCR1 was correlated with aggressive tumor stage and poor survival in hepatocellular carcinoma (11). The current study demonstrated a pivotal role and potential mechanism of FLVCR1 in regulating the proliferation and migration of ESCC for the first time. Our results exhibited that downregulation of FLVCR1 remarkably induced apoptosis and suppressed the proliferation, migration, invasion, and colony formation of ESCC cells. These findings indicated the function of FLVCR1 in the formation and progression of the malignant phenotype of ESCC.

Interestingly, wound healing assays showed that FLVCR1 knockdown had not begun to affect the migration of ESCC cells after as little as 8 hours. The inhibition of FLVCR1 knockdown on ESCC migration was investigated by a Transwell migration assay after more than 20 hours of

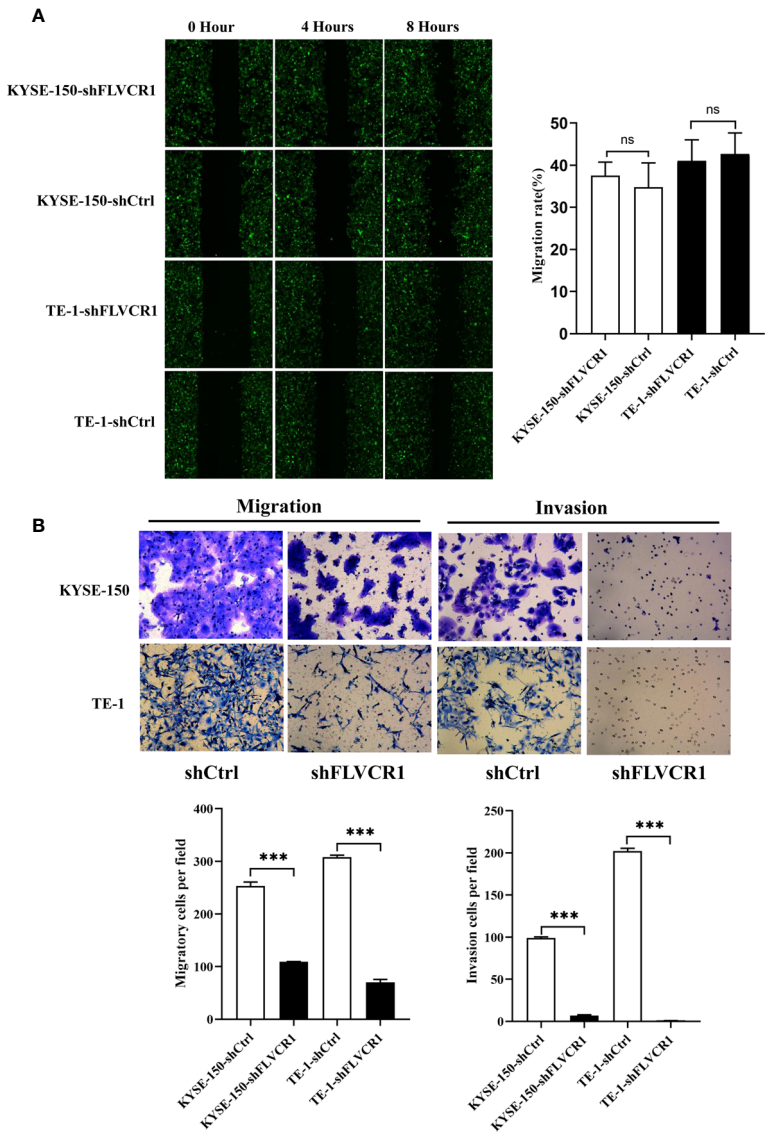


FIGURE 3 | Knockdown of FLVCR1 decelerates ESCC cell migration and invasion. **(A)** Wound healing assay showed that inhibition of FLVCR1 did not significantly affect the motility of ESCC cells at 8 hours. **(B)** Transwell assays revealed that knockdown of FLVCR1 attenuated cell migration at 20 hours and reduced cell invasion at 48 hours. ***P < 0.001; ns, no significance.

incubation, suggesting that FLVCR1 may have a time-dependent effect on the cellular migration of ESCC. Tumor metastasis is a comprehensive process during which cells detach from the primary site, disseminate throughout the body, and form new metastases by adhesion. By wound healing assay, cells moved to the wound by laterally creeping and displayed a horizontal migration. However, a Transwell migration assay showed stereoscopic migration (12, 13). Also, wound healing assays showed the motility of clusters of cells, while Transwell migration assays revealed the migratory ability of a single cell (14, 15). Furthermore, the vital role of FLVCR1 in tumorigenesis and migration of ESCC was confirmed in a tumor-bearing mouse model.

Next, Co-IP and LC-MS were utilized to explore the potential FLVCR-binding protein, which might influence the proto-oncogene FLVCR1. In this study, CSE1L was identified as a novel downstream target of FLVCR1 and implicated in the regulation of tumor growth and migration in ESCC. CSE1L, a human homolog of the yeast chromosome segregation protein, is up-regulated in multiple types of carcinomas, including colorectal cancer, lung cancer, breast cancer, and so on (16). CSE1L was uncovered to promote the tumor progression by affecting cell proliferation, apoptosis, and invasion of colorectal cancer (17). CSE1L silencing could induce apoptosis and repress the proliferation and invasion of gastric cancer cells by regulating the PI3K/Akt/mTOR and MEK/ERK signaling pathways (18). In

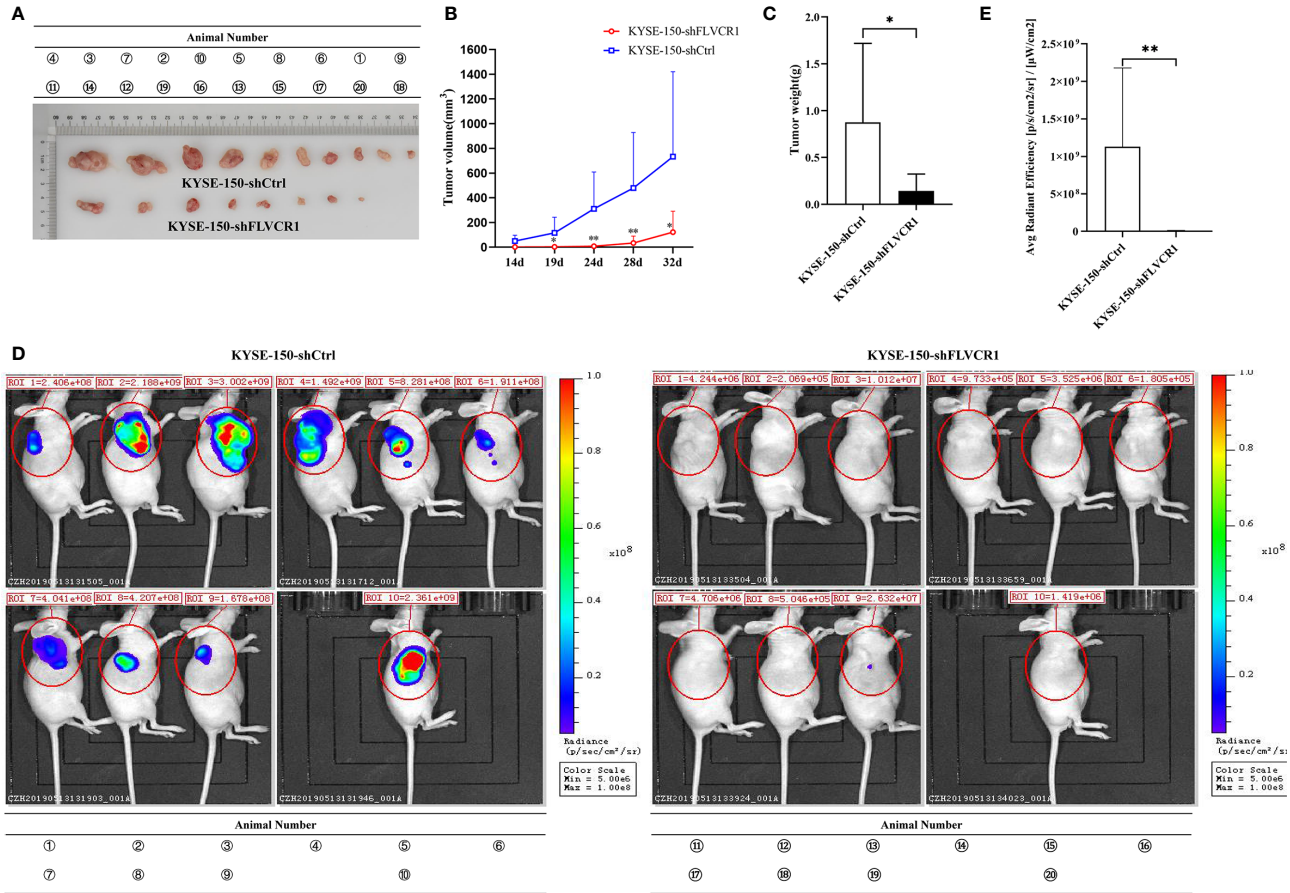


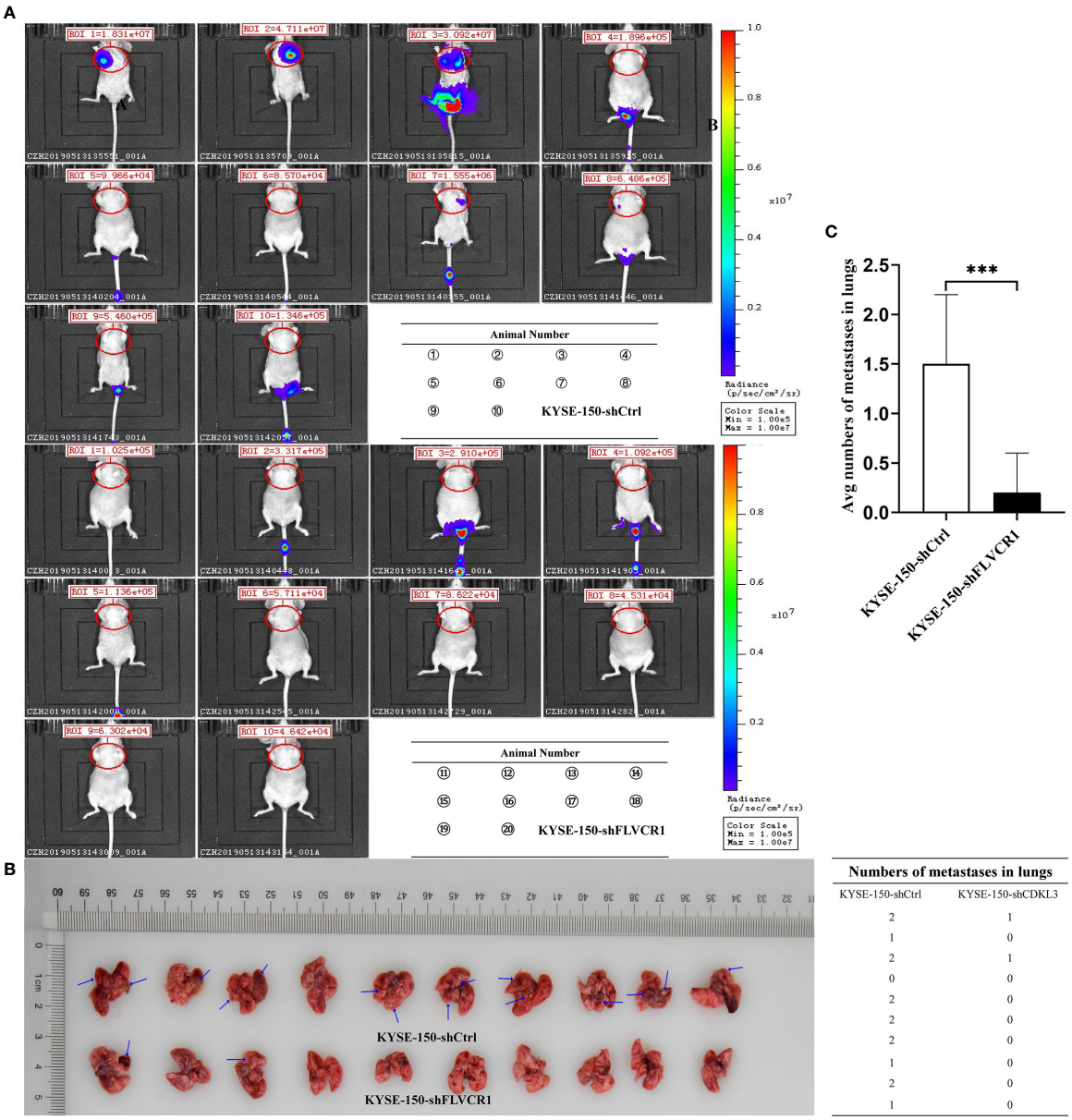
FIGURE 4 | FLVCR1 knockdown retarded tumor growth in vivo. **(A)** KYSE-150-shFLVCR1 cells generated smaller tumors in nude mice than the KYSE-150-shCtrl cells. **(B)** The growth curve showed that FLVCR1 knockdown suppresses growth. **(C)** KYSE-150-shFLVCR1-formed tumors had significantly lighter weight than KYSE-150-shCtrl cell-formed tumors. **(D)** *In vivo* imaging of tumor-bearing mice. **(E)** Seriously weakened fluorescent radiant efficiency was found in KYSE-150-shFLVCR1 cell-injected mice compared to that of control mice. Data were represented as the mean ± SD. *P<0.05, **P<0.01.

addition, a recent study revealed that the abnormal expression of CSE1L was correlated with neoplastic progression in Barrett's esophagus (19). However, the role of CSE1L has not been studied in esophageal cancer.

CSE1L was one of the proteins that interacted with FLVCR1 based on our data. Therefore it was selected as an essential candidate for further ESCC research. The results of Co-IP showed the interaction between CSE1L and FLVCR1, indicating that CSE1L might be required for FLVCR1-mediated tumor promotion in ESCC. CSE1L expression was downregulated by FLVCR1 knockdown in KYSE-150 cells. The inhibition of ESCC induced by FLVCR1 knockdown can be partly restored by CSE1L overexpression. However, more comprehensive and in-depth studies are needed in our next plan to explore the precise molecular mechanisms involved in interactions between FLVCR1 and CSE1L. Besides, CSE1L can be detected not only in tumor tissues but also in body fluids, especially in the blood, indicating that CSE1L can be used as a tumor serum biomarker (20–22). Therefore, our further research plan is to investigate the diagnostic and prognostic value of serum CSE1L in ESCC.

Moreover, there is no information about the indicative function of FLVCR1 in ESCC prognosis in previous reports. In this paper, the correlation between FLVCR1 expression and the clinicopathological features of ESCC patients was identified by immunohistochemical analysis. The levels of FLVCR1 expression in ESCC tissues were significantly higher than those in corresponding normal tissues. Additionally, the upregulation of FLVCR1 expression was positively correlated with the aggressive characteristics of ESCC, including advanced lymph node metastasis and advanced TNM stage. In a previous study, overexpression of FLVCR1 was known to be strongly associated with poorer survival outcomes of patients with hepatocellular carcinoma (11). In this paper, FLVCR1 was not an independent prognostic factor in ESCC, but a prognostic marker for worse survival in ESCC patients.

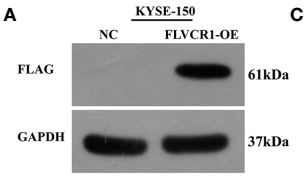
Furthermore, another analysis was conducted to evaluate the relationship between FLVCR1 expression status and the expression of p53, Ki67, PDL-1a, and CD8. Ki67, a proliferation indicator, is related to the poor prognosis of patients with various tumors (23–25). P53 is the most frequently mutated gene in ESCC,

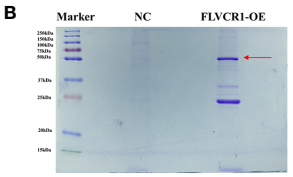


Numbers of metastases in lungs	
KYSE-150-shCtrl	KYSE-150-shCDKL3
2	1
1	0
2	1
0	0
2	0
2	0
2	0
1	0
2	0
1	0

FIGURE 5 | FLVCR1 knockdown reduced tumor metastasis *in vivo*. **(A)** *In vivo* imaging of tumor-bearing mice. **(B)** FLVCR1 knockdown suppressed the metastatic potential of KYSE-150 cells in nude mice. Red arrow: metastases **(C)** Severely weakened fluorescent radiant efficiency was found in KYSE-150-shFLVCR1 cells that formed metastases compared to that of control mice, ***P < 0.001.

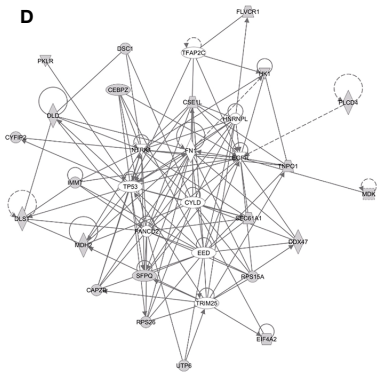
and p53 mutation shows excellent prediction performance on treatment response and overall survival in ESCC (26–28). Abnormal PDL-1 expression has been identified in various human malignancies, and antibody-mediated blockade of PDL-1 shows efficient antitumor activity (29). The PD−1/PDL-1 pathway can prevent effective antitumor immunity by downregulating the function of CD8+ T lymphocytes, so ESCC patients who are PDL-1−positive have a poorer prognosis than negative patients (30). However, our results showed that there was no association between FLVCR1 expression and the positive expression of Ki67, p53, PDL-1, and CD8. Our data on the expression of Ki67, p53, PDL-1, and CD8 from IHC are qualitative results, and more accurate quantitative analysis should be performed in further studies. In addition, $CD8^+$ T cells can be observed in peripheral blood (peripheral $CD8^+$ T cells) and tumor microenvironments ($CD8^+$ tumor-infiltrating lymphocytes). $CD8^+$ tumor-infiltrating lymphocytes ($CD8^+$ TILs) were found to be positively associated with the prognosis of ESCC patients (31, 32). FLVCR1, a major facilitator superfamily metabolite transporter, has been reported to contribute to the development and survival of peripheral $CD4^+$ and $CD8^+$ T cells (33). Studies on the relationship between FLVCR1 and $CD8^+$ TILs are still lacking.

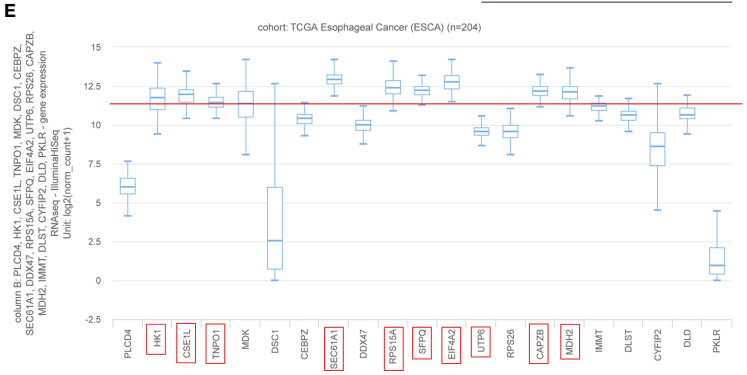


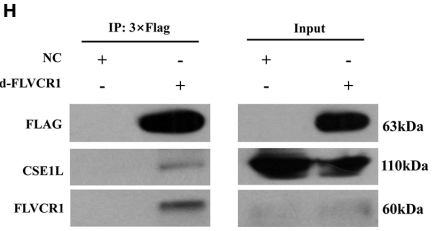


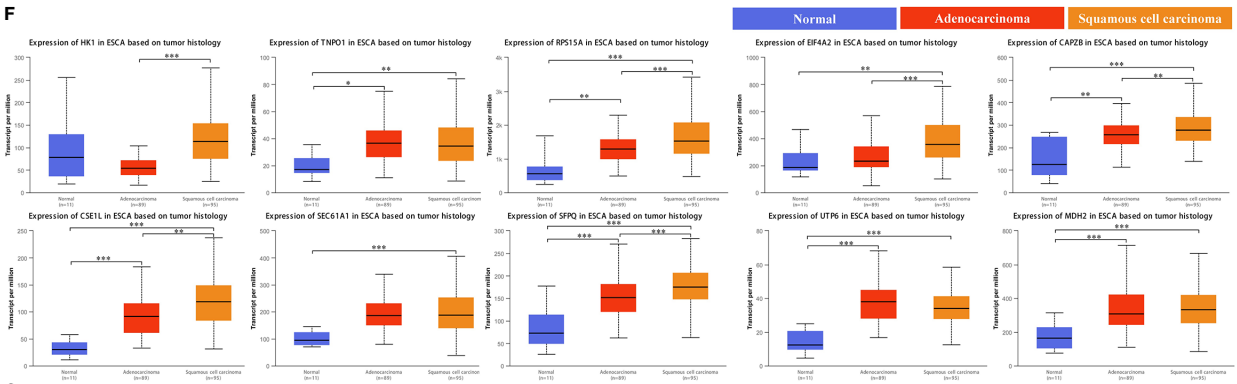
C

Protein Name	Coverage	Peptides	Score Mascot
PLCD4	3.28	1	0
RPS15A	6.92	1	47.64
RPS26	7.83	1	37.80
CEBPZ	1.61	1	0
CYFIP2	0.86	1	48.45
DSC1	1.68	1	0
DLD	2.16	1	43.17
DLST	1.77	1	37.02
EIF4A2	3.44	1	18.77
CSE1L	1.03	1	49.93
CAPZB	3.61	1	32.82
HK1	0.87	1	13.84
LYPLAL1	4.22	1	37.86
MDH2	2.66	1	31.19
MAGEA8	2.83	1	19.61
MT1G	19.35	1	32.74
IMMT	2.77	2	41.59
MDK	13.99	2	52.70
OPALIN	6.38	1	0
ATP4A	1.45	1	34.52
DDX47	2.64	1	33.51
SEC61A1	2.31	1	33.99
ANXA2P2	2.65	1	20.46
PKLR	1.92	1	49.71
SFPQ	0.99	1	43.07
TNPO1	1.22	1	30.02
UTP6	1.17	1	26.85









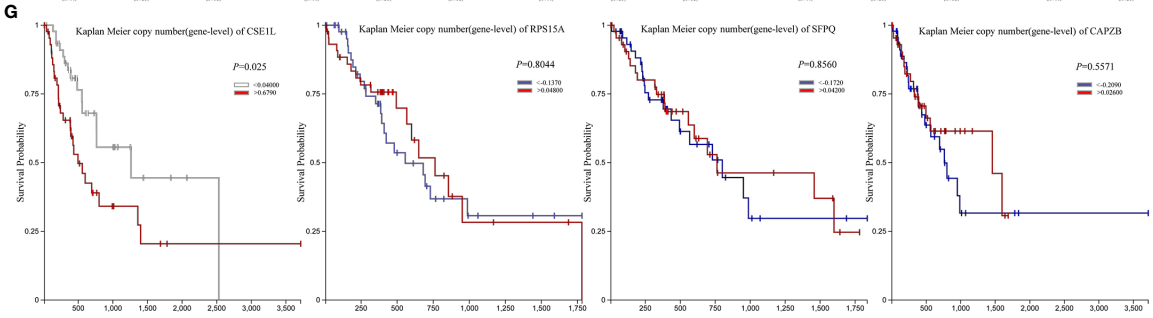


FIGURE 6 | Continued

FIGURE 6 | FLVCR1 interacts with CSE1L. **(A)** Western blotting was used to confirm that FLAG-FLVCR1 was successfully overexpressed in KYSE-150 cells. **(B)** Coomassie blue staining was used to detect immunoprecipitated Flag-FLVCR1-binding proteins. **(C)** MS analysis was used to identify target proteins interacting with FLVCR1, and a list of specific proteins that interact with FLVCR1 protein is shown. **(D)** The gene interaction network diagram demonstrated the interaction among multiple genes surrounding the regulation of FLVCR1. The 21 genes in gray represented the selected genes from MS analysis. Solid lines indicated direct interactions, while broken lines indicated indirect interactions. **(E)** The expression profile of 21 genes in the TCGA esophageal carcinoma cohort. Above the red line are the top-ten highly expressed genes. **(F)** The expression patterns of the top-ten highly expressed genes in the TCGA database according to tumor histology. **(G)** Comparison of overall survival of esophageal cancer patients with different CSE1L/RPS15A/SFPQ/CAPZB expression based on TCGA data. **(H)** The interaction between FLVCR1 and CSE1L identified by mass spectrometry analysis was confirmed by immunoprecipitation in KYSE-150 cells stably expressing FLVCR1. Red arrow: 3×FLAG-target. *P < 0.05, **P < 0.01, ***P < 0.001.

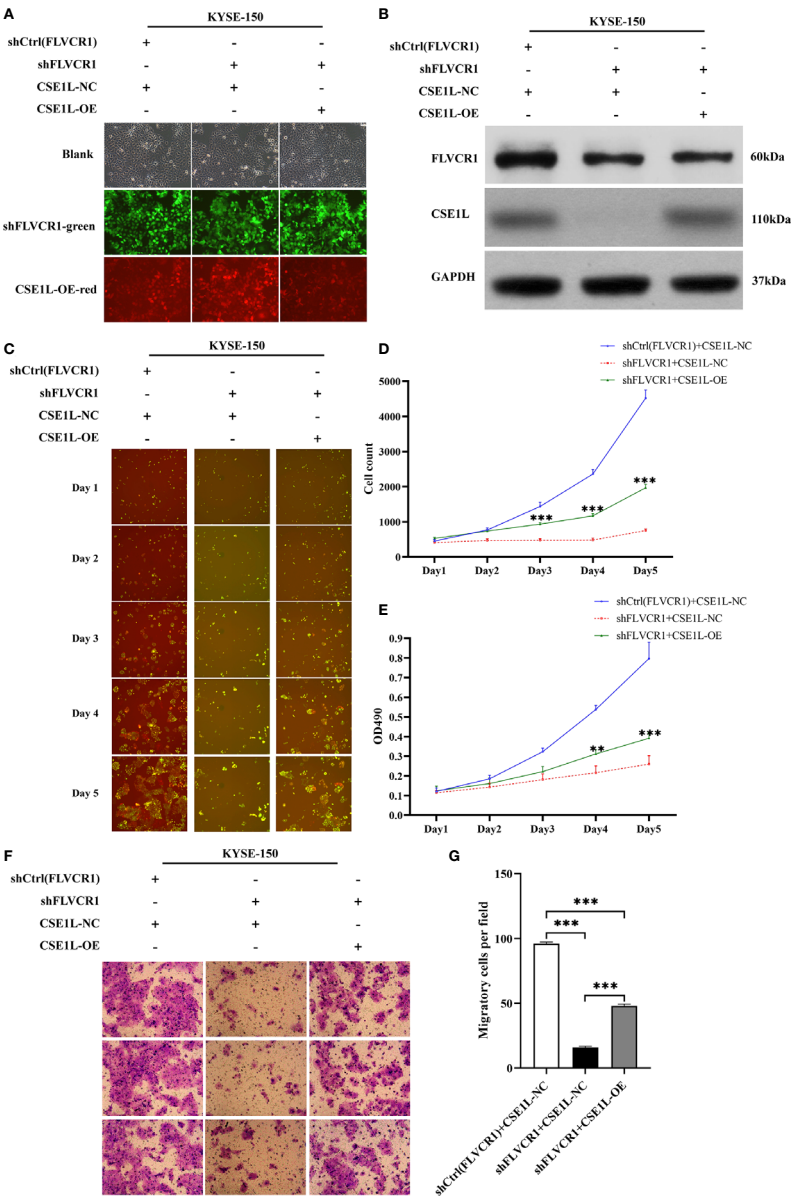


FIGURE 7 | CSE1L re-expression partly counteracted the negative effect of FLVCR1 inhibition on the proliferation and migration of KYSE-150 cells. **(A)** KYSE-150 cells expressed green and red fluorescence after successful infection. Magnification 100×. **(B)**, Western blot analysis showed that the expression of CSE1L was restored in the shFLVCR1+CSE1L-OE group and downregulated in the ShFLVCR1+CSE1L-NC group. **(C–E)** Re-expressed CSE1L partially diminished the proliferation inhibitory effect of FLVCR1 knockdown compared to shFLVCR1+CSE1L-NC group. **(F, G)** Compared to the shFLVCR1+CSE1L-NC group, the migration ability was restored by ectopic expression of CSE1L. Data were represented as the mean ± SD. **P < 0.01, ***P < 0.001.

Therefore, we will focus on the relationship between FLVCR1 and tumor-infiltrating $CD8^+$ T cells or peripheral $CD8^+$ T cells in ESCC in further studies.

In summary, these results indicate that upregulated FLVCR1 has been verified as a poor prognostic indicator in patients with ESCC. Functional experiments confirmed that FLVCR1 may function as an oncogene in ESCC by promoting cell growth, migration, and invasion and repressing apoptosis. We preliminarily confirmed the involvement of CSE1L in the effects of FLVCR1 on the cellular proliferation and migration of ESCC, indicating that FLVCR1-targeting treatment may be promising in the therapies of patients with ESCC.

AUTHOR'S NOTE

Parts of this work have been presented as an abstract at European-Society-for-Medical-Oncology (ESMO) 21st World Congress on Gastrointestinal Cancer and the ESMO Virtual Congress 2020.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Institutional Review Board of Taizhou Hospital. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by Institutional Animal Care and Use Committee (IACUC) of Wenzhou Medical University.

AUTHOR CONTRIBUTIONS

SZ, WY, and HY conceived and participated in the design of the study. The manuscript was written and revised by SZ and MZ. CZ recruited samples. SZ, CZ, and YM performed all the experimental work. SZ, WY, and HY participated in data analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by National Natural Science Foundation of China (NSFC 81872458), Natural Science Foundation of Zhejiang Province (LY19H160017), and Medical Science and Technology Project of Zhejiang Province (2020KY358).

ACKNOWLEDGMENTS

We would like to thank the many clinical doctors and experimenter from the Affiliated Taizhou Hospital, Wenzhou Medical University, who were involved in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.660955/full#supplementary-material

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* (2018) 68:7–30. doi: 10.3322/caac.21442
2. Lin Y, Totsuka Y, He Y, Kikuchi S, Qiao Y, Ueda J, et al. Epidemiology of Esophageal Cancer in Japan and China. *J Epidemiol* (2013) 23:233–42. doi: 10.2188/jea.je20120162
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin* (2016) 66:115–32. doi: 10.3322/caac.21338
4. Zeng H, Zheng R, Guo Y, Zhang S, Zou X, Wang N, et al. Cancer survival in China, 2003-2005: A population-based study. *Int J Cancer* (2015) 136:1921–30. doi: 10.1002/ijc.29227
5. Wei L, Yan N, Sun L, Bao C, Li D. Interplay between the NF−κB and hedgehog signaling pathways predicts prognosis in esophageal squamous cell carcinoma following neoadjuvant chemoradiotherapy. *Int J Mol Med* (2018) 41:2961–67. doi: 10.3892/ijmm.2018.3447
6. Chiabrando D, Vinchi F, Fiorito V, Mercurio S, Tolosano E. Heme in pathophysiology: a matter of scavenging, metabolism and trafficking across cell membranes. *Front Pharmacol* (2014) 5:61. doi: 10.3389/fphar.2014.00061
7. Wiel C, Le Gal K, Ibrahim MX, Jahangir CA, Kashif M, Yao H, et al. BACH1 Stabilization by Antioxidants Stimulates Lung Cancer Metastasis. *Cell* (2019) 178:330–45. doi: 10.1016/j.cell.2019.06.005
8. Sohoni S, Ghosh P, Wang T, Kalainayakan SP, Vidal C, Dey S, et al. Elevated Heme Synthesis and Uptake Underpin Intensified Oxidative Metabolism and Tumorigenic Functions in Non-Small Cell Lung Cancer Cells. *Cancer Res* (2019) 79:2511–25. doi: 10.1158/0008-5472
9. Chiabrando D, Castori M, di Rocco M, Ungelenk M, Gießelmann S, Di Capua M, et al. Mutations in the Heme Exporter FLVCR1 Cause Sensory Neurodegeneration with Loss of Pain Perception. *PloS Genet* (2016) 12: e1006461. doi: 10.1371/journal.pgen.1006461
10. Peng C, Song Y, Chen W, Wang X, Liu X, Wang F, et al. FLVCR1 promotes the proliferation and tumorigenicity of synovial sarcoma through inhibiting apoptosis and autophagy. *Int J Oncol* (2018) 52:1559–68. doi: 10.3892/ijo.2018.4312
11. Shen Y, Li X, Zhao B, Xue Y, Wang S, Chen X, et al. Iron metabolism gene expression and prognostic features of hepatocellular carcinoma. *J Cell Biochem* (2018) 119:9178–204. doi: 10.1002/jcb.27184
12. Mccord RP, Golloshi R. Abstract 5373: 3D genome architecture changes during cancer cell migration and metastasis. *Cancer Res* (2017) 77(13 Supplement):5373–73. doi: 10.1158/1538-7445.AM2017-5373
13. Yang Y, Hu S, Xu X, Li J, Liu A, Han J, et al. The Vascular Endothelial Growth Factors-Expressing Character of Mesenchymal Stem Cells Plays a Positive Role in Treatment of Acute Lung Injury In Vivo. *Mediators Inflamm* (2016) 2016:2347938. doi: 10.1155/2016/2347938
14. Rodriguez LG, Wu X, Guan JL. Wound-healing assay. *Methods Mol Biol* (2005) 294:23–9. doi: 10.1385/1-59259-860-9:023
15. Justus CR, Leffler N, Ruiz-Echevarria M, Yang LV. In vitro cell migration and invasion assays. *J Vis Exp* (2014) 1:51046. doi: 10.3791/51046
16. Jiang MC. CAS (CSE1L) signaling pathway in tumor progression and its potential as a biomarker and target for targeted therapy. *Tumour Biol* (2016) 37:13077–90. doi: 10.1007/s13277-016-5301-x

17. Alnabulsi A, Agouni A, Mitra S, Garcia-Murillas I, Carpenter B, Bird S, et al. Cellular apoptosis susceptibility (chromosome segregation 1-like, CSE1L) gene is a key regulator of apoptosis, migration and invasion in colorectal cancer. *J Pathol* (2012) 228(4):471–81. doi: 10.1002/path.4031
18. Li Y, Yuan S, Liu J, Wang Y, Zhang Y, Chen X, et al. CSE1L silence inhibits the growth and metastasis in gastric cancer by repressing GPNMB via positively regulating transcription factor MITF. *J Cell Physiol* (2020) 235:2071–79. doi: 10.1002/jcp.29107
19. Jiang K, Neill K, Cowden D, Klapman J, Eschrich S, Pimiento J, et al. Expression of CAS/CSE1L, the Cellular Apoptosis Susceptibility Protein, Correlates With Neoplastic Progression in Barrett's Esophagus. *Appl Immunohistochem Mol Morphol* (2018) 26:552–6. doi: 10.1097/PAI.0000000000000464
20. Tai CJ, Liao CF, Su TC, Shen KH, Chang CC, Lin SH, et al. Presence of CSE1L protein in urine of patients with urinary bladder urothelial carcinomas. *Int J Biol Markers* (2012) 27:e280–4. doi: 10.5301/JBM.2012.9310
21. Liao CF, Lin SH, Chen HC, Tai CJ, Chang CC, Li LT, et al. CSE1L, a novel microvesicle membrane protein, mediates Ras-triggered microvesicle generation and metastasis of tumor cells. *Mol Med* (2012) 18:1269–80. doi: 10.2119/molmed.2012.00205
22. Lee WR, Shen SC, Shih YH, Chou CL, Tseng JT, Chin SY, et al. Early decline in serum phospho-CSE1L levels in vemurafenib/sunitinib-treated melanoma and sorafenib/lapatinib-treated colorectal tumor xenografts. *J Transl Med* (2015) 13:191. doi: 10.1186/s12967-015-0553-6
23. Peng Y, Wang L, Gu J. Elevated preoperative carcinoembryonic antigen (CEA) and Ki67 is predictor of decreased survival in IIA stage colon cancer. *World J Surg* (2013) 37:208–13. doi: 10.1007/s00268-012-1814-7
24. Li H, Han X, Liu Y, Liu G, Dong G. Ki67 as a predictor of poor prognosis in patients with triple-negative breast cancer. *Oncol Lett* (2015) 9:149–52. doi: 10.3892/ol.2014.2618
25. Turkel Kucukmetin N, Cicek B, Saruc M, Ersoy O, Vardareli E, Onder O, et al. Ki67 as a prognostic factor for long-term outcome following surgery in gastrointestinal stromal tumors. *Eur J Gastroenterol Hepatol* (2015) 27:1276–80. doi: 10.1097/MEG.0000000000000454
26. Liu X, Zhang M, Ying S, Zhang C, Lin R, Zheng J, et al. Genetic alterations in esophageal tissues from squamous dysplasia to carcinoma. *Gastroenterology* (2017) 153:166–77. doi: 10.1053/j.gastro.2017.03.033
27. Kandioler D, Schoppmann SF, Zwrtek R, Kappel S, Wolf B, Mittlböck M, et al. The biomarker TP53 divides patients with neoadjuvantly treated esophageal cancer into 2 subgroups with markedly different outcomes. A p53 Research Group study. *J Thorac Cardiovasc Surg* (2014) 148:2280–6. doi: 10.1016/j.jtcvs.2014.06.079
28. Zhao Z, Wang P, Gao Y, He J. The high expression instead of mutation of p53 is predictive of overall survival in patients with esophageal squamous-cell carcinoma: a meta-analysis. *Cancer Med* (2017) 6:54–66. doi: 10.1002/cam4.945
29. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* (2012) 366:2455–65. doi: 10.1056/NEJMoa1200694
30. Leng C, Li Y, Qin J, Ma J, Liu X, Cui Y, et al. Relationship between expression of PD-L1 and PD-L2 on esophageal squamous cell carcinoma and the antitumor effects of $CD8^+$ T cells. *Oncol Rep* (2016) 35:699–708. doi: 10.3892/or.2015.4435
31. Han L, Gao QL, Zhou XM, Shi C, Chen GY, Song YP, et al. Characterization of CD103(+) CD8(+) tissue-resident T cells in esophageal squamous cell carcinoma: may be tumor reactive and resurrected by anti-PD-1 blockade. *Cancer Immunol Immunother* (2020) 69:1493–504. doi: 10.1007/s00262-020-02562-3
32. Nishimura J, Tanaka H, Yamakoshi Y, Hiramatsu S, Tamura T, Toyokawa T, et al. Impact of tumor-infiltrating LAMP-3 dendritic cells on the prognosis of esophageal squamous cell carcinoma. *Esophagus* (2019) 16:333–44. doi: 10.1007/s10388-019-00669-w
33. Philip M, Funkhouser SA, Chiu EY, Phelps SR, Delrow JJ, Cox J, et al. Heme exporter FLVCR is required for T cell development and peripheral survival. *J Immunol* (2015) 194:1677–85. doi: 10.4049/jimmunol.1402172

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 28 May 2021
doi: 10.3389/fonc.2021.665388

Comprehensive Analysis of GLUT1 Immune Infiltrates and ceRNA Network in Human Esophageal Carcinoma

Xu-Sheng Liu[1†], Yan Gao[1†], Li-Bing Wu[1†], Hua-Bing Wan[1], Peng Yan[1], Yang Jin[1], Shi-Bo Guo[1], Ya-Lan Wang[1], Xue-Qin Chen[2], Lu-Meng Zhou[1], Jian-Wei Yang[1], Xue-Yan Kui[1], Xiao-Yu Liu[1] and Zhi-Jun Pei[1,3]*

[1] Department of Nuclear Medicine and Institute of Anesthesiology and Pain, Taihe Hospital, Hubei University of Medicine, Shiyan, China, [2] Hubei University of Medicine, Shiyan, China, [3] Hubei Key Laboratory of Embryonic Stem Cell Research, Shiyan, China

OPEN ACCESS

Edited by:
Bin Qiao,
Zhengzhou University, China

Reviewed by:
Simona D'Aguanno,
Istituti Fisioterapici Ospitalieri (IRCCS), Italy
Junwen Huang,
First Affiliated Hospital of Zhengzhou University, China

****Correspondence:***
Zhi-Jun Pei
pzjun1980@yeah.net

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *10 February 2021*
Accepted: *05 May 2021*
Published: *28 May 2021*

Citation:
Liu X-S, Gao Y, Wu L-B, Wan H-B, Yan P, Jin Y, Guo S-B, Wang Y-L, Chen X-Q, Zhou L-M, Yang J-W, Kui X-Y, Liu X-Y and Pei Z-J (2021) Comprehensive Analysis of GLUT1 Immune Infiltrates and ceRNA Network in Human Esophageal Carcinoma. Front. Oncol. 11:665388. doi: 10.3389/fonc.2021.665388

Background: Glucose transporter 1 (GLUT1) is encoded by the solute carrier family 2A1 (SLC2A1) gene and is one of the glucose transporters with the greatest affinity for glucose. Abnormal expression of GLUT1 is associated with a variety of cancers. However, the biological role of GLUT1 in esophageal carcinoma (ESCA) remains to be determined.

Methods: We analyzed the expression of GLUT1 in pan-cancer and ESCA as well as clinicopathological analysis through multiple databases. Use R and STRING to perform GO/KEGG function enrichment and PPI analysis for GLUT1 co-expression. TIMER and CIBERSORT were used to analyze the relationship between GLUT1 expression and immune infiltration in ESCA. The TCGA ESCA cohort was used to analyze the relationship between GLUT1 expression and m6A modification in ESCA, and to construct a regulatory network in line with the ceRNA hypothesis.

Results: GLUT1 is highly expressed in a variety of tumors including ESCA, and is closely related to histological types and histological grade. GO/KEGG functional enrichment analysis revealed that GLUT1 is closely related to structural constituent of cytoskeleton, intermediate filament binding, cell-cell adheres junction, epidermis development, and P53 signaling pathway. PPI shows that GLUT1 is closely related to TP53, GIPC1 and INS, and these three proteins all play an important role in tumor proliferation. CIBERSORT analysis showed that GLUT1 expression is related to the infiltration of multiple immune cells. When GLUT1 is highly expressed, the number of memory B cells decreases. ESCA cohort analysis found that GLUT1 expression was related to 7 m6A modifier genes. Six possible crRNA networks in ESCA were constructed by correlation analysis, and all these ceRNA networks contained GLUT1.

Conclusion: GLUT1 can be used as a biomarker for the diagnosis and treatment of ESCA, and is related to tumor immune infiltration, m6A modification and ceRNA network.

Keywords: GLUT1, esophageal carcinoma, immune infiltration, m6A modification, ceRNA

INTRODUCTION

Esophageal carcinoma (ESCA) is one of the most common malignancies of the upper digestive tract worldwide, and esophageal squamous cell carcinoma (ESCC) is the most common pathological subtype of ESCA. ESCA has the characteristics of high malignancy, poor prognosis, and high mortality, which has seriously threatened human life and health. Although some progress has been made in the diagnosis and treatment of esophageal diseases, the prognosis of patients with middle and advanced ESCA is still extremely poor, with a 5-year survival rate of only 15%-20%. The molecular mechanism of the formation and progression of ESCA is an extremely complex process involving cell cycle regulation and signal transduction, which makes effective treatment of ESCA more difficult. Therefore, further studying the pathogenesis of ESCA and providing new molecular targets for the early diagnosis and treatment of tumors has more important practical significance and theoretical value.

Glucose transporter 1 (GLUT1) is the first glucose transporter discovered and one of the glucose transporters with the greatest affinity for glucose (1). GLUT1 is encoded by the solute carrier family 2A1 (SLC2A1) gene (2), and its crystal structure was analyzed for the first time in 2014 (3). GLUT1 is widely present in most tissues of the human body, but in normal tissues and benign lesions, the expression level of GLUT1 is low, and high expression is often related to cancer, which may indicate a poor prognosis or recurrence (4). GLUT1 not only maintains the normal functions of the human body, but also plays an important role in the occurrence and development of tumors, especially in the glycolysis process of tumor cells (5). Tumor cells need to consume a lot of glucose when they grow and proliferate. However, GLUT1 is one of the key proteins that transport glucose into cells. Currently, studies have shown that GLUT1 is highly expressed in oral (4), gastric (5), breast (6), colorectal (7), and ovarian cancers (8). In previous studies (9), we found that GLUT1 was highly expressed in ESCA and strongly correlated with metabolic parameters of ^{18}F-FDG PET/CT imaging, but we failed to carry out more studies on the biological function of GLUT1 in ESCA.

Tumor immunotherapy, N6-methyladenosine (m6A) and ceRNA regulatory network are the new directions of tumor gene therapy and are widely used in the study of the mechanism of ESCA. Na et al. (10) found that GLUT1 is highly expressed in lung adenocarcinoma, and the expression level is negatively correlated with immune score. ^{18}F-FDG uptake in immunocompromised lung adenocarcinoma patients is positively correlated with GLUT1 expression, but negatively correlated with immune score. This result indicates that the competitive uptake of glucose by cancer cells and immune cells in the tumor microenvironment may be caused by the different expression levels of GLUT1 in the cells. Chen et al (11). found that the key gene of m6A, METTL3, induced GLUT1 translation in an m6a-dependent manner, which enhanced the glucose uptake and lactate production of colorectal cancer cells, which in turn led to the activation of mTORC1 signaling and the development of colorectal cancer. Chen et al. (12) found that the CircRNA_100290/miR-378a/GLUT1 ceRNA regulatory network plays an important role in oral squamous cell carcinoma. Overexpression of GLUT1 can rescue the reduction of tumor cell proliferation and glycolysis caused by down-regulation of circRNA. However, there is little research on the overall understanding of GLUT1 in ESCA, especially the relationship between GLUT1 and tumor immunotherapy, m6A modification and ceRNA regulatory network.

In this study, we analyzed the differences in GLUT1 expression in different cancers by analyzing The Cancer Genome Atlas (TCGA) and various public databases. Using multi-dimensional analysis to evaluate the gene and functional network related to the expression of GLUT1 in ESCA, and to explore the relationship between its expression differences and tumor immunity, m6A modification, and ceRNA regulatory network, provide a theoretical basis for discovering possible molecular pathways.

MATERIALS AND METHODS

Ethics Statement

This study proposal has been approved by the Ethics Committee of Taihe Hospital Affiliated of Hubei University of Medicine (Shiyan, China) and conducted in accordance with the research principles described in the Helsinki Declaration.

Oncomine Analysis

Oncomine (www.oncomine.org) is a publicly accessible database of oncogene chip information used to analyze the transcription level of GLUT1 in various cancers (13). This database uses Student's t test to compare the transcription levels of GLUT1 in normal controls and clinical cancer specimens. In this study, the fold change > 2 and cut-off of P-value < 0.0001.

TIMER Analysis

Tumor immune to assess resource (TIMER, www.cistrome.shinyapps.io) is a reliable and convenient database, including gene expression profiles from the TCGA database. The TIMER tool can be used to estimate immune cell infiltration and evaluate its clinical impact (14, 15).

In this study, we used TIMER to evaluate the transcription level of GLUT1 in a variety of tumors and analyzed the correlation between GLUT1 transcription level and immune cell infiltration, including B cells, neutrophils, CD4 + T cells, macrophages, CD8 + T cells, and dendritic cells, as well as the tumor purity. We used the TIMER tool to analyze the correlation between GLUT1 and immune cell markers to evaluate the role of GLUT1 in tumor immunity. Immune cell gene markers are selected from the website of R&D Systems (www.rndsystems.com/cn/resources/cell-markers/immune-cells). These gene markers include markers of B cells, CD8 + T cells, follicular helper T cells (Tfh), T-helper 1 (Th1) cells, T-helper 2 (Th2) cells, T-helper 17 (Th17) cells, Treg, T cells exhausted, macrophages, M1 macrophages, M2 macrophages, tumor-associated macrophages (TAM), monocytes, natural killer (NK) cells, neutrophils, and dendritic cells (DC). In addition,

we used the somatic copy number alteration (SCNA) module of the TIMER tool to link the genetic copy number variations (CNV) of GLUT1 with the relative abundance of tumor infiltrating cells.

TCGA Data

TCGA (www.tcga-data.nci.nih.gov/tcga/) contains more than 10,000 samples of 39 tumor types (16). We downloaded ESCA RNA-seq data from the Genomic Data Commons (GDC, https://portal.gdc.cancer.gov/) database, which included 162 tumor samples and 11 normal samples. In this study, we used TCGA-ESCA data to analyze the expression of GLUT1 and the correlation between GLUT1 expression and clinicopathological characteristics. In addition, we also analyzed the correlation between the expression level of GLUT1 and the expression of m6A-related genes in ESCA samples and the differences expression in m6A-related gene between the high and low GLUT1 expression groups. m6A related genes include METTL3, YTHDC1, YTHDC2, METTL14, RBM15, RBM15B, IGF2BP1, IGF2BP2, IGF2BP3, VIRMA, WTAP, YTHDF1, YTHDF2, YTHDF3, ZC3H13, HNRNPA2B1, HNRNPC, RBMX, FTO and ALKBH5 (17).

GEO Data

We downloaded the RNA sequencing data of ESCA from the Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo) database to analyze the transcription level of GLUT1. (GSE38129, n=60; GSE23400, n=106).

Cell Lines and Cell Culture Reagents

Human ESCA cell lines ECA109 and KYSE-150 and normal human squamous esophageal cell line Het-1A were obtained from the American Type Culture Collection (Manassas, VA, USA). The cells were maintained in DMEM high glucose medium (Hyclone, Logan, UT, USA) supplemented with 10% FBS (Gibco, USA) and 1% antibiotics (penicillin-streptomycin, Gibco, USA).

RNA Extraction and qRT-PCR

The implementation method refers to previous study (18). Total RNA was isolated from cells using Trizol reagent (Invitrogen, Carlsbad, CA, USA). Use Prime Script RT reagent kit (Takara, Dalian, China) for reverse transcription, and then use SYBR Prime Script RT PCR kit (Takara, Dalian, China) for qRT-PCR. Use GAPDH as an internal reference and use the $2^{-\triangle\triangle Ct}$ method to calculate the results. GLUT1 primer sequences: forward primer CTTTGTGGCCTTCTTTGAAGT and reverse primer CCACACAGTTGCTCCACAT. GAPDH primer sequences: forward primer GGAGCGAGATCCCTCCAAAAT and reverse primer GGCTGTTGTCATACTTCTCATGG.

Immunohistochemistry

Clinical samples were obtained from 50 patients with ESCA who were surgically treated at Taihe Hospital Affiliated of Hubei University of Medicine from February 2016 to September 2017. The content of GLUT1 was detected by IHC according to the method previously described (9). The ESCA tissue and the paracarcinoma tissues were prepared into 3 μm paraffin sections and incubated with mouse monoclonal antibodies of GLUT1 (1:200, Abcam, USA) at 4°C overnight in a refrigerator. The sections were coupled with the goat anti-mouse IgG-HRP secondary antibody (1:2000, Abcam, USA) at room temperature for 1.5 h, then each incubated section was stained with DAB reagent, and finally counterstained with hematoxylin.

LinkedOmics Analysis

The LinkedOmics database (http://www.linkedomics.org/login.php) is a web-based platform that can provide comprehensive multi-omics data analysis tools for the TCGA database (19). The Pearson correlation coefficient was used for statistical analysis of GLUT1 co-expression and displayed in the form of volcano map and heat map. The rank criterion was an FDR<0.05.

R Software

The ClusterProfiler package (version: 3.18.0) of R was employed to analyze the GO function and KEGG pathway enrichment of potential targets. Use the ggplot2 software package to visualize the analysis data. The CIBERSORT package was used to evaluate the relative proportion of 22 immune infiltrating cells in tumor samples when GLUT1 was high or low expression.

STRINGS Analysis

STRINGS (www.string-db.org) is an online analysis website that contains all publicly available protein-protein interaction (PPI) data. In this study (20), we used STRING to perform PPI network analysis on GLUT1.

Prediction of miRNA

Use starBase3.0 (www.starbase.sysu.edu.cn) online website to predict the target miRNA of GLUT1, and the prediction results include the analysis of PITA, miRanda and TargetScan (21). In addition, we analyze the correlation between target miRNA expression and GLUT1 expression to screen for miRNAs that are more in line with ceRNA conditions. Finally, use TargetScan (http://www.targetscan.org/vert_72) online tool to predict the potential binding site of target miRNA and GLUT1 (22).

Prediction of lncRNA and ceRNA Network Construction

Use miRNet2.0 (www.mirnet.ca/miRNet/home.xhtml) (23) and starBase to predict the target lncRNA of miRNA, miRNA includes has-miR-140-5p, has-miR-148a-3p and has-miR-148b-3p. In addition, we analyze the correlation between target miRNA expression and lncRNA expression to screen for lncRNAs that are more in line with ceRNA conditions. Comprehensive analysis of miRNA-mRNA and miRNA-lncRNA with negative correlation between expression levels to establish a key lncRNA-miRNA-mRNA (GLUT1) ceRNA network for ESCA.

Statistical Analysis

Unpaired samples used unpaired t-test and paired samples used paired t-test. Multi-group cell experiments use One-way ANOVA. The correlation of gene expression was evaluated

using Spearman's correlation. The threshold of $P < 0.05$ indicates the significance of correlation.

RESULTS

Pan-Cancer Analysis of GLUT1 mRNA Expression in Different Databases

To determine the difference in the expression of GLUT1 in tumors and normal tissues, the Oncomine database was used to analyze the levels of GLUT1 mRNA in tumors and normal tissues of various cancer types. This analysis showed that bladder (24, 25), breast (26–28), colorectal (29, 30), esophageal (31, 32), gastric (33), head and neck (34, 35), kidney (36–39), leukemia (40), lung (41–46), lymphoma (47), ovarian (48) and pancreatic cancer (49–52) have higher expression of GLUT1 compared with normal tissues. In addition, there are data showing that the expression is lower in breast (53), esophageal cancer (54) and leukemia tumors (55) (**Figure 1A**). **Supplementary Table 1** summarizes the details of GLUT1 expression in various cancers.

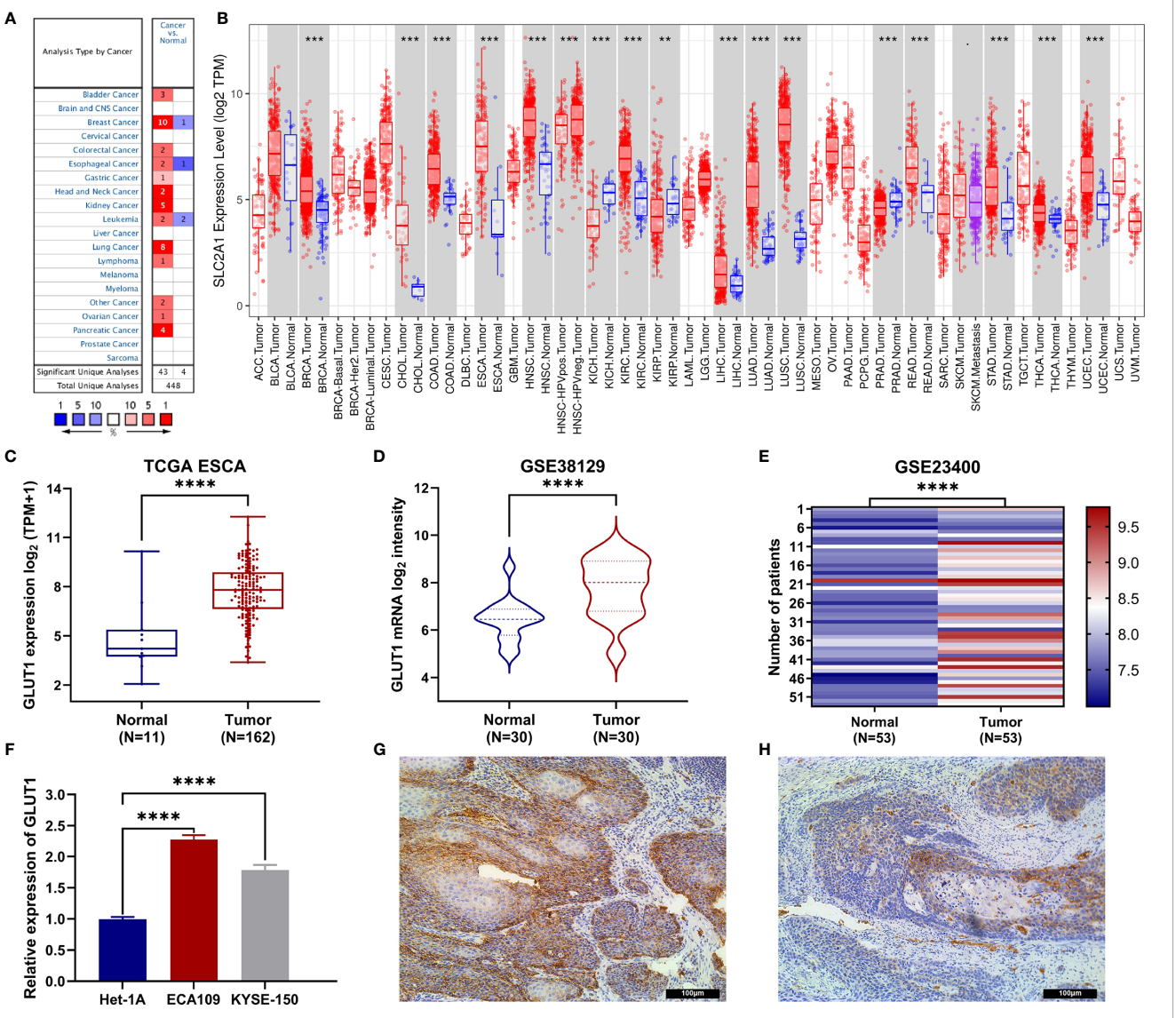


FIGURE 1 | The expression of GLUT1 in esophageal carcinoma (ESCA) and pan-carcinoma. **(A)** The Oncomine database shows that GLUT1 is up-regulated in multiple tumor types. **(B)** GLUT1 expression levels in different tumor types were measured using TIMER. **(C)** TCGA cohort analysis of the expression level of GLUT1 between ESCA and normal tissues. **(D)** The GSE38129 data set was used to analyze the expression level of GLUT1 between ESCA and normal tissues. **(E)** The GSE23400 data set was used to analyze the expression level of GLUT1 between ESCA and paired normal adjacent tissues. **(F)** The expression of GLUT1 in human esophageal carcinoma ECA109 cell line, KYSE-150 cell line and human normal esophageal epithelial cells Het-1A. Immunohistochemistry assay was used to analyze the expression of GLUT1 in ESCA tissues **(G)**, in paracarcinoma tissues **(H)**. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

To further evaluate the expression of GLUT1 in human cancers, we used the TIMER database for analysis. The differential expression of GLUT1 in different tumors and adjacent normal tissues is shown in **Figure 1B**. Compared with adjacent normal tissues, BRCA (breast invasive carcinoma), CHOL (cholangiocarcinoma), COAD (colon adenocarcinoma), ESCA (esophageal carcinoma), HNSC (head and neck squamous cell carcinoma), KIRC (kidney renal clear cell carcinoma), KIRP (kidney renal papillary cell carcinoma), LIHC (liver hepatocellular carcinoma), LUAD (lung adenocarcinoma), LUSC (lung squamous cell carcinoma), READ (rectal adenocarcinoma), STAD (gastric adenocarcinoma), THCA (thyroid cancer) and UCEC (endometrial cancer) expression were significantly increased expression. However, compared with adjacent normal tissues, KICH (kidney chromosome) and PRAD (prostate Adenocarcinoma) have significantly lower GLUT1 expression than normal tissues.

Transcriptional Levels of GLUT1 in Patients With ESCA

To further evaluate the expression of GLUT1 in ESCA, we used TCGA RNA sequencing data and GEO dataset for analysis and found that the level of GLUT1 mRNA was significantly increased in cancer tissues (**Figure 1C**). In unpaired or paired data sets, the expression of GLUT1 mRNA in the tumor group was significantly higher than that in normal tissues (**Figures 1D, E**). To verify the accuracy of data analysis, we also used qRT-PCR and IHC to detect the expression of GLUT1 mRNA and protein in ESCA cells. The results of qRT-PCR showed that the expression of GLUT1 mRNA in ESCA ECA109 and KYSE-150 cell lines was significantly higher than that in human normal esophageal epithelial cells Het-1A cell line (**Figure 1F**). IHC results showed that the protein level of GLUT1 in tumor tissues was significantly higher than that in adjacent normal tissues (**Figures 1G, H**). These results indicate that GLUT1 has a potential carcinogenic effect on the progression of ESCA.

Relationship Between GLUT1 mRNA Expression and Clinicopathological Parameters in Patients with ESCA

To better understand the relevance of GLUT1 expression in cancer, we used the TCGA cohort to analyze its underlying mechanism and correlate it with certain clinical aspects. Chi-square test was performed on samples of ESCA with qualified clinical information, and it was found that the high expression of GLUT1 was significantly correlated with tissue type ($P < 0.001$) and histological grade ($P = 0.006$) (**Table 1**).

Enrichment Analysis of GLUT1 Gene Co-Expression Network and PPI Analysis in ESCA

To further understand the biological significance of GLUT1 in ESCA, we used the LinkedOmics database to analyze the GLUT1 co-expression in ESCA. As shown in **Figure 2A**, 4300 genes are positively correlated with GLUT1, and 6056 genes are significantly negatively correlated with GLUT1 (FDR < 0.05). The heat map shows the top 50 significant genes that are positively correlated (**Figure 2B**) and negatively correlated with GLUT1 (**Figure 2C**), respectively. See **Supplementary Table 2** for detailed descriptions of related genes.

TABLE 1 | Correlation of GLUT1 mRNA expression with clinicopathological features in the TCGA cohort.

Characteristic	levels	GLUT1 expression		*P*
		Low (%)	High (%)	
Age	<=60	36 (22.2%)	47 (29%)	0.116
	>60	45 (27.8%)	34 (21%)	
Gender	Female	11 (6.8%)	12 (7.4%)	1.000
	Male	70 (43.2%)	69 (42.6%)	
T stage	T1	16 (11%)	11 (7.6%)	0.085
	T2	12 (8.3%)	25 (17.2%)	
	T3	39 (26.9%)	38 (26.2%)	
	T4	3 (2.1%)	1 (0.7%)	
N stage	N0	27 (18.8%)	39 (27.1%)	0.291
	N1	35 (24.3%)	28 (19.4%)	
	N2	4 (2.8%)	5 (3.5%)	
	N3	4 (2.8%)	2 (1.4%)	
M stage	M0	58 (45%)	63 (48.8%)	0.486
	M1	5 (3.9%)	3 (2.3%)	
Pathologic stage	Stage I	8 (5.6%)	8 (5.6%)	0.064
	Stage II	26 (18.3%)	43 (30.3%)	
	Stage III	30 (21.1%)	19 (13.4%)	
	Stage IV	5 (3.5%)	3 (2.1%)	
Histological type	Adenocarcinoma	61 (37.7%)	19 (11.7%)	**<0.001**
	Squamous Cell Carcinoma	20 (12.3%)	62 (38.3%)	
Histologic grade	G1	2 (1.6%)	14 (11.1%)	**0.006**
	G2	33 (26.2%)	33 (26.2%)	
	G3	26 (20.6%)	18 (14.3%)	

Bold values indicate P < 0.05.

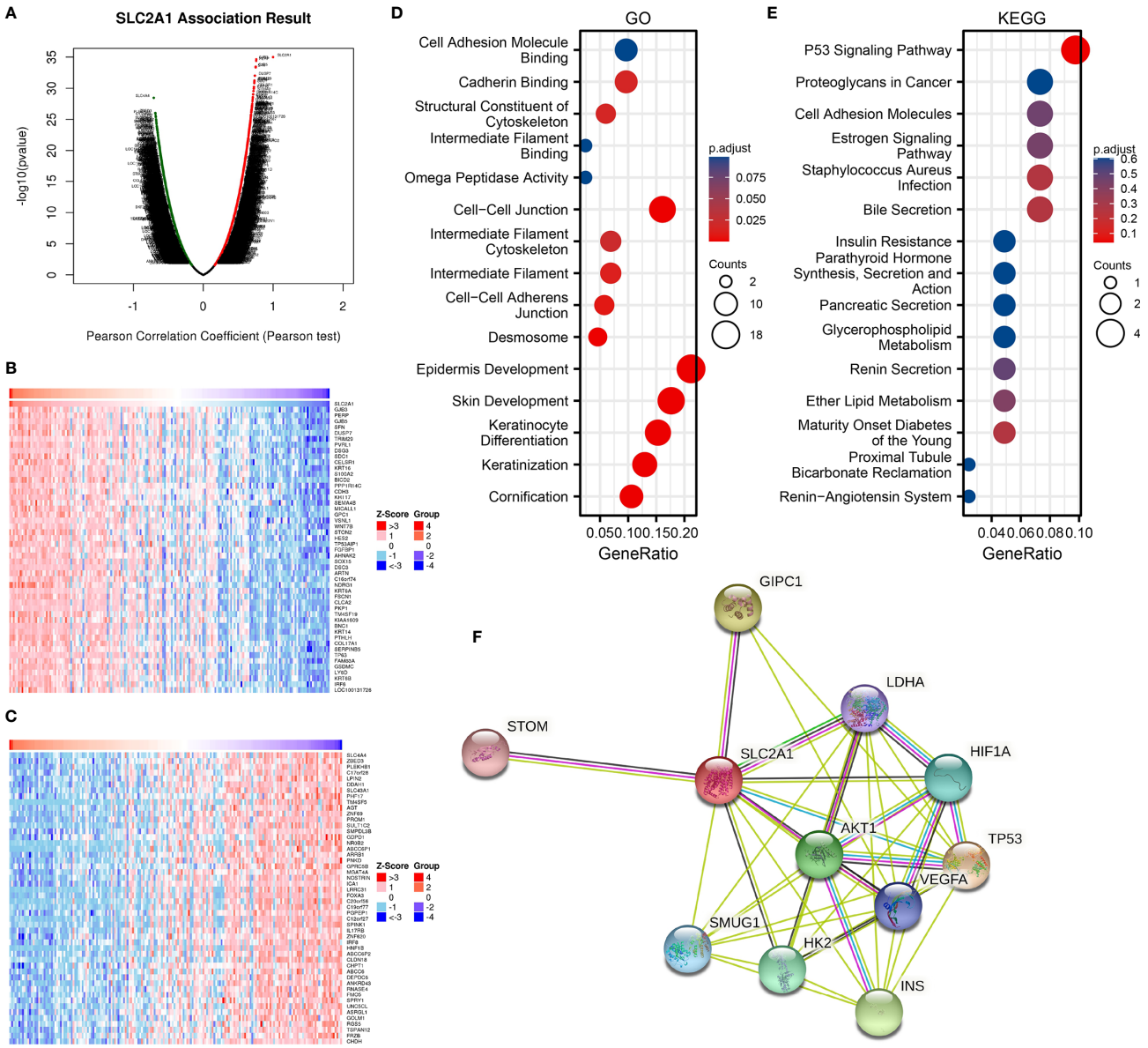


FIGURE 2 | Enrichment analysis of GLUT1 functional networks in esophageal carcinoma (ESCA). **(A)** Genes highly related to GLUT1 identified in the ESCA cohort by Pearson test. **(B)** The heat map shows the top 50 genes positively related to GLUT1 in the ESCA cohort. **(C)** The heat map shows the top 50 genes negatively related to GLUT1 in the ESCA cohort. **(D)** Enrichment of gene ontology (GO) terms for genes related to GLUT1. **(E)** Enrichment of Kyoto Encyclopedia of Genes and Genomes (KEGG) terms for genes related to GLUT1. **(F)** Protein–protein interaction network of GLUT1.

We use R software package to perform Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of GLUT1 related genes. Under the condition of p.adj < 0.1, there are 84 biological process (GO-BP), 8 cellular component (GO-CC), 5 biological process (GO-MF), and 1 KEGG. The bubble chart shows the first 15 pieces of information about GO and KEGG, including 5 pieces of BP, CC, and MF. GO function annotation shows that GLUT1 co-expression are mainly involved in structural constituent of cytoskeleton, intermediate filament binding, cell-cell adheres junction, epidermis development (**Figure 2D**). KEGG pathway analysis showed that GLUT1 co-expression are mainly related to the P53 signaling pathway (**Figure 2E**). **Supplementary Table 3** summarizes the GO and KEGG enrichment analysis details of GLUT1 co-expression.

To further understand the potential mechanism of GLUT1, the STRING database was used to study the PPI network of GLUT1. The analysis showed that GLUT1 is associated with Cellular tumor antigen p53 (TP53), PDZ domain-containing protein GIPC1 (GIPC1), Insulin (INS), RAC-alpha serine

(AKT1), Hexokinase-2 (HK2), Hypoxia-inducible factor 1-alpha (HIF1A), Single-strand selective monofunctional uracil DNA glycosylase (SMUG1), Vascular endothelial growth factor A (VEGFA), Lactate dehydrogenase A (LDHA) and Erythrocyte band 7 integral membrane protein (STOM) were 0.95, 0.893, 0.873, 0.872, 0.867, 0.852, 0.831, 0.831, 0.818 and 0.817 (**Figure 2F** and **Supplementary Table 4**). TP53 plays an inhibitory role in most tumors, and GIPC1 and INS play a promoting role in some tumors. We found that the above three proteins have the highest correlation with GLUT1. These results may indicate that GLUT1 is closely related to the occurrence and development of ESCA.

GLUT1 Expression Is Associated With Immune Signatures in ESCA

Studies have shown that tumor-infiltrating lymphocytes can be used as an independent predictor of the status and prognosis of cancer sentinel lymph nodes (56). Therefore, we used TIMER to analyze whether the expression of GLUT1 is related to the level of immune infiltration in ESCA. As shown in **Figure 3A**, GLUT1 expression showed a negative correlation with the levels of B cells ($P = 8.65\times10^{-6}$), CD4 + T cells ($P = 5.02\times10^{-3}$), Macrophages ($P = 2.53\times10^{-3}$) and Dendritic cells ($P = 3.81\times10^{-2}$). These results indicate that GLUT1 plays a key role in the immune infiltration of ESCA. In addition, it was also found that GLUT1 CNV has a significant correlation with the infiltration level of CD4 + T cells, Neutrophils and Dendritic Cells (**Figure 3B**).

To study the relationship between GLUT1 and various immune-infiltrating cells in ESCA, the TIMER tool was used to analyze the correlation between GLUT1 and immune markers of various immune cells in ESCA (**Table 2**). The results showed that the expression of GLUT1 was significantly correlated with the immune markers CD20 and CD19 of B cells in ESCA ($P < 0.05$, **Table 2**). We also analyzed a variety of T cells with different functions, such as CD8 + T cells, Tfh cells, Th1 cells, Th2 cells, Th17 cells, Treg, and exhausted T cells. The results after adjustment of tumor purity showed that the expression level of GLUT1 was significantly correlated with most of the immune

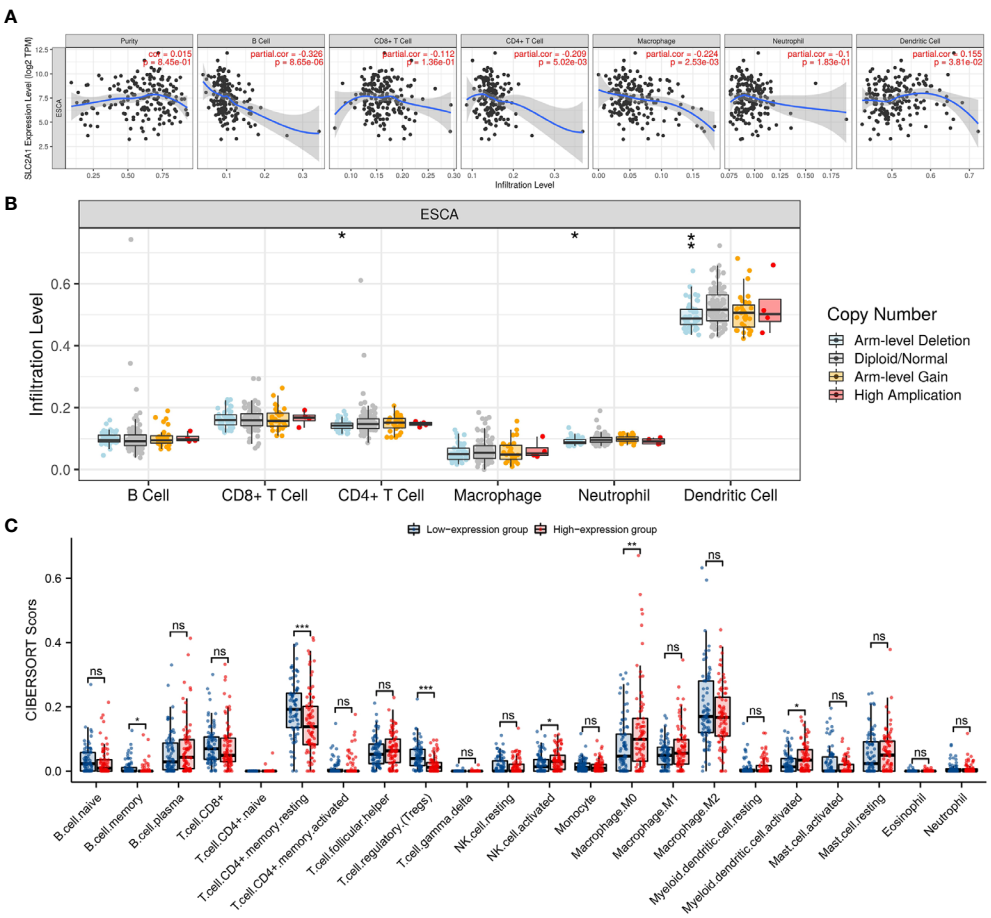


FIGURE 3 | Correlations of GLUT1 expression with immune infiltration level in esophageal carcinoma (ESCA). **(A)** The expression of GLUT1 was significantly correlated with infiltrating levels of B cell, CD4+T cells, macrophages, and dendritic cells in ESCA. **(B)** GLUT1 CNV affects the infiltrating levels of CD4+T cells, neutrophils, and dendritic cells in ESCA. **(C)** The change ratio of 22 immune cell subtypes in the high and low GLUT1 expression groups in ESCA tumor samples. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$. ns, not significant.

TABLE 2 | Correlation analysis between GLUT1 and relate genes and markers of immune cells in TIMER.

Gene markers	Gene markers	rho	p	adj.p
B cell	CD19	-0.29406	**6.16E-05***	**4.83E-04***
	CD20	-0.32504	**8.51E-06***	**8.10E-05***
	CD70	0.019655	7.93E-01	8.78E-01
CD8+ T Cell	CD8A	-0.17305	**2.02E-02***	5.51E-02
	CD8B	-0.25575	**5.30E-04***	**2.90E-03***
	CD25	-0.11128	1.37E-01	2.48E-01
Tfh	CD183	-0.35906	**7.43E-07***	**1.02E-05***
	CD185	-0.20494	**5.78E-03***	**2.07E-02***
	CD278	-0.09349	2.12E-01	3.37E-01
Th1	CD212	-0.28929	**8.19E-05***	**5.96E-04***
	CD191	-0.10567	1.58E-01	2.75E-01
	CD195	-0.2203	**2.96E-03***	**1.23E-02***
Th2	CD194	-0.3358	**4.06E-06***	**4.22E-05***
	CD198	-0.15153	**4.23E-02***	9.99E-02
	CD365	-0.44608	**3.48E-10***	**9.93E-09***
Th17	CD360	-0.16464	**2.72E-02***	6.98E-02
	IL23R	-0.33128	**5.56E-06***	**5.56E-05***
	CD196	-0.57397	**3.68E-17***	**5.88E-15***
Treg	FOXP3	-0.1535	**3.97E-02***	9.50E-02
	CD73	-0.11494	1.24E-01	2.28E-01
	CD127	0.105172	1.60E-01	3.00E-01
T cell exhaustion	PD-1	-0.20453	**5.89E-03***	**2.53E-02***
	CTLA4	-0.16284	**2.90E-02***	8.42E-02
	LAG3	-0.02181	7.71E-01	8.79E-01
Macrophage	CD68	0.003556	9.62E-01	9.86E-01
	CD11b	-0.11248	1.33E-01	2.60E-01
M1 Macrophage	NOS2	-0.37043	**3.08E-07***	**4.93E-06***
	IRF5	0.384121	**1.02E-07***	**1.85E-06***
M2 Macrophage	CD163	-0.12681	8.98E-02	1.94E-01
	CD206	-0.00262	9.72E-01	9.89E-01
TAM	CCL2	0.005064	9.46E-01	9.79E-01
	CD86	0.107275	1.52E-01	2.87E-01
Monocyte	CD14	-0.01487	8.43E-01	9.31E-01
	CD33	-0.08073	2.81E-01	4.45E-01
Natural killer cell	CD57	-0.26358	**3.50E-04***	**2.46E-03***
	KIR3DL1	-0.13801	6.47E-02	1.53E-01
	CD7	-0.25512	**5.48E-04***	**3.66E-03***
Neutrophil	CD16	0.046633	5.34E-01	6.82E-01
	CD55	-0.38651	**8.36E-08***	**1.59E-06***
Dendritic cell	CD1C	-0.0649	3.87E-01	5.51E-01
	CD141	0.545347	**2.46E-15***	**4.92E-14***

** and bold values indicate P < 0.05.*

markers of different T cells in ESCA. including CD8B, CD183, CD185, CD212, CD195, CD194, CD365, IL23R, CD196, and PD-1 ($P < 0.05$, **Table 2**). It indicates that GLUT1 may be involved in the T cell immune response in ESCA. We also found that the expression level of GLUT1 was significantly correlated with the immune markers nitric oxide synthase 2 (NOS2) and interferon regulatory factor 5 (IRF5) of M1 macrophage in ESCA ($P < 0.05$, **Table 2**). It indicates that GLUT1 may regulate macrophage polarization in ESCA. We also found that the expression of GLUT1 was significantly correlated with immune markers of NK cells, Neutrophil and DC in ESCA, including CD57, CD7, CD55 and CD141 ($P < 0.05$, **Table 2**). These results indicate that the expression of GLUT1 in ESCA is related to immune cell infiltration in different ways.

In addition, we divided 162 tumor samples into two groups based on GLUT1 expression, with 81 samples in the high-expression group and 81 samples in the low-expression group. We tried to analyze the differential expression of 22 immune cells between different GLUT1 expression groups to determine whether the tumor immune microenvironment is different between high GLUT1 expression level and low GLUT1 expression level in ESCA (**Figure 3C**). The results show that the expression of memory B cell, resting memory CD4 + T cell, regulatory T cell (Tregs), activated NK cell, M0 macrophage and activated myeloid dendritic cell are quite different between the high and low GLUT1 expression groups. The results showed that, compared with the low expression group, activated NK cell, M0 macrophage and activated myeloid dendritic cell increased in the high expression group of GLUT1 ($P < 0.05$), while the memory B cell, resting memory CD4 + T cell and regulatory T cell (Tregs) decreased ($P < 0.05$).

GLUT1 Expression Is Associated With m6A RNA Methylation Regulators in ESCA

The m6a modification plays an important role in the occurrence and development of ESCA. We tried to analyze whether GLUT1

expression is related to m6A modification. We analyzed the TCGA ESCA data set to study the correlation between the expression of GLUT1 and 20 m6A-related genes in ESCA (**Figure 4A**). The results showed that GLUT1 expression was significantly positively correlated with 7 m6A-related genes in ESCA, including IGF2BP2 (r = 0.1984, P = 0.0114), YTHDF2 (r = 0.3135, P < 0.0001), HNRNPC (r = 0.3758, P < 0.0001), METTL3 (r = 0.1689, P = 0.0317), VIRMA (r = 0.1857, P = 0.018), FTO (r = 0.3, P = 0.0001) and ALKBH5 (r = 0.2744, P = 0.0004). Draw a scatter plot to show the correlation between GLUT1 and m6A related genes (**Figure 4B**). In addition, we divided 162 tumor samples into two groups based on GLUT1 expression, with 81 samples in the high-expression group and 81 samples in the low-expression group. We tried to analyze the differential expression of 20 m6A related genes between different GLUT1 expression groups to determine whether the m6A modification is different between high GLUT1 expression level and low GLUT1 expression level in ESCA (**Figure 4C**). The results showed that, compared with the low expression group, the expression of METTL3, VIRMA, YTHDC1, IGF2BP2, YTHDF2, HNRNPA2B1, HNRNPC, FTO and ALKBH5 increased in the high expression group of GLUT1 (P < 0.05). The above results indicate that GLUT1 is closely related to m6A modification in ESCA.

GLUT1 Related ceRNA Network Construction in ESCA

There is growing evidence that the lncRNA-miRNA-mRNA ceRNA network plays a key role in a variety of human cancers, so we tried to analyze and construct a ceRNA network involving GLUT1 in ESCA. We use PITA, miRanda and TargetScan databases to analyze and predict 79, 28 and 18 GLUT1 target miRNAs, respectively. Venn diagram shows the prediction results of GLUT1 target miRNA in PITA, miRanda and TargetScan software. A total of 14 target miRNAs are common predicted by 3 databases, including hsa-miR-19a-3p, hsa-miR-19b-3p, hsa-miR-22-3p, hsa-miR-148a-3p, hsa-miR-130a-3p, hsa-miR-140-5p, hsa-miR-152-3p, hsa-miR-301a-3p, hsa-miR-130b-3p, hsa-miR-328-3p, hsa-miR-148b-3p, hsa-miR-410-3p, hsa-miR-454-3p and hsa-miR-301b-3p (**Figure 5A**). In addition, we analyze the correlation between target miRNA expression and GLUT1 expression to screen for miRNAs that are more in line with ceRNA conditions. As shown in **Figure 5B**, correlation analysis proved that there are 3 target miRNAs expression levels negatively correlated with GLUT1, namely hsa-

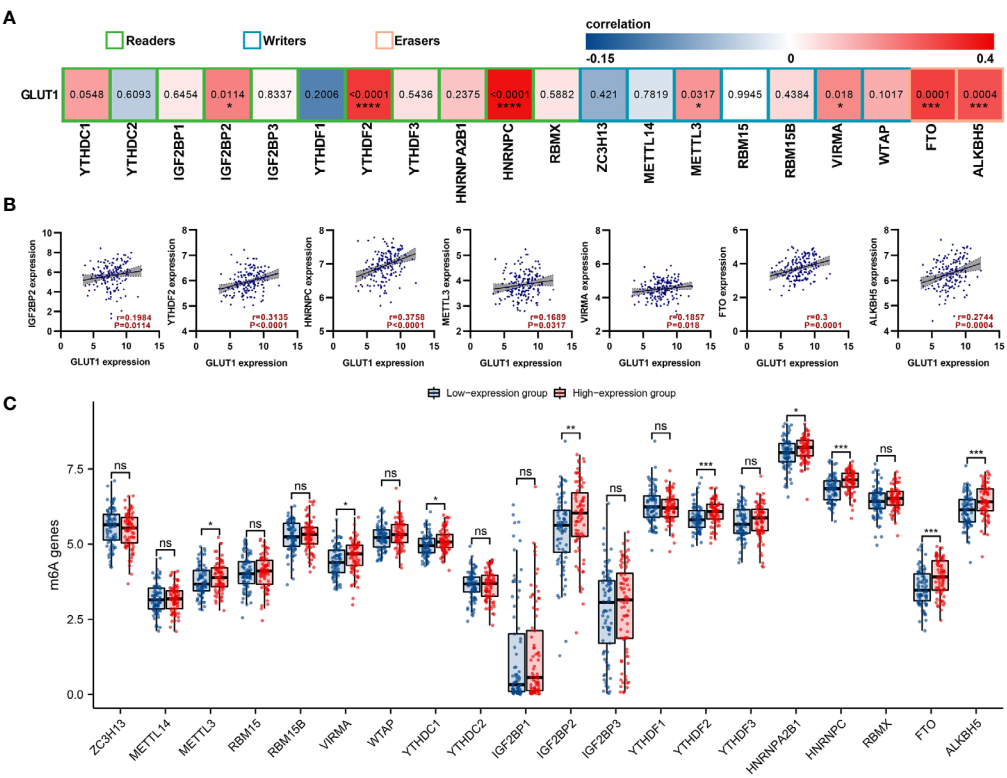


FIGURE 4 | Correlations of GLUT1 expression with m6A related genes in esophageal carcinoma (ESCA). **(A)** TCGA cohort analyzed the correlation between the expression level of GLUT1 and the expression of m6A-related genes in ESCA. **(B)** Draw a scatter plot to show the correlation between GLUT1 and m6A related genes. Related m6A related genes include IGF2BP2, YTHDF2, HNRNPC, METTL3, VIRMA, FTO and ALKBH5. **(C)** The differential expression of m6A related genes in the high and low GLUT1 expression groups in ESCA tumor samples. *p < 0.05; **p < 0.01; ***p < 0.001; ****p <0.0001. ns, not significant.

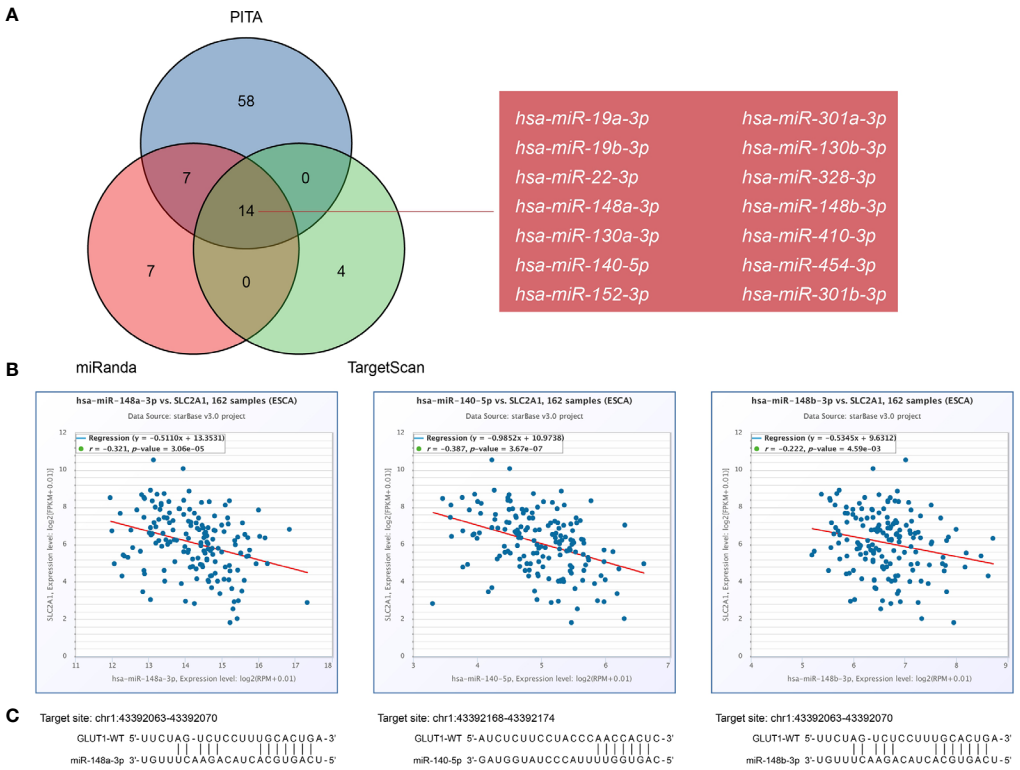


FIGURE 5 | Prediction of miRNAs targeting GLUT1 in esophageal carcinoma (ESCA). **(A)** Venn graph showing the prediction results of GLUT1 targets in PITA, miRanda, and TargetScan software packages. **(B)** Use starBase software to analyze the correlation between GLUT1 and the target miRNA. Use scatter plots to show miRNA-mRNA with significant correlation. **(C)** TargetScan predicts the potential binding site of GLUT1 to the target miRNA.

miR-148a-3p (r = -0.321, $P < 0.0001$), hsa-miR-140-5p (r = -0.387, $P < 0.0001$) and hsa-miR-148b-3p (r = -0.222, $P < 0.0001$). TargetScan predicts the potential binding site of GLUT1 to the target miRNA (**Figure 5C**).

We used the miRNet and starBase online database to further predict the lncRNA that may bind to the three target miRNAs (hsa-miR-148a-3p, hsa-miR-140-5p and hsa-miR-148b-3p) and display them through the Venn diagram (**Figures 6A–C**). Based on the ceRNA network hypothesis, there is a negative correlation between lncRNA and miRNA. Therefore, we used the starBase database to analyze the correlation between target lncRNA expression and miRNA in ESCA. As shown in **Figure 6D**, correlation analysis proved that there are 4 target lncRNAs expression levels that are negatively correlated with hsa-miR-148b-3p, namely HOTAIRM1, LINC00174, OIP5-AS1 and A1BG-AS1. However, only the expression level of DHRS4-AS1 was negatively correlated with hsa-miR-140-5p (**Figure 6E**), and the expression of A1BG-AS1 was negatively correlated with hsa-miR-148a-3p (**Figure 6F**). Based on the ceRNA hypothesis, there is an inverse relationship between miRNA and lncRNA or mRNA, so we can construct 6 pairs of ceRNA networks (A1BG-AS1- miR-148a-3p-GLUT1, HOTAIRM1-miR-148b-3p-GLUT1, LINC00174-miR-148b-3p-GLUT1, OIP5-AS1-miR-148b-3p-GLUT1, A1BG-AS1-miR-148b-3p-GLUT1 and DHRS4-AS1-miR-140-5p-GLUT1) based on the correlation analysis results (**Figure 6G**).

DISCUSSION

GLUTs are mainly distributed in the cell membrane area and can mediate glucose transport across the membrane. The up-regulation of GLUTs mainly exists in most tumor cells. It is a key factor for the uptake of glucose by tumor cells and may also be one of the causes of early carcinogenesis (57). Among the known diseases in humans, mutations in the GLUT1 gene affect the normal uptake of glucose by cells, which in turn triggers a series of diseases such as brain atrophy and developmental delay. But in cancer, tumor cells need to take in a lot of glucose to maintain the malignant proliferation of cells, so the expression of GLUT1 will eventually affect the development of cancer (11). Studies have shown that GLUT1 is highly expressed in a variety of cancers, and the expression level is closely related to clinical pathological characteristics such as tumor stage and tumor grade (4–8). In this study, we verified the differential expression of GLUT1 in tumors through experiments and bioinformatics analysis.

Analysis of the Oncomine database found that GLUT1 was highly expressed in 12 cancers, and the TCGA cohort showed that GLUT1 was highly expressed in 13 cancers, which is consistent with previous studies. Through the analysis of GEO and TCGA ESCA cohort, it was found that compared with normal tissues, the expression of GLUT1 in ESCA samples was significantly increased ($P < 0.05$). We also detected the

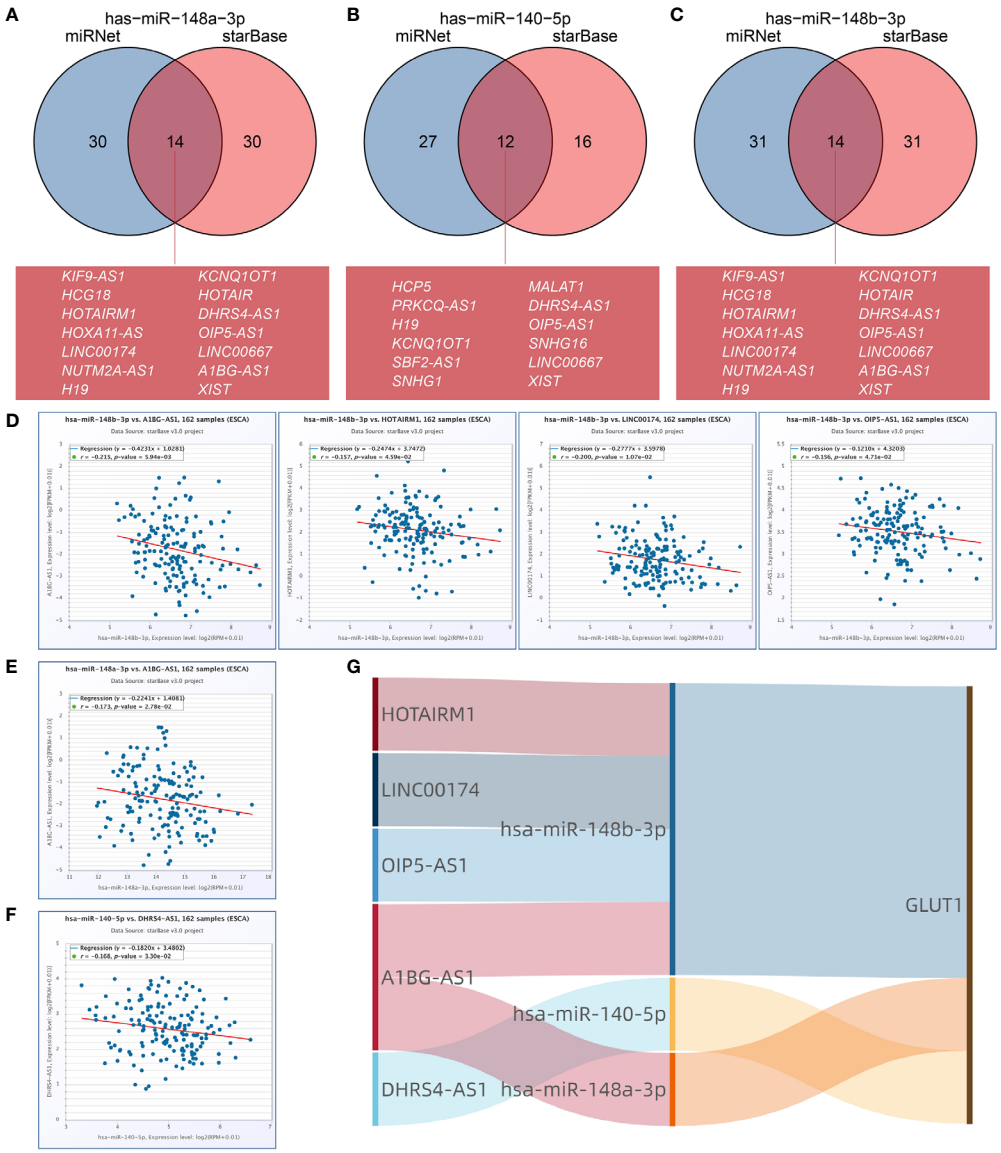


FIGURE 6 | Prediction of lncRNA and ceRNA network construction in esophageal carcinoma (ESCA). The Venn diagram respectively shows the target lncRNA of has-miR-148a-3p **(A)**, has-miR-140-5p **(B)**, and has-miR-148b-3p **(C)**. Use starBase software to analyze the correlation between miRNA and the target lncRNA. Use scatter plots to show miRNA-mRNA with significant correlation. LncRNA related to has-miR-148b-3p **(D)**. LncRNA related to has-miR-148a-3p **(E)**. LncRNA related to has-miR-140-5p **(F)**. **(G)** The Sankey diagram shows the lncRNA-miRNA-mRNA (GLUT1) regulatory network in line with the ceRNA hypothesis.

expression of GLUT1 in ESCA samples and normal tissue samples by qRT-PCR and IHC, and the analysis results were the same as above. Kato et al. (58) found that the expression of GLUT1 in ESCA is closely related to the tumor status, metastatic status, lymph node status and pathological stage of patients. This is different from the results of our analysis using TCGA ESCA cohort, which may be due to the large difference in sample size. However, many researchers have found that ESCA patients with high GLUT1 expression have a worse prognosis than those with low GLUT1 expression (58–60). In summary, GLUT1 can be used as a potential diagnostic and prognostic marker for ESCA.

At present, the research on the function and mechanism of GLUT1 in tumors mainly focuses on the glycolysis process of tumor cells. Zheng et al. (61) found that Circ_0058063 up-regulates GLUT1 expression and promotes glucose uptake in ESCC, thereby promoting cell proliferation. However, no study on the functional enrichment analysis of GLUT1 co-expression in ESCA has been reported. In this study, we used the LinkedOmics database to analyze the GLUT1 co-expression in

ESCA. Through the GO and KEGG functional enrichment analysis of 100 genes related to GLUT1, it is found that the GLUT1 co-expression is mainly related to the structural constituent of cytoskeleton, intermediate filament binding, cell-cell adheres junction, and epidermis development. KEGG pathway analysis showed that the GLUT1 co-expression was mainly related to the P53 signaling pathway. These functions and pathways are all related to the occurrence and development of ESCA. PPI analysis found that GLUT1 has the strongest correlation with TP53, GIPC1 and INS, and these three proteins all play important roles in the proliferation of tumor. The above results suggest that GLUT1 may not only participate in glycolysis in ESCA, but may also have multiple biological functions.

To explore the relationship between GLUT1 and immune infiltration in ESCA, we used the TIMER database to reveal the relationship between GLUT1 expression and the level of immune infiltration in ESCA. We found that GLUT1 expression is significantly correlated with B cells, CD4 + T cells, macrophages, and dendritic cells. In addition, it was also found that GLUT1 CNV was significantly correlated with the infiltration level of CD4 + T cells, neutrophils, and dendritic cells. Moreover, the correlation between GLUT1 expression and immune cell marker genes suggests that GLUT1 plays an important role in regulating ESCA tumor immunity. TIMER database analysis found that GLUT1 expression has a significant correlation with the gene markers of B cells (CD19, CD20), CD8 + T Cell (CD8B), Tfh (CD183, CD185), Th1 (CD212, CD195), Th2 (CD194, CD365), Th17 (IL23R, CD196), T cell exhaustion (PD-1), M1 macrophage (NOS2, IRF5), NK cell (CD57), neutrophil (CD55), and DC (CD141). We also found that the level of NK cell, M0 macrophage and activated myeloid dendritic cell increased in the GLUT1 high expression group, while the levels of memory B cell, resting memory CD4 + T cell and regulatory T cell decreased. We believe that GLUT1 can affect the immune infiltration of ESCA by regulating memory B cell. After activating a human's primary immunity, B cells will proliferate in large numbers, most of which will differentiate into effector B cells, and the rest will differentiate into memory B cells. Memory B cells exist in the human body for a long time and can play a long-term role. Ledesma et al. (62) found that the number of memory B cells in immune responders increased greatly, and they played a role in inhibiting tumor cell proliferation and metastasis. We can infer that the overexpression of GLUT1 inhibits the immune response and infiltration of memory B cells. We believe that excessive GLUT1 in patients with ESCA will trigger an anti-tumor immune response. These findings suggest that GLUT1 plays an important role in the regulation and recruitment of immune infiltrating cells in ESCA. However, controlled trials and clinical trials are needed to explain the relationship more accurately between GLUT1 and memory B cells *in vivo*.

M6A methylation is the most common form of mRNA modification in eukaryotes, and it plays a vital role in promoting tumor proliferation and migration. Chen et al. (11) found that the key m6A gene METTL3 can enhance the stability of GLUT1 mRNA, increase the glucose uptake and lactate production of colorectal cancer cells, and promote the development of colorectal cancer. Huang et al. (63) found that IGF2BP2 directly binds to GLUT1 mRNA and stabilizes GLUT1 mRNA, thereby promoting aerobic glycolysis and proliferation of pancreatic ductal adenocarcinoma cells. In this study, we tried to analyze whether GLUT1 expression is related to m6A modification in ESCA. We found that GLUT1 expression was significantly correlated with IGF2BP2, YTHDF2, HNRNPC, METTL3, VIRMA, FTO and ALKBH5. We also found that the levels of METTL3, VIRMA, YTHDC1, IGF2BP2, YTHDF2, HNRNPA2B1, HNRNPC, FTO and ALKBH5 increased in the GLUT1 high expression group. We believe that the GLUT1 gene may be modified by m6A to enhance mRNA stability, which in turn enhances the glycolysis and proliferation of ESCA.

The crosstalk between ceRNA is achieved by long non-coding RNA (lncRNA) or circular RNA (circRNA) competitively binding miRNA to affect mRNA expression. lncRNA and circRNA play the role of miRNA sponge, reducing the abundance of miRNA in the body, thereby reducing the inhibitory effect of miRNA on downstream target genes. Li et al. (64) found that lncRNA RAD51-AS1 can inhibit the miR-29b/c-3p/NDRG2 signal axis and the expression of hexokinase 2 and GLUT1, thereby inhibiting the progression of colorectal cancer. Chen et al. (12) found that circRNA_100290 can be used as ceRNA to eliminate the inhibitory effect of mir-378a on GLUT1, thereby promoting glycolysis and proliferation of oral squamous cell carcinoma. In this study, we first predicted some upstream miRNAs of GLUT1. The 3 databases jointly predicted 14 potential upstream miRNAs, but the expression of only 3 miRNAs (hsa-miR-148a-3p, hsa-miR-140-5p, hsa-miR-148b-3p) was significantly negatively correlated with GLUT1 in ESCA. Mari et al. (65) reported that overexpression of miR-148a-3p in ESCA can enhance tumor immune response. Chen et al. (66) found that miR-140-5p is under-expressed in ESCA, and it may regulate the cell invasion of ESCA by regulating the expression of Slug. Then, we further predicted the upstream lncRNA of these key miRNAs. Through correlation analysis, only 5 lncRNAs (HOTAIRM1, LINC00174, OIP5-AS1, A1BG-AS1, DHRS4-AS1) can be defined as key lncRNAs. HOTAIRM1 is up-regulated in glioblastoma and promotes tumor cell migration and invasion (67). LINC00174 targets miR-4500 in laryngeal papilloma to inhibit BZW2 and promote tumor cell proliferation (68). OIP5-AS1 promotes the invasion and migration of ovarian cancer cells (69). These reports further hint at the feasibility of our analysis. Of course, although the ceRNA network of GLUT1 was obtained through bioinformatics analysis, more experiments are needed to confirm our prediction.

In summary, this is the first comprehensive analysis of the relationship between GLUT1 expression and tumor immune infiltration, m6A modification, and ceRNA network in ESCA. The GLUT1 gene may be modified by m6A to enhance the stability of its mRNA, thereby enhancing the effect of promoting glycolysis and cell proliferation in ESCA. The expression of GLUT1 is related to a variety of immune cells and may affect

ESCA tumor immunity by inhibiting the infiltration of memory B cells. The construction of ceRNA network of GLUT1 indicates that GLUT1 may participate in a variety of molecular regulatory mechanisms in ESCA. GLUT1 can be used as an effective biomarker for the diagnosis and treatment of ESCA.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Taihe Hospital Affiliated of Hubei University of Medicine. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

X-SL, YG, and L-BW conceived the project and wrote the manuscript. X-SL, H-BW, PY, YJ, S-BG, Y-LW, and X-QC participated in data analysis. L-MZ, J-WY, X-YK, and X-YL participated in discussion and language editing. Z-JP reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Hubei province's Outstanding Medical Academic Leader program, the Foundation for Innovative Research Team of Hubei Provincial Department of Education T2020025, the Hubei Provincial Department of Science and Technology Innovation Group Program (grant no. 2019CFA034), Free-exploring Foundation of Hubei University of Medicine (grant no. FDFR201903), Open Project of Hubei Key Laboratory of Embryonic Stem Cell Research (grant no.2020ESOF009), and the Key Discipline Project of Hubei University of Medicine.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.665388/full#supplementary-material

Supplementary Table 1 | GLUT1 expression in cancerous versus normal tissue in ONCOMINE.

Supplementary Table 2 | GLUT1 co-expressed genes.

Supplementary Table 3 | The GO and KEGG enrichment analysis of GLUT1 related genes.

Supplementary Table 4 | The PPI network of GLUT1.

REFERENCES

1. Kasahara M, Hinkle PC. Reconstitution and Purification of the D-glucose Transporter From Human Erythrocytes. *J Biol Chem* (1977) 252(20):7384–90. doi: 10.1016/S0021-9258(19)66976-0
2. Uldry M, Thorens B. The SLC2 Family of Facilitated Hexose and Polyol Transporters. *Pflugers Archiv Eur J Physiol* (2004) 447(5):480–9. doi: 10.1007/s00424-003-1085-0
3. Deng D, Xu C, Sun P, Wu J, Yan C, Hu M, et al. Crystal Structure of the Human Glucose Transporter Glut1. *Nature* (2014) 510(7503):121–5. doi: 10.1038/nature13306
4. Pereira KMA, Chaves FN, Viana TSA, Carvalho FSR, Costa FWG, Alves APNN, et al. Oxygen Metabolism in Oral Cancer: HIF and GLUTs (Review). *Oncol Lett* (2013) 6(2):311–6. doi: 10.3892/ol.2013.1371
5. Berlth F, Mönig S, Pinther B, Grimminger P, Maus M, Schlösser H, et al. Both GLUT-1 and GLUT-14 Are Independent Prognostic Factors in Gastric Adenocarcinoma. *Ann Surg Oncol* (2015) 22(Suppl 3):S822–31. doi: 10.1245/s10434-015-4730-x
6. Avanzato D, Pupo E, Ducano N, Isella C, Bertalot G, Luise C, et al. High USP6NL Levels in Breast Cancer Sustain Chronic AKT Phosphorylation and GLUT1 Stability Fueling Aerobic Glycolysis. *Cancer Res* (2018) 78(13):3432–44. doi: 10.1158/0008-5472.CAN-17-3018
7. van Laarhoven HWM, Kaanders JHAM, Lok J, Peeters WJM, Rijken PFJW, Wiering B, et al. Hypoxia in Relation to Vasculature and Proliferation in Liver Metastases in Patients With Colorectal Cancer. *Int J Radiat Oncol Biol Phys* (2006) 64(2):473–82. doi: 10.1016/j.ijrobp.2005.07.982
8. Rudlowski C, Moser M, Becker AJ, Rath W, Buttner R, Schroder W, et al. Glut1 mRNA and Protein Expression in Ovarian Borderline Tumors and Cancer. *Oncology* (2004) 66(5):404–10. doi: 10.1159/000079489
9. Liu X, Yuan L, Gao Y, Zhou L, Yang J, Pei Z. Overexpression of METTL3 Associated With the Metabolic Status on 18F-FDG PET/CT in Patients With Esophageal Carcinoma. *J Cancer* (2020) 11(16):4851–60. doi: 10.7150/jca.44754
10. Na KJ, Choi H, Oh HR, Kim YH, Lee SB, Jung YJ, et al. Reciprocal Change in Glucose Metabolism of Cancer and Immune Cells Mediated by Different Glucose Transporters Predicts Immunotherapy Response. *Theranostics* (2020) 10(21):9579–90. doi: 10.7150/thno.48954
11. Chen H, Gao S, Liu W, Wong C, Wu J, Wu J, et al. RNA M(6)a Methyltransferase METTL3 Facilitates Colorectal Cancer by Activating M (6)a-GLUT1-Mtorc1 Axis and Is a Therapeutic Target. *Gastroenterology* (2021) 160(4):1284–300.e16. doi: 10.1053/j.gastro.2020.11.013
12. Chen X, Yu J, Tian H, Shan Z, Liu W, Pan Z, et al. Circle RNA Hsa_Circrna_100290 Serves as a ceRNA for miR-378a to Regulate Oral Squamous Cell Carcinoma Cells Growth Via Glucose Transporter-1 (GLUT1) and Glycolysis. *J Cell Physiol* (2019) 234(11):19130–40. doi: 10.1002/jcp.28692
13. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, et al. Oncomine 3.0: Genes, Pathways, and Networks in a Collection of 18,000 Cancer Geneexpression Profiles. *Neoplasia (New York NY)* (2007) 9 (2):166–80. doi: 10.1593/neo.07112
14. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, et al. Comprehensive Analyses of Tumor Immunity: Implications for Cancer Immunotherapy. *Genome Biol* (2016) 17(1):174. doi: 10.1186/s13059-016-1028-7
15. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. Timer: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immunecells. *Cancer Res* (2017) 77(21):e108–10. doi: 10.1158/0008-5472.CAN-17-0307
16. Tomczak K, Czerwińska P, Wiznerowicz M. Review The Cancer Genome Atlas (Tcga): An Immeasurable Source of Knowledge. *Współczesna Onkol* (2015) 1A(1A):68–77. doi: 10.5114/wo.2014.47136

17. Li Y, Xiao J, Bai J, Tian Y, Qu Y, Chen X, et al. Molecular Characterization and Clinical Relevance of M(6)a Regulators Across 33 Cancer Types. *Mol Cancer* (2019) 18(1):137. doi: 10.1186/s12943-019-1066-3
18. Liu X, Huang H, Gao Y, Zhou L, Yang J, Li X, et al. Visualization of Gene Therapy With a Liver Cancer-Targeted Adeno-Associated Virus 3 Vector. *J Cancer* (2020) 11(8):2192–200. doi: 10.7150/jca.39579
19. Vasaikar SV, Straub P, Wang J, Zhang B. Linkedomics: Analyzing Multi-Omics Data Within and Across 32 Cancer Types. *Nucleic Acids Res* (2018) 46 (D1):D956–63. doi: 10.1093/nar/gkx1090
20. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. String V11: Protein-Protein Association Networks With Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res* (2019) 47(D1):D607–13. doi: 10.1093/nar/gky1131
21. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase V2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and Protein-RNA Interaction Networks From Large-Scale CLIP-Seq Data. *Nucleic Acids Res* (2014) 42(Database issue): D92–7. doi: 10.1093/nar/gkt1248
22. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting Effective MicroRNA Target Sites in Mammalian Mrnas. *Elife* (2015) 4(12):e5005. doi: 10.7554/eLife.05005
23. Chang L, Zhou G, Soufan O, Xia J. miRNet 2.0: Network-Based Visual Analytics for Mirna Functional Analysis and Systems Biology. *Nucleic Acids Res* (2020) 48(W1):W244–51. doi: 10.1093/nar/gkaa467
24. Sanchez-Carbayo M, Socci ND, Lozano J, Saint F, Cordon-Cardo C. Defining Molecular Profiles of Poor Outcome in Patients With Invasive Bladder Cancer Using Oligonucleotide Microarrays. *J Clin Oncol Off J Am Soc Clin Oncol* (2006) 24(5):778–89. doi: 10.1200/JCO.2005.03.2375
25. Dyrskjøt L, Kruhøffer M, Thykjaer T, Marcussen N, Jensen JL, Møller K, et al. Gene Expression in the Urinary Bladder: A Common Carcinoma In Situ Gene Expression Signature Exists Disregarding Histopathological Classification. *Cancer Res* (2004) 64(11):4040–8. doi: 10.1158/0008-5472.CAN-03-3620
26. Zhao H, Langerød A, Ji Y, Nowels KW, Nesland JM, Tibshirani R, et al. Different Gene Expression Patterns in Invasive Lobular and Ductal Carcinomas of the Breast. *Mol Biol Cell* (2004) 15(6):2523–36. doi: 10.1091/mbc.e03-11-0786
27. Richardson AL, Wang ZC, De Nicolo A, Lu X, Brown M, Miron A, et al. X Chromosomal Abnormalities in Basal-Like Human Breast Cancer. *Cancer Cell* (2006) 9(2):121–32. doi: 10.1016/j.ccr.2006.01.013
28. Curtis C, Shah SP, Chin S, Turashvili G, Rueda OM, Dunning MJ, et al. The Genomic and Transcriptomic Architecture of 2,000 Breast Tumours Reveals Novel Subgroups. *Nature* (2012) 486(7403):346–52. doi: 10.1038/nature10983
29. Ki DH, Jeung H, Park CH, Kang SH, Lee GY, Lee WS, et al. Whole Genome Analysis for Liver Metastasis Gene Signatures in Colorectal Cancer. *Int J Cancer* (2007) 121(9):2005–12. doi: 10.1002/ijc.22975
30. Skrzypczak M, Goryca K, Rubel T, Paziewska A, Mikula M, Jarosz D, et al. Modeling Oncogenic Signaling in Colon Tumors by Multidirectional Analyses of Microarray Data Directed for Maximization of Analytical Reliability. *PloS One* (2010) 5(10):e13091. doi: 10.1371/journal.pone.0013091
31. Hu N, Clifford RJ, Yang HH, Wang C, Goldstein AM, Ding T, et al. Genome Wide Analysis of DNA Copy Number Neutral Loss of Heterozygosity (CNNLOH) and its Relation to Gene Expression in Esophageal Squamous Cell Carcinoma. *BMC Genomics* (2010) 11(1):576. doi: 10.1186/1471-2164-11-576
32. Su H, Hu N, Yang HH, Wang C, Takikita M, Wang Q, et al. Global Gene Expression Profiling and Validation in Esophageal Squamous Cell Carcinoma and Its Association With Clinical Phenotypes. *Clin Cancer Res* (2011) 17 (9):2955–66. doi: 10.1158/1078-0432.CCR-10-2724
33. Chen X, Leung SY, Yuen ST, Chu K, Ji J, Li R, et al. Variation in Gene Expression Patterns in Human Gastric Cancers. *Mol Biol Cell* (2003) 14 (8):3208–15. doi: 10.1091/mbc.E02-12-0833
34. Ye H, Yu T, Temam S, Ziober BL, Wang J, Schwartz JL, et al. Transcriptomic Dissection of Tongue Squamous Cell Carcinoma. *BMC Genomics* (2008) 9 (1):69. doi: 10.1186/1471-2164-9-69
35. Pyeon D, Newton MA, Lambert PF, den Boon JA, Sengupta S, Marsit CJ, et al. Fundamental Differences in Cell Cycle Deregulation in Human Papillomavirus-Positive and Human Papillomavirus-Negative Head/Neck and Cervical Cancers. *Cancer Res* (2007) 67(10):4605–19. doi: 10.1158/0008-5472.CAN-06-3619
36. Beroukhim R, Brunet J, Di Napoli A, Mertz KD, Seeley A, Pires MM, et al. Patterns of Gene Expression and Copy-Number Alterations in Von-Hippel Lindau Disease-Associated and Sporadic Clear Cell Carcinoma of the Kidney. *Cancer Res* (2009) 69(11):4674–81. doi: 10.1158/0008-5472.CAN-09-0146
37. Gumz ML, Zou H, Kreinest PA, Childs AC, Belmonte LS, LeGrand SN, et al. Secreted Frizzled-Related Protein 1 Loss Contributes to Tumor Phenotype of Clear Cell Renal Cell Carcinoma. *Clin Cancer Res* (2007) 13(16):4740–9. doi: 10.1158/1078-0432.ccr-07-0143
38. Jones J, Otu H, Spentzos D, Kolia S, Inan M, Beecken WD, et al. Gene Signatures of Progression and Metastasis in Renal Cell Cancer. *Clin Cancer Res* (2005) 11(16):5730–9. doi: 10.1158/1078-0432.CCR-04-2225
39. Yusenko MV, Kuiper RP, Boethe T, Ljungberg B, van Kessel AG, Kovacs G. High-Resolution DNA Copy Number and Gene Expression Analyses Distinguish Chromophobe Renal Cell Carcinomas and Renal Oncocytomas. *BMC Cancer* (2009) 9(1):152. doi: 10.1186/1471-2407-9-152
40. Andersson A, Ritz C, Lindgren D, Eden P, Lassen C, Heldrup J, et al. Microarray-Based Classification of a Consecutive Series of 121 Childhood Acute Leukemias: Prediction of Leukemic and Genetic Subtype as Well as of Minimal Residual Disease Status. *Leukemia* (2007) 21(6):1198–203. doi: 10.1038/sj.leu.2404688
41. Su L, Chang C, Wu Y, Chen K, Lin C, Liang S, et al. Selection of DDX5 as a Novel Internal Control for Q-RT-PCR From Microarray Data Using a Block Bootstrap Re-Sampling Scheme. *BMC Genomics* (2007) 8(1):140. doi: 10.1186/1471-2164-8-140
42. Selamat SA, Chung BS, Girard L, Zhang W, Zhang Y, Campan M, et al. Genome-Scale Analysis of DNA Methylation in Lung Adenocarcinoma and Integration With mRNA Expression. *Genome Res* (2012) 22(7):1197–211. doi: 10.1101/gr.132662.111
43. Hou J, Aerts J, den Hamer B, van IJcken W, den Bakker M, Riegman P, et al. Gene Expression-Based Classification of Non-Small Cell Lung Carcinomas and Survival Prediction. *PloS One* (2010) 5(4):e10312. doi: 10.1371/journal.pone.0010312
44. Okayama H, Kohno T, Ishii Y, Shimada Y, Shiraishi K, Iwakawa R, et al. Identification of Genes Upregulated Inalk-Positive Andegfr/KRAS/ALK-Negative Lung Adenocarcinomas. *Cancer Res* (2012) 72(1):100–11. doi: 10.1158/0008-5472.CAN-11-1403
45. Beer DG, Kardia SLR, Huang C, Giordano TJ, Levin AM, Misek DE, et al. Gene-Expression Profiles Predict Survival of Patients With Lung Adenocarcinoma. *Nat Med* (2002) 8(8):816–24. doi: 10.1038/nm733
46. Landi MT, Dracheva T, Rotunno M, Figueroa JD, Liu H, Dasgupta A, et al. Gene Expression Signature of Cigarette Smoking and Its Role in Lung Adenocarcinoma Development and Survival. *PloS One* (2008) 3(2):e1651. doi: 10.1371/journal.pone.0001651
47. Eckerle S, Brune V, Doring C, Tiacci E, Bohle V, Sundstrom C, et al. Gene Expression Profiling of Isolated Tumour Cells From Anaplastic Large Cell Lymphomas: Insights Into its Cellular Origin, Pathogenesis and Relation to Hodgkin Lymphoma. *Leukemia* (2009) 23(11):2129–38. doi: 10.1038/leu.2009.161
48. Yoshihara K, Tajima A, Komata D, Yamamoto T, Kodama S, Fujiwara H, et al. Gene Expression Profiling of Advanced-Stage Serous Ovarian Cancers Distinguishes Novel Subclasses and Implicates ZEB2 in Tumor Progression and Prognosis. *Cancer Sci* (2009) 100(8):1421–8. doi: 10.1111/j.1349-7006.2009.01204.x
49. Pei H, Li L, Fridley BL, Jenkins GD, Kalari KR, Lingle W, et al. Fkbp51 Affects Cancer Cell Response to Chemotherapy by Negatively Regulating Akt. *Cancer Cell* (2009) 16(3):259–66. doi: 10.1016/j.ccr.2009.07.016
50. Iacobuzio-Donahue CA, Maitra A, Olsen M, Lowe AW, Van Heek NT, Rosty C, et al. Exploration of Global Gene Expression Patterns in Pancreatic Adenocarcinoma Using cDNA Microarrays. *Am J Pathol* (2003) 162 (4):1151–62. doi: 10.1016/S0002-9440(10)63911-9
51. Logsdon CD, Simeone DM, Binkley C, Arumugam T, Greenson JK, Giordano TJ, et al. Molecular Profiling of Pancreatic Adenocarcinoma and Chronic Pancreatitis Identifies Multiple Genes Differentially Regulated in Pancreatic Cancer. *Cancer Res* (2003) 63(10):2649–57.
52. Badea L, Herlea V, Dima SO, Dumitrascu T, Popescu I. Combined Gene Expression Analysis of Whole-Tissue and Microdissected Pancreatic Ductal

Adenocarcinoma Identifies Genes Specifically Overexpressed in Tumor Epithelia. *Hepato-gastroenterology* (2008) 55(88):2016.
53. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal Gene Expression Predicts Clinical Outcome in Breast Cancer. (2008) 14(5):518–27. doi: 10.1038/nm1764
54. Kim SM, Park Y, Park ES, Cho JY, Izzo JG, Zhang D, et al. Prognostic Biomarkers for Esophageal Adenocarcinoma Identified by Analysis of Tumor Transcriptome. *PloS One* (2010) 5(11):e15074. doi: 10.1371/journal.pone.0015074
55. Haferlach T, Kohlmann A, Wieczorek L, Basso G, Kronnie GT, Béné M, et al. Clinical Utility of Microarray-Based Gene Expression Profiling in the Diagnosis and Subclassification of Leukemia: Report From the International Microarray Innovations in Leukemia Study Group. *J Clin Oncol Off J Am Soc Clin Oncol* (2010) 28(15):2529–37. doi: 10.1200/JCO.2009.23.4732
56. Sudo T, Nishida R, Kawahara A, Saisho K, Mimori K, Yamada A, et al. Clinical Impact of Tumor-Infiltrating Lymphocytes in Esophageal Squamous Cell Carcinoma. *Ann Surg Oncol* (2017) 24(12):3763–70. doi: 10.1245/s10434-017-5796-4
57. Ancey P, Contat C, Meylan E. Glucose Transporters in Cancer - From Tumor Cells to the Tumor Microenvironment. *FEBS J* (2018) 285(16):2926–43. doi: 10.1111/febs.14577
58. Kato H, Takita J, Miyazaki T, Nakajima M, Fukai Y, Masuda N, et al. Glut-1 Glucose Transporter Expression in Esophageal Squamous Cell Carcinoma Is Associated With Tumor Aggressiveness. *Anticancer Res* (2002) 22 (5):2635–9.
59. Tohma T, Okazumi S, Makino H, Cho A, Mochizuki R, Shuto K, et al. Overexpression of Glucose Transporter 1 in Esophageal Squamous Cell Carcinomas: A Marker for Poor Prognosis. *Dis Esophagus Off J Int Soc Dis Esophagus* (2005) 18(3):185–9. doi: 10.1111/j.1442-2050.2005.00489.x
60. Kato H, Takita J, Miyazaki T, Nakajima M, Fukai Y, Masuda N, et al. Correlation of 18-F-fluorodeoxyglucose (Fdg) Accumulation With Glucose Transporter (Glut-1) Expression in Esophageal Squamous Cell Carcinoma. *Anticancer Res* (2003) 23(4):3263–72.
61. Zheng Y, Chen Y, Jiang H, Zhang H, Wang H, Xu J, et al. Circ_0058063 Upregulates Glut1 Expression and Promotes Glucose-Uptake in Esophageal Squamous-Cell Carcinomas. *J Thorac Dis* (2020) 12(3):925–31. doi: 10.21037/jtd.2019.12.57
62. Helmink BA, Reddy SM, Gao J, Zhang S, Basar R, Thakur R, et al. B Cells and Tertiary Lymphoid Structures Promote Immunotherapy Response. *Nature* (2020) 577(7791):549–55. doi: 10.1038/s41586-019-1922-8
63. Huang S, Wu Z, Cheng Y, Wei W, Hao L. Insulin-Like Growth Factor 2 Mrna Binding Protein 2 Promotes Aerobic Glycolysis and Cell Proliferation in Pancreatic Ductal Adenocarcinoma Via Stabilizing Glut1 Mrna. *Acta Bioch Bioph Sin* (2019) 51(7):743–52. doi: 10.1093/abbs/gmz048
64. Li C, Wang P, Du J, Chen J, Liu W, Ye K. LncRNA RAD51-AS1/miR-29b/c-3p/NDRG2 Crosstalk Repressed Proliferation, Invasion and Glycolysis of Colorectal Cancer. *IUBMB Life* (2021) 73(1):286–98. doi: 10.1002/iub.2427
65. Mari L, Hoefnagel SJM, Zito D, van de Meent M, van Endert P, Calpe S, et al. microRNA 125a Regulates Mhc-I Expression on Esophageal Adenocarcinoma Cells, Associated With Suppression of Antitumor Immune Response and Poor Outcomes of Patients. *Gastroenterology* (2018) 155(3):784–98. doi: 10.1053/j.gastro.2018.06.030
66. Li W, Jiang G, Zhou J, Wang H, Gong Z, Zhang Z, et al. Down-Regulation of Mir-140 Induces EMT and Promotes Invasion by Targeting Slug in Esophageal Cancer. *Cell Physiol Biochem Int J Exp Cell Physiol Biochem Pharmacol* (2014) 34(5):1466–76. doi: 10.1159/000366351
67. Xie P, Li X, Chen R, Liu Y, Liu D, Liu W, et al. Upregulation of HOTAIRM1 Increases Migration and Invasion by Glioblastoma Cells. *Aging* (2020) 13 (2):2348–64. doi: 10.18632/aging.202263
68. Liu J, Yang T, Zhang Y, Wang S. Promotion of BZW2 by LINC00174 Through Mir-4500 Inhibition Enhances Proliferation and Apoptosis Evasion in Laryngeal Papilloma. *Cancer Cell Int* (2020) 20:471. doi: 10.1186/s12935-020-01559-3
69. Jiang X, Ye Z, Jiang Y, Yu W, Fang Q. Lncrna OIP5-AS1 Upregulates Snail Expression by Sponging miR-34a to Promote Ovarian Carcinoma Cell Invasion and Migration. *Biol Res* (2020) 53(1):49. doi: 10.1186/s40659-020-00315-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 29 June 2021
doi: 10.3389/fonc.2021.684941

Combined Impacts of Genetic Variants of Long Non-Coding RNA MALAT1 and the Environmental Carcinogen on the Susceptibility to and Progression of Oral Squamous Cell Carcinoma

Yi-Fang Ding[1,2†], Yu-Ching Wen[3,4†], Chun-Yi Chuang[5,6], Chiao-Wen Lin[7,8], Yi-Chieh Yang[9,10], Yu-Fan Liu[11], Wei-Min Chang[12], Lun-Ching Chang[13], Shun-Fa Yang[14,15] and Ming-Hsien Chien[9,16,17,18*]*

[1] Graduate Institute of Medical Sciences, College of Medicine, Taipei Medical University, Taipei, Taiwan, [2] Department of Otolaryngology, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan, [3] Department of Urology, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan, [4] Department of Urology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan, [5] School of Medicine, Chung Shan Medical University, Taichung, Taiwan, [6] Department of Otolaryngology, Chung Shan Medical University Hospital, Taichung, Taiwan, [7] Institute of Oral Sciences, Chung Shan Medical University, Taichung, Taiwan, [8] Department of Dentistry, Chung Shan Medical University Hospital, Taichung, Taiwan, [9] Graduate Institute of Clinical Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan, [10] Department of Medical Research, Tungs' Taichung MetroHarbor Hospital, Taichung, Taiwan, [11] Department of Biomedical Sciences, College of Medicine Sciences and Technology, Chung Shan Medical University, Taichung, Taiwan, [12] School of Oral Hygiene, College of Oral Medicine, Taipei Medical University, Taipei, Taiwan, [13] Department of Mathematical Sciences, Florida Atlantic University, Boca Raton, FL, United States, [14] Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan, [15] Department of Medical Research, Chung Shan Medical University Hospital, Taichung, Taiwan, [16] Pulmonary Research Center, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan, [17] Traditional Herbal Medicine Research Center, Taipei Medical University Hospital, Taipei, Taiwan, [18] Taipei Medical University (TMU) Research Center of Cancer Translational Medicine, Taipei Medical University, Taipei, Taiwan

OPEN ACCESS

Edited by:
Di Wu,
University of North Carolina at Chapel Hill, United States

Reviewed by:
Show-Mei Chuang,
National Chung Hsing University, Taiwan
Shaolei Teng,
Howard University, United States

***Correspondence:**
Ming-Hsien Chien
mhchien1976@gmail.com
Shun-Fa Yang
ysf@csmu.edu.tw

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *24 March 2021*
Accepted: *11 June 2021*
Published: *29 June 2021*

Citation:
Ding Y-F, Wen Y-C, Chuang C-Y, Lin C-W, Yang Y-C, Liu Y-F, Chang W-M, Chang L-C, Yang S-F and Chien M-H (2021) Combined Impacts of Genetic Variants of Long Non-Coding RNA MALAT1 and the Environmental Carcinogen on the Susceptibility to and Progression of Oral Squamous Cell Carcinoma. Front. Oncol. 11:684941. doi: 10.3389/fonc.2021.684941

Oral squamous cell carcinoma (OSCC) is the most common malignant tumor of the oral cavity, and long non-coding (lnc)RNA of metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) was recently reported to play a crucial role in OSCC development and progression. However, potential effects of genetic variants of MALAT1 on the development of OSCC are still unclear. Herein, we performed a case-control study in 1350 patients with OSCC and 1199 healthy controls to evaluate the association between functional single-nucleotide polymorphisms (SNPs) of MALAT1 and OSCC susceptibility, as well as its clinicopathologic characteristics. A TaqMan allelic discrimination assay was used to genotype four tagging SNPs, viz., rs3200401 C>T, rs619586 A>G, rs1194338 C>A, and rs7927113 G>A, and results showed that the MALAT1 rs3200401 T allele had a lower risk of OSCC (adjusted odds ratio (AOR): 0.779, 95% confidence interval (CI): 0.632~0.960, p=0.019) and a higher risk of developing moderately (grade II)/poorly (grade III) differentiated OSCC (AOR: 1.508-fold, 95% CI: 1.049~2.169, p=0.027) under a dominant model. According to environmental carcinogen exposure, patients with a betel quid-chewing habit who carried the T allele of rs3200401 more easily developed

high-grade (II/III) OSCC (AOR: 1.588, 95% CI: 1.055~2.390, p=0.027), and patients with the same genotype but who did not chew betel quid had a lower risk of developing lymph node metastasis (AOR: 0.437, 95% CI: 0.255~0.749, p=0.003). In addition to rs3200401, the rs619586 AG/GG genotype was associated with increased risks of developing advanced stages (III+IV) and larger tumor sizes (>T2) compared to the AA genotype, especially in the subgroup of betel quid chewers. Furthermore, analyses of clinical datasets revealed that the MALAT1 expression level was upregulated in OSCC compared to normal tissues, especially in the betel quid-chewing population. These results indicated involvement of MALAT1 SNPs rs3200401 and rs619586 in the development of OSCC and support the interaction between MALAT1 gene polymorphisms and the environmental carcinogen as a predisposing factor for OSCC progression.

Keywords: oral squamous cell carcinoma, MALAT1, single-nucleotide polymorphisms, susceptibility, progression

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is one of the six most frequent cancers in the world, the causes of OSCC are complex, and a lot of factors contribute to its development and progression. For example, continuous exposure to tobacco, alcohol use, and human papillomavirus infection are common, major risk factors for OSCC worldwide (1, 2), and betel nut chewing is another predominant risk factor causing OSCC in Taiwan (3). Until now, the pathogenesis of OSCC has remained unclear, but it was generally reported that OSCC is attributed to combined effects of various risk factors and genetic and epigenetic changes. Although great advances have been made in treating OSCC, including surgery, chemotherapy, and radiotherapy, the 5-year survival rate is only 50% (4), mainly due to delays in diagnoses which allows the cancer to metastasize. Thus, finding notable prognostic factors, metastatic predictors, and therapeutic targets in OSCC is urgently needed.

Long non-coding (lnc)RNAs are a type of RNA molecule with a length of >200 nucleotides (nt), which are unable to encode proteins. Expanding evidence indicates that lncRNAs are notable molecular markers involved in modulating gene expressions and cancer progression, such as tumor cell proliferation, invasion, metastasis, and angiogenesis (5–7). Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) is one of the most widely studied nuclear-retained lncRNAs that has garnered much attention in recent years due to its abundance and apparent role in various diseases. MALAT1 was shown to act as a competing endogenous (ce)RNA or micro (mi)RNA sponge which sequesters miRNAs under various conditions (8). In terms of cancer, MALAT1 was initially identified as an RNA whose expression is upregulated in primary lung tumors that had higher metastatic abilities (9). In addition to lung cancer, pro-oncogenic and prometastatic roles of MALAT1 were reported in a wide range of solid and non-solid tumors including OSCC. For instance, MALAT1 can function as a ceRNA to modulate signal transduction and activator of transcription 3 (STAT3) expression by absorbing miR-125b and further promote the growth of OSCC (10). The miR-101/enhancer of zeste homolog 2 (EZH2) axis is another pathway regulated by MALAT1 to facilitate proliferation and invasion of OSCC cells (11).

Recently, increasing evidence has indicated that single-nucleotide polymorphisms (SNPs) are universally present in lncRNA genes, and they may directly or indirectly influence lncRNA expression levels through various means and thus are likely to regulate the development and progression of cancer (12). To date, a number of SNPs in lncRNAs were found to be related to the development, progression, and prognosis of OSCC such as *maternally expressed 3* (*MEG3*) (13), *phosphatase and tensin homolog pseudogene 1* (*PTENP1*) (14), *H19* (15), *HOX transcript antisense RNA* (*HOTAIR*) (16), and so on. Although MALAT1 SNPs such as rs619586 were reported to be associated with the risk or progression of several cancer types and expression of MALAT1 (17), little is known about the effects of polymorphisms of MALAT1 on the development and progression of OSCC, especially in Asian populations. In the present study, a case-control study in a Taiwanese population was performed to identify roles of MALAT1 SNPs in the risk and clinical characteristics of OSCC.

MATERIALS AND METHODS

Selection of Study Subjects

In total, 1350 male patients with OSCC treated at Chung Shan Medical University Hospital (Taichung, Taiwan) between 2007 and 2019 were recruited for this study. All participants provided informed consent. In total, 1199 anonymized healthy controls were randomly selected from the Taiwan Biobank Project; none had a previous history of cancer at any site. Moreover, subjects with oral precancerous disease, including oral submucosal fibrosis, leukoplakia, erythroplakia, verrucous hyperplasia, etc., were excluded from the control group. A questionnaire was

completed by all participants *via* face-to-face interviews to obtain information about the patient's exposure to betel quid chewing, tobacco use, and alcohol consumption. Medical information of OSCC patients was obtained from their medical records including tumor, node, metastasis (TNM) clinical staging, the primary tumor size, lymph node involvement, distal metastasis, and histologic grade. This study was approved by the Ethics Committee of Chung Shan Medical University Hospital (no. CS15125).

Cell Lines and Culture

HSC3 and SAS cell lines were obtained from Japanese Collection of Research Bioresources (JCRB) Cell Bank (Osaka, Japan) and cultured in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12; Life Technologies, Grand Island, NY). CAL27 cell line was obtained from American Type Culture Collection (ATCC) (Manassas, VA, USA) and cultured in DMEM medium (Life Technologies). In addition, OECM1 cell line was cultured in RPMI-1640 medium (Life Technologies). All cell culture media were all supplemented with 10% fetal bovine serum (FBS) (Gibco, Grand Island, NY) and all the cells were maintained at 37°C in a humidified atmosphere of 5% CO_2.

Genomic DNA Extraction and MALAT1 SNP Selection

Venous blood of all participants was placed in tubes containing ethylenediaminetetraacetic acid (EDTA) and further centrifuged at 3000 rpm for buffy coat isolation. DNA from blood leukocytes or OSCC cells was further extracted using a QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. The purity and concentration of DNA were determined with a Nanodrop-2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), with DNA stored at -20°C before genotyping. TagSNPs of lncRNA MALAT1 were selected based on data from the National Center for Biotechnology Information (NCBI) dbSNP database (http://www.ncbi.nlm.nih.gov/) (18) and International HapMap Project (http://hapmap.org). As a result three tagging SNPs (rs3200401, rs7927113, and rs619586) in MALAT1 were selected, which were representative and could capture all other common SNPs. We also included another novel MALAT1 promoter SNP, rs1194338, which was recently reported to be associated with a risk of colorectal carcinoma in a Han Chinese population (19). The reconstructed linkage disequilibrium (LD) plot of these four SNPs is shown in **Figure 1**. We determined one observed haploblock in which rs1194338 was in strong LD with rs619586 in our study. Most importantly, these four SNPs were selected for this study since they were reported to affect risk or progression of various cancer types in the Han Chinese population (20, 21).

Genotyping of MALAT1 SNPs

Candidate SNPs including rs3200401 (assay ID: C_3246069_10), rs619586 (assay ID: C_1060479_10), rs1194338 (assay ID: C_11661801_10), and rs7927113 (assay ID: C_ 29370554_10) were genotyped with a TaqMan SNP Genotyping Assay on an ABI StepOnePlus™ Real-Time PCR platform (Thermo Fisher Scientific). The final collected data were further analyzed with ABI StepOnePlus™ Software v2.3 (Applied Biosystems, Foster City, CA).

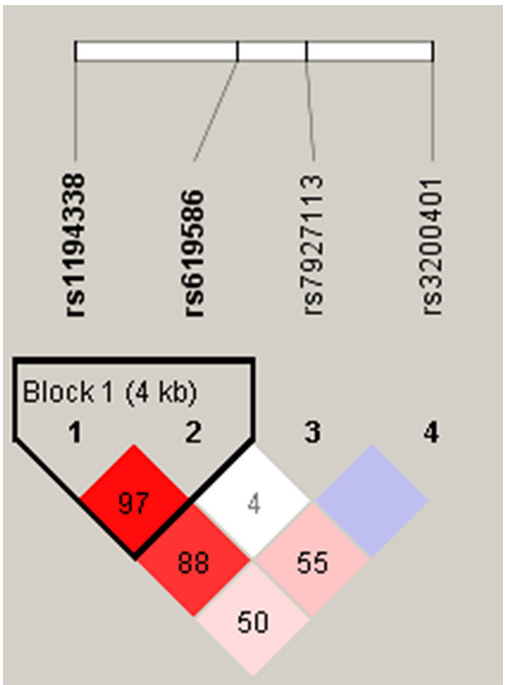


FIGURE 1 | Linkage disequilibrium (LD) map for single nucleotide polymorphisms in the *MALAT1* gene. Block is pairwise D' plots and haplotype blocks obtained from HAPLOVIEW.

Bioinformatics Analysis

Transcriptomic data from Taiwanese samples of adjacent non-tumor tissues and OSCC tissues (n=40) were analyzed using microarray datasets with accession number GSE37991 obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) to investigate whether MALAT1 expression (probe ID: ILMN_2141650) was associated with OSCC tumor formation. Gene expression levels of MALAT1 in normal tissues and head and neck cancerous tissues from Americans were obtained and analyzed from The Cancer Genome Atlas (TCGA)-HNSC database, which was downloaded from the UCSC Xena browser (https://xenabrowser.net/).

Predicting the Structure of the MALAT1/miRNA Duplex

Based on the lncRNASNP2 database (http://bioinfo.life.hust.edu.cn/lncRNASNP2) and the criterion of a minor allele frequency (MAP) of >5% in Asian populations, the targeted rs619586 SNP, NC_000011.10:g.65498698 A>G, was searched in MALAT1 lncRNA. Search results in the section of SNP caused the miRNA target gain represented three miRNAs candidates including hsa-miR-3619-5p, hsa-miR-761 and hsa-miR-214-3p.

We focused on the short transcript form (NR_002019.4, NONHSAT022125.2) of MALAT1 lncRNA for predicting potential lncRNA-miRNA interactions. The minimum free energy (MFE) are -19.5, -22.6 and -18.8 kcal/mol calculated by BiBiserv2 RNAhybrid that G variant on MALAT1 lncRNA could induce target interaction sites gain from miRNAs of hsa-miR-3619-5p, hsa-miR-761 and hsa-miR-214-3p, respectively.

RNA Preparation and SYBR Green Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

Total RNA was isolated from OSCC cells using Total RNA mini kit (Geneaid Biotech Ltd, New Taipei City, Taiwan). Reverse transcription was provided in the High-Capacity cDNA Reverse Transcription Kit (Termo Fisher Scientifc, Waltham, MA, USA), and mRNA expression was detected by qRT-PCR analysis using SYBR Green Fast qPCR System (Termo Fisher Scientifc). GAPDH was selected as the internal reference. The primer sequences were designed as follows: GAPDH-forward (5'-GGAGCGAGATCCCTCCAAAAT-3'), GAPDH-reverse (5'-GGCTGTTGTCATACTTCTCATGG-3'), MALAT1-forward (5'-GATTGAGGAGGCTGTGCTGT-3'), MALAT1-reverse (5'- CAGCTGCCTGCTGTTTTCTG-3'), CTNNB1-forward (5'-GTGCTATCTGTCTGCTCTAGTA-3'), CTNNB1-reverse (5'-CTTCCTGTTTAGTTGCAGCATC-3').

Statistical Analysis

Differences in demographic variables between the cancer-free (control) and OSCC groups were analyzed by the Mann-Whitney *U*-test and Fisher's exact test. The adjusted odds ratios (AORs) and 95% confidence intervals (CIs) of associations of different MALAT1 SNP distributions with OSCC risk or clinical pathological characteristics of OSCC were estimated by multiple logistic regression models after controlling for other covariates, including age, betel quid chewing, cigarette smoking, and alcohol consumption. All statistics were performed using SAS statistical software (vers. 9.1, 2005; SAS Institute, Cary, NC, USA), and $p<0.05$ was considered statistically significant.

RESULTS

Characteristics of Study Participants

Demographic and lifestyle characteristics of the 1350 OSCC cases and 1199 cancer-free controls enrolled in our study were first compared, and results are shown in **Table 1**. The age distributions between the control group and case group were similar. Consistent with previous studies of OSCC patients in Asia (22, 23), higher frequencies of betel nut chewing ($p<0.001$), alcohol consumption ($p<0.001$), and tobacco use ($p<0.001$) were observed in OSCC patients compared to the control group, indicating that these lifestyle characteristics might be critical in the pathogenesis of oral carcinogenesis. The majority of OSCC cases exhibited no lymph node invasion (65.9%) or distal metastasis (99.3%), and their tumors were graded as moderately/poorly differentiated (86.1%).

TABLE 1 | Distributions of demographic characteristics in 1199 healthy controls and 1350 male patients with oral cancer.

Variable	Controls (*N*=1199)	Patients (*N*=1350)	*p* value
Age (years)			
≤55	610 (50.9%)	681 (50.4%)	*p*=0.828
>55	589 (49.1%)	669 (49.6%)	
Betel quid chewing			
No	1000 (83.4%)	343 (25.4%)	
Yes	199 (16.6%)	1007 (74.6%)	*p*<0.001*
Cigarette smoking			
No	564 (47.0%)	210 (15.6%)	
Yes	635 (53.0%)	1140 (84.4%)	*p*<0.001*
Alcohol consumption			
No	962 (80.2%)	712 (52.7%)	
Yes	237 (19.8%)	638 (47.3%)	*p*<0.001*
Stage			
I+II		634 (47.0%)	
III+IV		716 (53.0%)	
Tumor T status			
T1+T2		686 (50.8%)	
T3+T4		664 (49.2%)	
Lymph node status			
N0		890 (65.9%)	
N1+N2+N3		460 (34.1%)	
Metastasis			
M0		1340 (99.3%)	
M1		10 (0.7%)	
Cell differentiation			
Well differentiated		188 (13.9%)	
Moderately or poorly differentiated		1162 (86.1%)	

*The Mann-Whitney U-test or Fisher's exact test was used for comparisons between healthy controls and patients with oral cancer. *Statistically significant at p < 0.05.*

Associations Between MALAT1 Genetic Polymorphisms and OSCC Risks

To determine associations of selected MALAT1 SNPs (rs3200401, rs619586, rs1194338, and rs7927113) with OSCC in this Taiwanese population, we utilized AORs (with 95% CIs) which were estimated by multiple logistic regression models after adjusting for other variables (age, betel nut chewing, cigarette smoking, and alcohol consumption), together with the OR (with 95% CI) of each comparison. As shown in **Table 2**, distributions of MALAT1 genotypes revealed that the most frequent alleles were homozygous C/C, A/A, and G/G for the rs3200401, rs619586, and rs7927113 loci, respectively, and heterozygous C/A for the rs1194338 locus. We observed that subjects with MALAT1 polymorphic rs3200401 T/T and a combination of the C/T and T/T genotypes respectively exhibited significantly lower risks of 0.548- (95% CI: 0.319~0.940) and 0.779-fold (95% CI: 0.632~0.960) of having OSCC compared to those with the C/C wild-type (WT) homozygotes.

Relationships of Clinicopathological Characteristics With MALAT1 Genetic Polymorphisms in OSCC Patients

Since MALAT1 genetic polymorphisms were found to be correlated with susceptibility to OSCC, we further explored the effects of MALAT1 SNPs on the clinical status of OSCC patients, such as the clinical stage, primary tumor size, lymph node

TABLE 2 | Odds ratio (OR) and 95% confidence interval (CI) of oral cancer associated with *MALAT1* genotypic frequencies.

Variable	Controls (N = 1199) (%)	Patients (N = 1350) (%)	OR (95% CI)	AOR (95% CI)[a]
rs3200401				
CC	807 (67.3%)	948 (70.2%)	1.000 (reference)	1.000 (reference)
CT	347 (28.9%)	363 (26.9%)	0.890 (0.748-1.060)	0.872 (0.654-1.010)
				p = 0.061
TT	45 (3.8%)	39 (2.9%)	0.737 (0.475-1.144)	**0.548 (0.319-0.940)**
				***p* = 0.029**
CT+TT	392 (32.7%)	402 (29.8%)	0.873 (0.738-1.032)	**0.779 (0.632-0.960)**
				***p* = 0.019**
rs619586				
AA	1015 (84.7%)	1135 (84.1%)	1.000 (reference)	1.000 (reference)
AG	177 (14.8%)	202 (15.0%)	1.021 (0.820-1.270)	0.995 (0.760-1.304)
				p = 0.973
GG	7 (0.5%)	13 (0.9%)	1.661 (0.660-4.179)	1.113 (0.346-3.578)
				p = 0.858
AG+GG	184 (15.3%)	215 (15.9%)	1.045 (0.843-1.295)	1.000 (0.767-1.304)
				p = 0.998
rs1194338				
CC	505 (42.1%)	588 (43.6%)	1.000 (reference)	1.000 (reference)
CA	544 (45.4%)	625 (46.3%)	0.987 (0.836-1.164)	1.009 (0.822-1.238)
				p = 0.934
AA	150 (12.5%)	137 (10.1%)	0.784 (0.605-1.018)	0.729 (0.527-1.008)
				p = 0.056
CA+AA	694 (57.9%)	762 (56.4%)	0.943 (0.806-1.104)	0.946 (0.779-1.150)
				p = 0.579
rs7927113				
GG	1191 (99.3%)	1338 (99.1%)	1.000 (reference)	1.000 (reference)
GA	8 (0.7%)	12 (0.9%)	1.335 (0.544-3.277)	0.932 (0.297-2.926)
				p = 0.904
AA	-	-	-	-
GA+AA	8 (0.7%)	12 (0.9%)	1.335 (0.544-3.277)	0.932 (0.297-2.926)
				p = 0.904

The odds ratio (OR) with their 95% confidence intervals were estimated by logistic regression models.

[a]*The adjusted odds ratio (AOR) with their 95% confidence intervals were estimated by multiple logistic regression models after controlling for age, betel quid chewing, cigarette smoking, and alcohol drinking.*

Bold values mean the p value is significant ($p < 0.05$).

involvement, metastatic status, and histopathologic grading. We found that patients with at least one minor allele (CT or TT) of rs3200401 exhibited a significantly higher risk of developing moderately (grade II)/poorly (grade III) differentiated OSCC (AOR: 1.508-fold; 95% CI: 1.049~2.169; p=0.027) compared to their counterparts with the corresponding WT homozygotes (**Table 3**). In addition to rs3200401, MALAT1 rs619586 polymorphisms presented significant differences in terms of clinical stage (AOR: 1.358-fold; 95% CI: 1.009~1.827; p=0.044) and tumor size (AOR: 1.429-fold; 95% CI: 1.064~1.919; p=0.018) in OSCC patients with at least one minor allele (AG or GG) (**Table 4**).

Stratified Analysis of MALAT1 Genetic Associations With Clinicopathological Characteristics of OSCC Patients

Actually, chewing betel quid was reported to be the strongest risk factors for causing OSCC in East Asian population (24). Taiwan is also an endemic betel quid-chewing area, and betel quid chewing was reported to be correlated with a poor prognosis in OSCC (25). In this study, we further divided our recruited OSCC patients into betel quid-chewing and non-betel quid-chewing groups, and further investigated the difference between MALAT1 SNPs and the OSCC clinicopathological status in these two subgroups. Compared to the overall OSCC group, development of advance clinical (III+IV) and tumor T stage (>T2) was further strengthened in the betel quid-chewing subgroup who harbored at least one polymorphic G allele of MALAT1 rs619586 (**Table 5**). The risk of developing high-grade OSCC was further heightened in individuals who chewed betel nut and also carried at least one mutant allele of rs3200401 (**Table 6**). In contrast to the betel quid-chewing subgroup, OSCC patients with at least one T allele of MALAT1 rs3200401 were at a lower risk for developing lymph node metastasis if they did not chew betel nuts (**Table 7**). These results suggested that a potential interaction between betel nut chewing and the existence of at least one polymorphic allele of these two MALAT1 SNPs was shown to be correlated with OSCC progression. In addition to betel quid-chewing, two other common environmental carcinogens-alcohol consumption and tobacco use-were also selected to investigate their impacts with MALAT1 SNPs on the clinicopathological status of OSCC. We observed that OSCC patients with the smoking or drinking habit who had at least one G allele of MALAT1 rs619586 were at higher risk of developing advance clinical stage and larger tumor sizes compared to those patients with AA homozygotes (**Supplementary Tables 1**, **2**). Moreover, OSCC patients with

TABLE 3 | Adjusted odds ratios (AORs) and 95% confidence intervals (CIs) of clinical statuses associated with genotypic frequencies of *MALAT1* rs3200401 in male oral cancer patients (*N*=1350).

Variable			AOR (95% CI)	*p* value
	Clinical stage			
rs3200401	Stage I+II	Stage III+IV		
	(*n*=634) (%)	(*n*=716) (%)		
CC	429 (67.7%)	519 (72.5%)	1.00	
CT+TT	205 (32.3%)	197 (27.5%)	0.796 (0.630~1.007)	*p*=0.057
	Tumor size			
rs3200401	≤T2	>T2		
	(*n*=686) (%)	(*n*=664) (%)		
CC	484 (70.5%)	464 (69.9%)	1.00	
CT+TT	202 (29.5%)	200 (30.1%)	1.023 (0.809~1.293)	*p*=0.849
	Lymph node metastasis			
rs3200401	No	Yes		
	(*n*=890) (%)	(*n*=460) (%)		
CC	613 (68.9%)	335 (72.8%)	1.00	
CT+TT	277 (31.1%)	125 (27.2%)	0.829 (0.645~1.066)	*p*=0.143
	Metastasis			
rs3200401	M0	M1		
	(*n*=1340) (%)	(*n*=10) (%)		
CC	942 (70.3%)	6 (60.0%)	1.00	
CT+TT	398 (29.7%)	4 (40.0%)	1.619 (0.452~5.795)	*p*=0.459
	Cell differentiation grade			
rs3200401	≤Grade I	>Grade I		
	(*n*=188) (%)	(*n*=1162) (%)		
CC	145 (77.1%)	803 (69.1%)	1.00	
CT+TT	43 (22.9%)	359 (30.9%)	**1.508 (1.049~2.169)**	***p*=0.027***

Cell differentiation grade: grade I, well differentiated; grade II, moderately differentiated; grade III, poorly differentiated.
*The AORs with their 95% CIs were estimated by multiple logistic regression models after controlling for age, betel quid chewing, cigarette smoking, and alcohol consumption. *Statistically significant at p < 0.05.*
Bold values mean the p value is significant (p < 0.05).

the smoking or drinking habit who had at least one T allele of MALAT1 rs3200401 were at higher risk of developing high-grade OSCC compared to those patients with CC homozygotes (**Supplementary Tables 3**, **4**). Compared to the overall OSCC group, the risk of OSCC within advanced-stage development were strengthened in the alcohol consumption or tobacco use subgroups who harbored at least one polymorphic G allele of MALAT1 rs619586 or T allele of MALAT1 rs3200401 (**Supplementary Tables 1**–**4**).

MALAT1 Expression Is Upregulated in OSCC, Especially in Patients With the Habit of Chewing Betel Quid

To further dissect expression levels of MALAT1 and their clinical significance in oral cancer, cases of head and neck squamous cell carcinomas (HNSCC) were analyzed from TCGA dataset. Compared to noncancerous tissues, MALAT1 expression was prone to be upregulated in HNSCC (p=0.065) (**Figure 2A**). Interestingly, from the GSE37991 dataset of the GEO database, we further observed that MALAT1 expression levels were significantly higher in OSCC specimens than their corresponding matched normal tissues from an OSCC N/T paired Taiwanese cohort with the habit of betel quid chewing (p=0.0014) (**Figure 2B**). Moreover, relative levels of MALAT1 transcripts were higher in HNSCC patients with larger tumors (T4 status) than in patients with smaller tumors (T1, T2, or T3 status) (**Figure 2C**).

DISCUSSION

Recently, lncRNAs were shown to play important roles in cancer development, and a large number of lncRNAs associated with multiple cancers were identified, including OSCC (26, 27). Oncogenic roles of lncRNA MALAT1 in OSCC were previously reported, including promotion of growth, metastasis, and chemoresistance of OSCC through inducing the epithelial-to-mesenchymal transition (EMT) and sponging tumor-suppressive miRNAs such as miR-143, miR-125b, and miR-101 (10, 11, 28–30). However, OSCC risks and clinicopathologic characteristics conferred by genetic variants on loci of MALAT1 have rarely been the focus of epidemiological investigations.

In the present molecular epidemiology study conducted in 1350 OSCC patients and 1199 healthy controls, we found that mutant base T of rs3200401 was significantly associated with a lower risk of OSCC, regardless of the codominant model (TT) or dominant model (CT+TT). Similar to our results, previous studies showed that the risk of prostate cancer (PCa) development in a Ukrainian population was significantly lower in those with the rs3200401 TT genotype compared to CC genotype (31). In breast cancer, the CT genotype of rs3200401 imparted a lower risk of breast cancer compared to CC in a Han Chinese population (20). In addition, Wang et al. indicated that advanced lung cancer patients with the rs3200401 CT and TT genotypes had significantly longer overall survival times than did patients with the CC genotype (32). Volkogon et al. found that bladder cancer patients with the

TABLE 4 | Adjusted odds ratios (AORs) and 95% confidence intervals (CIs) of clinical statuses associated with genotypic frequencies of *MALAT1* rs619586 in male oral cancer patients (*N*=1350).

Variable			AOR (95% CI)	*p* value
	Clinical stage			
rs619586	Stage I+II	Stage III+IV		
	(*n*=634) (%)	(*n*=716) (%)		
AA	546 (86.1%)	589 (82.3%)	1.00	
AG+GG	88 (13.9%)	127 (17.7%)	**1.358 (1.009~1.827)**	***p*=0.044***
	Tumor size			
rs619586	≤T2	>T2		
	(*n*=686) (%)	(*n*=664) (%)		
AA	593 (86.4%)	542 (81.6%)	1.00	
AG+GG	93 (13.6%)	122 (18.4%)	**1.429 (1.064~1.919)**	***p*=0.018***
	Lymph node metastasis			
rs619586	No	Yes		
	(*n*=890) (%)	(*n*=460) (%)		
AA	748 (84.0%)	387 (84.1%)	1.00	
AG+GG	142 (16.0%)	73 (15.9%)	1.005 (0.738~1.370)	*p*=0.973
	Metastasis			
rs619586	M0	M1		
	(*n*=1340) (%)	(*n*=10) (%)		
AA	1125 (84.0%)	10 (100.0%)	1.00	
AG+GG	215 (16.0%)	0 (40.0%)	-	-
	Cell differentiation grade			
rs619586	≤Grade I	>Grade I		
	(*n*=188) (%)	(*n*=1162) (%)		
AA	159 (84.6%)	976 (84.0%)	1.00	
AG+GG	29 (15.4%)	186 (16.0%)	1.043 (0.681~1.599)	*p*=0.846

Cell differentiation grade: grade I, well differentiated; grade II, moderately differentiated; grade III, poorly differentiated.
*The AORs with their 95% CIs were estimated by multiple logistic regression models after controlling for age, betel quid chewing, cigarette smoking, and alcohol consumption. *Statistically significant at p < 0.05.*
Bold values mean the p value is significant (p < 0.05).

rs3200401 TT genotype had significantly longer disease-free survival times than did patients with the C allele (33). Taken together, results of our study were accordant with previous studies mentioned above and revealed that the T allele of rs3200401 may have a protective role against the development of cancer. Wang et al. showed that the rs3200401 C>T nucleotide replacement led to 1.62 kcal/mol MFE change, which alters MALAT1's spatial structure and impairs its interaction with serine/arginine-rich splicing factor 2 (SRSF2) (32). Violation of interactions between MALAT1 and SRSF2 ultimately lead to inhibition of pre-mRNA alternative splicing and expressions of genes involved in cancer development, and such an effect might explain both the decreased tumor aggression activity and better survival rates in patients with various cancer types who are minor T-allele carriers (32).

In contrast to the protective effect against malignant tumor development, the rs3200401 TT genotype was reported to increase the risk of esophageal squamous cell carcinoma (ESCC) and tumor width of bladder cancer compared to the rs3200401 CC genotype (34). Our present results showed that OSCC patients with at least one minor allele (CT or TT) of rs3200401 exhibited a significantly higher risk of developing high-grade tumors, especially in the subgroup of betel quid chewers. Detailed mechanisms of the opposite role of rs32000401 polymorphisms in different cancers remain to be identified. We hypothesized that environmental carcinogens might be the potential effector driving the protective or oncogenic role of rs32000401 polymorphisms in cancers, because our study showed that OSCC patients harboring at least one T allele of MALAT1 rs3200401 were at a lower risk of developing lymph node metastasis if they did not chew betel nut. The interaction between betel quid chewing and rs32000401 polymorphisms on OSCC progression should be further investigated in future work.

In addition to rs3200401, OSCC patients harboring at least one polymorphic G allele of MALAT1 rs619586 had a higher frequency of developing advance clinical and tumor T stages, and this phenomenon was further strengthened in the betel quid-chewing subgroup. Similar to our results, Wang et al. indicated that patients with differentiated thyroid carcinoma (DTC) carrying the AG/GG genotypes of the MALAT1 rs619586 polymorphism exhibited higher tumor grades and shorter survival times compared to AA-genotype patients (35). Interestingly, most previous studies indicated that the G allele of rs619586 could significantly decrease MALAT1 expression in different cancer types such as DTC (35), breast cancer (20), and papillary thyroid cancer (PTC) (36). MALAT1 was reported to play an oncogenic role in many solid tumors including OSCC (10, 11), and we actually observed that MALAT1 expression was prone to be upregulated in HNSCC from the Western society and significantly upregulated in OSCC specimens from Taiwanese cohort, especially in patients who chewed betel quid. This discrepancy may be due to the differences of genetic backgrounds and pathogenic mechanisms in head and neck cancer between Asian and Western societies. For example, chewing betel quid is the strongest risk factors for causing OSCC in East Asian population (24). In contrast, the tobacco smoking is the most-established risk factor for head and neck

TABLE 5 | Adjusted odds ratios (AORs) and 95% confidence intervals (CIs) of clinical statuses associated with genotypic frequencies of *MALAT1* rs619586 in male oral cancer patients who chewed betel nuts (*N*=1007).

Variable			AOR (95% CI)	*p* value
	Clinical stage			
rs619586	Stage I+II	Stage III+IV		
	(*n*=480) (%)	(*n*=527) (%)		
AA	414 (86.3%)	430 (81.6%)	1.00	
AG+GG	66 (13.7%)	97 (18.4%)	**1.419 (1.009~1.997)**	***p*=0.045***
	Tumor size			
rs619586	≤T2	>T2		
	(*n*=519) (%)	(*n*=488) (%)		
AA	448 (86.3%)	396 (81.2%)	1.00	
AG+GG	71 (13.7%)	92 (18.8%)	**1.452 (1.035~2.038)**	***p*=0.031***
	Lymph node metastasis			
rs619586	No	Yes		
	(*n*=675) (%)	(*n*=332) (%)		
AA	565 (83.7%)	279 (84.0%)	1.00	
AG+GG	110 (16.3%)	53 (16.0%)	0.972 (0.678~1.391)	*p*=0.875
	Metastasis			
rs619586	M0	M1		
	(*n*=1000) (%)	(*n*=7) (%)		
AA	837 (83.7%)	7 (100.0%)	1.00	
AG+GG	163 (16.3%)	0 (0.0%)	-	-
	Cell differentiation grade			
rs619586	≤Grade I	>Grade I		
	(*n*=151) (%)	(*n*=856) (%)		
AA	125 (82.8%)	719 (84.0%)	1.00	
AG+GG	26 (17.2%)	137 (16.0%)	0.901 (0.568~1.429)	*p*=0.657

Cell differentiation grade: grade I, well differentiated; grade II, moderately differentiated; grade III, poorly differentiated.
*The AORs with their 95% CIs were estimated by multiple logistic regression models after controlling for age, cigarette smoking, and alcohol consumption. *Statistically significant at p < 0.05. Bold values mean the p value is significant (p < 0.05).*

TABLE 6 | Adjusted odds ratios (AORs) and 95% confidence intervals (CIs) of clinical statuses associated with genotypic frequencies of *MALAT1* rs3200401 in male oral cancer patients who chewed betel nuts (*N*=1007).

Variable			AOR (95% CI)	*p* value
	Clinical stage			
rs3200401	Stage I+II	Stage III+IV		
	(*n*=480) (%)	(*n*=527) (%)		
CC	326 (67.9%)	374 (71.0%)	1.00	
CT+TT	154 (32.1%)	153 (29.0%)	0.861 (0.658~1.127)	*p*=0.277
	Tumor size			
rs3200401	≤T2	>T2		
	(*n*=519) (%)	(*n*=448) (%)		
CC	359 (69.2%)	341 (69.9%)	1.00	
CT+TT	160 (30.8%)	147 (30.1%)	0.960 (0.733~1.256)	*p*=0.764
	Lymph node metastasis			
rs3200401	No	Yes		
	(*n*=675) (%)	(*n*=332) (%)		
CC	470 (69.6%)	230 (69.3%)	1.00	
CT+TT	205 (30.4%)	102 (30.7%)	1.006 (0.756~1.340)	*p*=0.966
	Metastasis			
rs3200401	M0	M1		
	(*n*=1000) (%)	(*n*=7) (%)		
CC	696 (69.6%)	4 (57.1%)	1.00	
CT+TT	304 (30.4%)	3 (42.9%)	1.774 (0.393~7.997)	*p*=0.456
	Cell differentiation grade			
rs3200401	≤Grade I	>Grade I		
	(*n*=151) (%)	(*n*=856) (%)		
CC	117 (77.5%)	583 (68.1%)	1.00	
CT+TT	34 (22.5%)	273 (31.9%)	**1.588 (1.055~2.390)**	***p*=0.027***

Cell differentiation grade: grade I, well differentiated; grade II, moderately differentiated; grade III, poorly differentiated.
*The AORs with their 95% CIs were estimated by multiple logistic regression models after controlling for age, cigarette smoking, and alcohol consumption. *Statistically significant at p < 0.05. Bold values mean the p value is significant (p < 0.05).*

TABLE 7 | Adjusted odds ratios (AORs) and 95% confidence intervals (CIs) of clinical statuses associated with genotypic frequencies of *MALAT1* rs3200401 in male oral cancer patients who did not chew betel nuts (*N*=343).

Variable			AOR (95% CI)	*p* value
	Clinical stage			
rs3200401	Stage I+II	Stage III+IV		
	(*n*=154) (%)	(*n*=189) (%)		
CC	103 (66.9%)	145 (76.7%)	1.00	
CT+TT	51 (33.1%)	44 (23.3%)	0.618 (0.382~1.001)	*p*=0.051
	Tumor size			
rs3200401	≤T2	>T2		
	(*n*=167) (%)	(*n*=176) (%)		
CC	125 (74.9%)	123 (69.9%)	1.00	
CT+TT	42 (25.1%)	53 (30.1%)	1.254 (0.777~2.024)	*p*=0.355
	Lymph node metastasis			
rs3200401	No	Yes		
	(*n*=215) (%)	(*n*=128) (%)		
CC	143 (66.5%)	105 (82.0%)	1.00	
CT+TT	72 (33.5%)	23 (18.0%)	**0.437 (0.255-0.749)**	***p*=0.003***
	Metastasis			
rs3200401	M0	M1		
	(*n*=340) (%)	(*n*=3) (%)		
CC	246 (72.4%)	2 (66.7%)	1.00	
CT+TT	94 (27.6%)	1 (33.3%)	1.324 (0.117~14.989)	*p*=0.821
	Cell differentiation grade			
rs3200401	≤Grade I	>Grade I		
	(*n*=37) (%)	(*n*=306) (%)		
CC	28 (75.7%)	220 (71.9%)	1.00	
CT+TT	9 (24.3%)	86 (28.1%)	1.198 (0.540~2.658)	*p*=0.656

Cell differentiation grade: grade I, well differentiated; grade II, moderately differentiated; grade III, poorly differentiated.
*The AORs with their 95% CIs were estimated by multiple logistic regression models after controlling for age, cigarette smoking, and alcohol consumption. *Statistically significant at p < 0.05. Bold values mean the p value is significant (p < 0.05).*

cancer in the Western society (37), suggesting the various pathogenic mechanisms in the different ethnic cohorts may lead the different MALAT1 expression status. Herein, our results showed that the G allele of rs619586 might trigger lower MALAT1 expression but was correlated with advance clinical and tumor T stages. Recently, a growing number of reports suggested that MALAT1 acts as a 'sponge' to bind specific miRNAs and upregulate miRNAs' targets to modulate cancer progression. For example, MALAT1 promoted OSCC development *via* sponging miR-125b and subsequently upregulating miR-125b's target gene, *STAT3* (10). Moreover, MALAT1 induced OSCC invasion by negatively regulating miR-101 and upregulating miR-101's target gene, *EZH2* (11). Large numbers of SNPs were predicted to have potential impacts on miRNA-lncRNA interactions (38). For example, the G variant of the rs619586 SNP was reported to interact with mir-214 and upregulate its target gene, *CTNNB1* (encoding β-catenin) to promote proliferation of DTC (35). Actually, miR-214 and

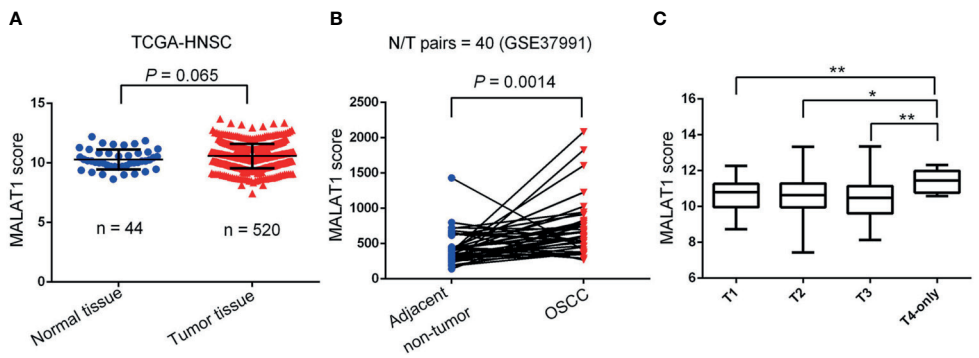


FIGURE 2 | Clinical relevance of metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) levels in head and neck squamous cell carcinomas (HNSCC) patients obtained from TCGA and GEO databases. **(A)** *MALAT1* gene expression levels in normal and HNSCC tissues were compared according to data from TCGA datasets. Statistical significance was analyzed by a *t*-test. **(B)** *MALAT1* gene expression levels in an oral squamous cell carcinoma normal/tumorous (N/T) paired cohort with the habit of betel quid chewing (GSE37991). Statistical significance was analyzed by a paired *t*-test. **(C)** *MALAT1* gene expression levels in HNSCC from TCGA were compared according to the tumor size (T stages). Statistical significance was analyzed by a *t*-test. *$p < 0.05$, **$p < 0.01$.

CTNNB1 were respectively reported as having tumor suppressive and oncogenic roles in OSCC (39, 40). By estimating the MFE of the miRNA/MALAT1 duplex, we observed that the A allele of rs619586 led to a far worse energy of miR-214 hybridization of −12.2 kcal/mole than did the G allele (−18.8 kcal/mole) (**Figure 3A**), suggesting that the G allele of MALAT1 rs619586 enhances miR-214 binding to MALAT1. We next examined rs619586 genotypes of four OSCC cell lines (CAL27, HSC3, OECM1, and SAS) and found that OECM1 cells carrying AG genotype of rs619586 respectively expressed lower MALAT1 and higher CTNNB1 levels compared to CAL27 and HSC3 cells carrying AA genotype of rs619586 (**Figures 4A, B**). Furthermore, we analyzed 523 HNSCC human samples that were retrieved from TCGA by using the cBioportal platform and observed that MALAT1 expression inversely correlated with CTNNB1 (**Figure 5**). These phenomena suggested that although the G allele of rs619586 may cause a decrease in MALAT1 levels, this MALAT1 SNP can sponge miR-214 to upregulate CTNNB1 and further promote OSCC progression. In addition to miR-214, the association between miR-761 or miR-3619-5p with the G variant of rs619586 SNP was also predicted by BiBiServ2 RNAhybrid (https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid), and the G allele had better binding stability to these two miRNAs (**Figures 3B, C**). Interactions of different lncRNAs with miR-761 or miR-3619-5p were reported to promote progression or chemoresistance of cancers. For example, the lncRNA, HOXA11-AS, was reported to be a molecular sponge that binds to miR-761 and subsequently upregulates miR-761's target of *tripartite motif-containing protein 29* (*TRIM29*) to promote progression of PTC (41). Moreover, the lncRNA, KCNQ1OT1, was reported to confer chemoresistance in gliomas *via* sponging miR-761 and subsequently upregulating peripheral interface module 1 (PIM1) (42). The exosomal lncRNA, HEIH, was shown to confer cisplatin resistance in OSCC by sponging miR-3619-5p and upregulating hepatoma-derived growth factor (HDGF) (43). According to those results, we suggest that MALAT1 might sponge miR-214, miR-761,

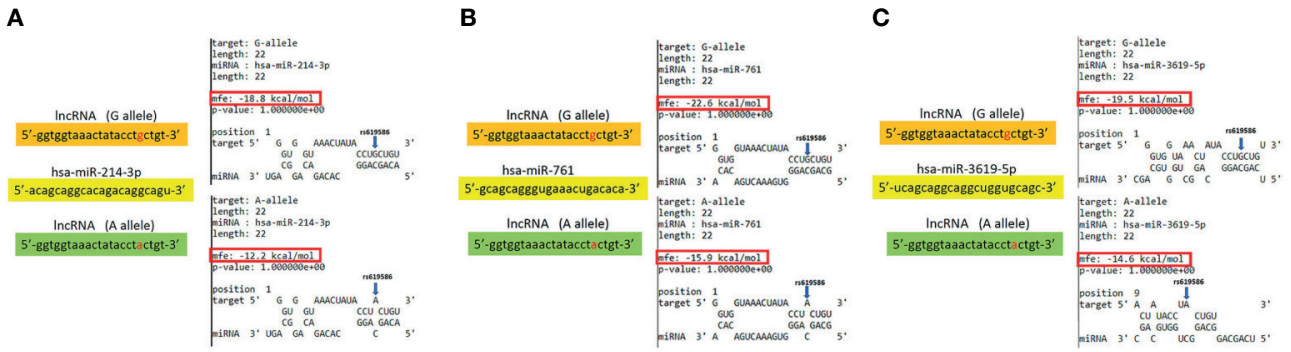


FIGURE 3 | Prediction of potential binding between metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) and its interacting microRNAs. The schematic diagram exhibits hybridization between MALAT1 harboring the rs619586 A or G allele and miR-214 **(A)**, miR-761**(B),** or miR-3619-5p **(C)**. The single-nucleotide polymorphism (SNP) rs619586 A allele reduced the binding affinities with these miRNAs. The positions of rs619586 SNPs are indicated by blue arrows. MFE, minimum free energy.

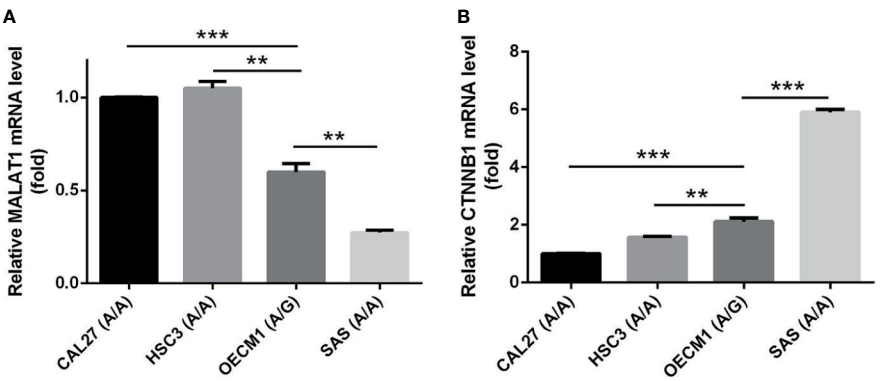


FIGURE 4 | The genotypes of *MALAT1* rs619586 and mRNA levels of *MALAT1* **(A)** and *CTNNB1* **(B)** in four oral squamous cell carcinoma (OSCC) cells (CAL27, HSC3, OECM1, and SAS) were detected by TaqMan SNP Genotyping Assay and RT-qPCR, respectively. Quantitative results of *MALAT1* and *CTNNB1* mRNA levels were adjusted to *GAPDH* mRNA levels. Values are presented as the mean ± standard deviation (SD) of three independent experiments. $^{**}p < 0.01$ and $^{***}p < 0.001$ compared to indicated mRNA levels in OECM1 cells.

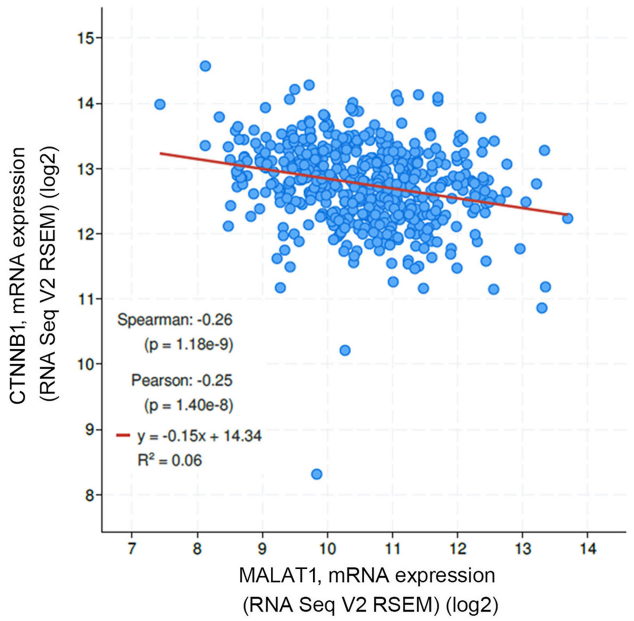


FIGURE 5 | Correlation analysis of TCGA head and neck squamous cell carcinoma (HNSCC) database (TCGA, PanCancer Atlas) using the cBioPortal showed an inverse correlation of MALAT1 with β-catenin (CTNNB1) mRNA levels.

or miR-3619-5p to promote OSCC development, and MALAT1 rs619586 SNPs may influence the interaction of MALAT1 with these miRNAs.

CONCLUSIONS

In conclusion, we first identified the diverse allelic effects of MALAT1 SNPs (rs3200401 and rs619586) which contribute to the susceptibility and clinicopathologic development of OSCC in a Taiwanese population. In addition, the combined effect of MALAT1 SNPs (rs3200401 and rs619586) with betel nut chewing causally contributes to the development of OSCC and this phenomenon was also observed in OSCC patients with the smoking or drinking habit. Furthermore, the rs619586 G variant may enhance the binding of miR-214, miR-761, or miR-3619-5p to MALAT1 to promote OSCC progression. However, these issues should be further confirmed in future studies, and downstream targets of these miRNAs in regulating OSCC progression also need to be further investigated.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

This study was approved (no. CS15125) by the Institutional Review Board of Chung Shan Medical University Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Y-FD and M-HC contributed to conception, drafted the manuscript, and critically revised the manuscript. S-FY, C-WL, and Y-CW contributed to conception, performed the experiments, and analyzed data. Y-CY, C-YC, Y-FL, L-CC, and W-MC performed the in silico data exploration, integration, and analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from Wan Fang Hospital, Taipei Medical University (110-wf-swf-06 to Y-CW and M-HC; 110-wf-phd-03 to Y-FD) and the TMU Research Center of Cancer Translational Medicine from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan (to M-HC).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.684941/full#supplementary-material

REFERENCES

1. Jiang X, Wu J, Wang J, Huang R. Tobacco and Oral Squamous Cell Carcinoma: A Review of Carcinogenic Pathways. *Tob Induc Dis* (2019) 17:29. doi: 10.18332/tid/105844
2. Kumar R, Rai AK, Das D, Das R, Kumar RS, Sarma A, et al. Alcohol and Tobacco Increases Risk of High Risk HPV Infection in Head and Neck Cancer Patients: Study From North-East Region of India. *PloS One* (2015) 10: e0140700. doi: 10.1371/journal.pone.0140700
3. Chen YK, Huang HC, Lin LM, Lin CC. Primary Oral Squamous Cell Carcinoma: An Analysis of 703 Cases in Southern Taiwan. *Oral Oncol* (1999) 35:173–9. doi: 10.1016/S1368-8375(98)00101-8
4. Argiris A, Karamouzis MV, Raben D, Ferris RL. Head and Neck Cancer. *Lancet* (2008) 371:1695–709. doi: 10.1016/S0140-6736(08)60728-X
5. Tsai MC, Spitale RC, Chang HY. Long Intergenic Noncoding RNAs: New Links in Cancer Progression. *Cancer Res* (2011) 71:3–7. doi: 10.1158/0008-5472.CAN-10-2483
6. Teppan J, Barth DA, Prinz F, Jonas K, Pichler M, Klec C. Involvement of Long Non-Coding RNAs (lncRNAs) in Tumor Angiogenesis. *Noncoding RNA* (2020) 6:42. doi: 10.3390/ncrna6040042
7. Li ZX, Zhu QN, Zhang HB, Hu Y, Wang G, Zhu YS. MALAT1: A Potential Biomarker in Cancer. *Cancer Manag Res* (2018) 10:6757–68. doi: 10.2147/CMAR.S169406
8. Arun G, Aggarwal D, Spector DL. Malat1 Long Non-Coding RNA: Functional Implications. *Noncoding RNA* (2020) 6:22. doi: 10.3390/ncrna6020022
9. Ji P, Diederichs S, Wang W, Böing S, Metzger R, Schneider PM, et al. MALAT-1, a Novel Noncoding RNA, and Thymosin Beta4 Predict Metastasis

and Survival in Early-Stage Non-Small Cell Lung Cancer. *Oncogene* (2003) 22:8031–41. doi: 10.1038/sj.onc.1206928
10. Chang SM, Hu WW. Long Non-Coding RNA MALAT1 Promotes Oral Squamous Cell Carcinoma Development Via microRNA-125b/STAT3 Axis. *J Cell Physiol* (2018) 233:3384–96. doi: 10.1002/jcp.26185
11. Xiao L, Wang W, Zhao J, Xu H, Li S, Yang X. LncRNA MALAT1 Promotes Cell Proliferation and Invasion by Regulating the miR-101/EZH2 Axis in Oral Squamous Cell Carcinoma. *Oncol Lett* (2020) 20:164. doi: 10.3892/ol.2020.12024
12. Lv Z, Xu Q, Yuan Y. A Systematic Review and Meta-Analysis of the Association Between Long Non-Coding RNA Polymorphisms and Cancer Risk. *Mutat Res* (2017) 771:1–14. doi: 10.1016/j.mrrev.2016.10.002
13. Hou Y, Zhang B, Miao L, Ji Y, Yu Y, Zhu L, et al. Association of Long Non-Coding RNA MEG3 Polymorphisms With Oral Squamous Cell Carcinoma Risk. *Oral Dis* (2019) 25:1318–24. doi: 10.1111/odi.13103
14. Xin C, Li JL, Zhang YX, Yu ZH. Polymorphisms in lncRNA PTENP1 and the Risk of Oral Squamous Cell Carcinoma in a Chinese Population. *Eur Rev Med Pharmacol Sci* (2018) 22:5583–7. doi: 10.26355/eurrev_201809_15822
15. Hashemi M, Moazeni-Roodi A, Sarabandi S, Karami S, Ghavami S. Association Between Genetic Polymorphisms of Long Noncoding RNA H19 and Cancer Risk: A Meta-Analysis. *J Genet* (2019) 98:81. doi: 10.1007/s12041-019-1126-x
16. Su SC, Hsieh MJ, Lin CW, Chuang CY, Liu YF, Yeh CM, et al. Impact of HOTAIR Gene Polymorphism and Environmental Risk on Oral Cancer. *J Dent Res* (2018) 97:717–24. doi: 10.1177/0022034517749451
17. Ni W, Wang X, Sun Y, Gao X. Meta-Analysis of the Association Between MALAT1 Rs619586 A>G Polymorphism and Cancer Risk. *J Int Med Res* (2020) 48:300060520941969. doi: 10.1177/0300060520941969
18. Smigielski EM, Sirotkin K, Ward M, Sherry ST. dbSNP: A Database of Single Nucleotide Polymorphisms. *Nucleic Acids Res* (2000) 28:352–5. doi: 10.1093/nar/28.1.352
19. Li Y, Bao C, Gu S, Ye D, Jing F, Fan C, et al. Associations Between Novel Genetic Variants in the Promoter Region of MALAT1 and Risk of Colorectal Cancer. *Oncotarget* (2017) 8:92604–14. doi: 10.18632/oncotarget.21507
20. Peng R, Luo C, Guo Q, Cao J, Yang Q, Dong K, et al. Association Analyses of Genetic Variants in Long Non-Coding RNA MALAT1 With Breast Cancer Susceptibility and mRNA Expression of MALAT1 in Chinese Han Population. *Gene* (2018) 642:241–8. doi: 10.1016/j.gene.2017.11.013
21. Zhao K, Jin S, Wei B, Cao S, Xiong Z. Association Study of Genetic Variation of lncRNA MALAT1 With Carcinogenesis of Colorectal Cancer. *Cancer Manag Res* (2018) 10:6257–61. doi: 10.2147/CMAR.S177244
22. Goud E, Malleedi S, Ramanathan A, Wong GR, Hwei Ern BT, Yean GY, et al. Association of Interleukin-10 Genotypes and Oral Cancer Susceptibility in Selected Malaysian Population: A Case- Control Study. *Asian Pac J Cancer Prev* (2019) 20:935–41. doi: 10.31557/APJCP.2019.20.3.935
23. Senghore T, Chien HT, Wang WC, Chen YX, Young CK, Huang SF, et al. Polymorphisms in ERCC5 rs17655 and ERCC1 Rs735482 Genes Associated With the Survival of Male Patients With Postoperative Oral Squamous Cell Carcinoma Treated With Adjuvant Concurrent Chemoradiotherapy. *J Clin Med* (2019) 8:33. doi: 10.3390/jcm8010033
24. Lee YA, Li S, Chen Y, Li Q, Chen CJ, Hsu WL, et al. Tobacco Smoking, Alcohol Drinking, Betel Quid Chewing, and the Risk of Head and Neck Cancer in an East Asian Population. *Head Neck* (2019) 41:92–102. doi: 10.1002/hed.25383
25. Yang J, Wang ZY, Huang L, Yu TL, Wan SQ, Song J, et al. Do Betel Quid and Areca Nut Chewing Deteriorate Prognosis of Oral Cancer? A Systematic Review, Meta-Analysis, and Research Agenda. *Oral Dis* (2020). doi: 10.1111/odi.13456
26. Jiang MC, Ni JJ, Cui WY, Wang BY, Zhuo W. Emerging Roles of lncRNA in Cancer and Therapeutic Opportunities. *Am J Cancer Res* (2019) 9:1354–66.
27. Zhang L, Meng X, Zhu XW, Yang DC, Chen R, Jiang Y, et al. Long Non-Coding RNAs in Oral Squamous Cell Carcinoma: Biologic Function, Mechanisms and Clinical Implications. *Mol Cancer* (2019) 18:102. doi: 10.1186/s12943-019-1021-3
28. Wang R, Lu X, Yu R. lncRNA MALAT1 Promotes EMT Process and Cisplatin Resistance of Oral Squamous Cell Carcinoma Via PI3K/AKT/m-TOR Signal Pathway. *Onco Targets Ther* (2020) 13:4049–61. doi: 10.2147/OTT.S251518
29. Yu L, Shao X, Huo L, Zhang T. Long Non-Coding RNA (lncRNA) Metastasis-Associated Lung Adenocarcinoma Transcript 1 (MALAT1) Promotes Cell Proliferation and Migration by Regulating miR-143-3p and MAGE Family Member A9 (MAGEA9) in Oral Squamous Cell Carcinoma. *Med Sci Monit* (2020) 26:e924187. doi: 10.12659/MSM.924187
30. Zhou X, Liu S, Cai G, Kong L, Zhang T, Ren Y, et al. Long Non Coding RNA MALAT1 Promotes Tumor Growth and Metastasis by Inducing Epithelial-Mesenchymal Transition in Oral Squamous Cell Carcinoma. *Sci Rep* (2015) 5:15972. doi: 10.1038/srep15972
31. Volkogon A, Chumachenko Y, Harbuzova V, Ataman A. Analysis of Association Between Rs3200401 Long NonCoding RNA MALAT1 Gene Polymorphism and Prostate Adenocarcinoma Development in Ukrainian Population. *J Urol Nephrol Stud* (2019) 1:99–102. doi: 10.29254/2077-4214-2019-2-1-150-109-112
32. Wang JZ, Xiang JJ, Wu LG, Bai YS, Chen ZW, Yin XQ, et al. A Genetic Variant in Long Non-Coding RNA MALAT1 Associated With Survival Outcome Among Patients With Advanced Lung Adenocarcinoma: A Survival Cohort Analysis. *BMC Cancer* (2017) 17:167. doi: 10.1186/s12885-017-3151-6
33. Volkogon A, Kolnoguz O, Harbuzova V, Ataman A. Long Non-Coding RNA MALAT1 Gene Polymorphismis Associated With Disease-Free Survival in Bladder Cancer Patients. *Galician Med J* (2020) 27:E202025. doi: 10.21802/gmj.2020.2.5
34. Qu Y, Shao N, Yang W, Wang J, Cheng Y. Association of Polymorphisms in MALAT1 With the Risk of Esophageal Squamous Cell Carcinoma in a Chinese Population. *Onco Targets Ther* (2019) 12:2495–503. doi: 10.2147/OTT.S191155
35. Wang ML, Liu JX. MALAT1 rs619586 Polymorphism Functions as a Prognostic Biomarker in the Management of Differentiated Thyroid Carcinoma. *J Cell Physiol* (2020) 235:1700–10. doi: 10.1002/jcp.29089
36. Wen J, Chen L, Tian H, Li J, Zhang M, Cao Q, et al. Effect of MALAT1 Polymorphisms on Papillary Thyroid Cancer in a Chinese Population. *J Cancer* (2019) 10:5714–21. doi: 10.7150/jca.28887
37. Hashim D, Genden E, Posner M, Hashibe M, Boffetta P. Head and Neck Cancer Prevention: From Primary Prevention to Impact of Clinicians on Reducing Burden. *Ann Oncol* (2019) 30:744–56. doi: 10.1093/annonc/mdz084
38. Gong J, Liu W, Zhang J, Miao X, Guo AY. lncRNASNP: A Database of SNPs in lncRNAs and Their Potential Functions in Human and Mouse. *Nucleic Acids Res* (2015) 43:D181–6. doi: 10.1093/nar/gku1000
39. Wang X, Li H, Shi J. LncRNA HOXA11-AS Promotes Proliferation and Cisplatin Resistance of Oral Squamous Cell Carcinoma by Suppression of miR-214-3p Expression. *BioMed Res Int* (2019) 2019:8645153. doi: 10.1155/2019/8645153
40. Reyes M, Flores T, Betancur D, Peña-Oyarzún D, Torres VA. Wnt/β-Catenin Signaling in Oral Carcinogenesis. *Int J Mol Sci* (2020) 21:4682. doi: 10.3390/ijms21134682
41. Yin X, Zhang J, Li C, Zhang Z, Jin T, Song L, et al. LncRNA HOXA11-AS Accumulation-Induced microRNA-761 Downregulation Regulates Cell Growth by Targeting TRIM29 in Papillary Thyroid Cancer. *Am J Transl Res* (2019) 11:6826–37.
42. Wang W, Han S, Gao W, Feng Y, Li K, Wu D. Long Noncoding RNA KCNQ1OT1 Confers Gliomas Resistance to Temozolomide and Enhances Cell Growth by Retrieving PIM1 From miR-761. *Cell Mol Neurobiol* (2020). doi: 10.1007/s10571-020-00958-4
43. Wang X, Yu H, Yu Z, Wang D. Exosomal LncRNA HEIH Promotes Cisplatin Resistance in Tongue Squamous Cell Carcinoma Via Targeting miR-3619-5p/HDGF Axis. *Acta Histochem* (2020) 122:151647. doi: 10.1016/j.acthis.2020.151647

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 15 July 2021
doi: 10.3389/fonc.2021.650891

Autophagy-Related Three-Gene Prognostic Signature for Predicting Survival in Esophageal Squamous Cell Carcinoma

Heyang Cui[1], Yongjia Weng[1], Ning Ding[1], Chen Cheng[1], Longlong Wang[1], Yong Zhou[1], Ling Zhang[1], Yongping Cui[1] and Weimin Zhang[1,2]**

[1] Department of Oncology, Cancer Institute, Peking University Shenzhen Hospital, Shenzhen Peking University-Hong Kong University of Science and Technology (PKU-HKUST) Medical Center, Shenzhen, China, [2] Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Laboratory of Molecular Oncology, Peking University Cancer Hospital & Institute, Beijing, China

OPEN ACCESS

Edited by:
Die Wang,
Hudson Institute of Medical Research, Australia

Reviewed by:
Yu Dou,
Shandong University, China
Ji-Feng Feng,
University of Chinese Academy of Sciences, China

***Correspondence:**
Yongping Cui
cuiyp@sphmc.org
Weimin Zhang
zhangweimin@bjmu.edu.cn

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: 08 January 2021
Accepted: 22 February 2021
Published: 15 July 2021

Citation:
Cui H, Weng Y, Ding N, Cheng C, Wang L, Zhou Y, Zhang L, Cui Y and Zhang W (2021) Autophagy-Related Three-Gene Prognostic Signature for Predicting Survival in Esophageal Squamous Cell Carcinoma. Front. Oncol. 11:650891. doi: 10.3389/fonc.2021.650891

Esophageal squamous cell carcinoma (ESCC) is one of the most aggressive malignant tumors in China, and its prognosis remains poor. Autophagy is an evolutionarily conserved catabolic process involved in the occurrence and development of ESCC. In this study, we described the expression profile of autophagy-related genes (ARGs) in ESCC and developed a prognostic prediction model for ESCC patients based on the expression pattern of ARGs. We used four ESCC cohorts, GSE53624 (119 samples) set as the discovery cohort, The Cancer Genome Atlas (TCGA) ESCC set (95 samples) as the validation cohort, 155 ESCC cohort, and Oncomine cohort were used to screen and verify differentially expressed ARGs. We identified 34 differentially expressed genes out of 222 ARGs. In the discovery cohort, we divided ESCC patients into three groups that showed significant differences in prognosis. Then, we analyzed the prognosis of 34 differentially expressed ARGs. Three genes [poly (ADP-ribose) polymerase 1 (PARP1), integrin alpha-6 (ITGA6), and Fas-associated death domain (FADD)] were ultimately obtained through random forest feature selection and were constructed as an ARG-related prognostic model. This model was further validated in TCGA ESCC set. Cox regression analysis confirmed that the three-gene signature was an independent prognostic factor for ESCC patients. This signature effectively stratified patients in both discovery and validation cohorts by overall survival ($P = 5.162E\text{-}8$ and $P = 0.052$, respectively). We also constructed a clinical nomogram with a concordance index of 0.713 to predict the survival possibility of ESCC patients by integrating clinical characteristics and the ARG signature. The calibration curves substantiated fine concordance between nomogram prediction and actual observation. In conclusion, we constructed a new ARG-related prognostic model, which shows the potential to improve the ability of individualized prognosis prediction in ESCC.

Keywords: ESCC, autophagy, PARP1, ITGA6, FADD, prognosis

INTRODUCTION

Esophageal cancer is one of the most common malignant tumors of the digestive system, with high morbidity and mortality (1). It has two major histological types: esophageal adenocarcinoma (EAC) and esophageal squamous cell carcinoma (ESCC) (2). ESCC is the principal histological type in China, which has the highest incidence and mortality compared with other countries (3). Despite the technical developments in diagnosis and treatment, this disease still tends to have a poor prognosis (2, 4) due to late diagnosis and lack of effective targets. Better understanding of the genetic and molecular disorders of the disease is the key to early diagnosis, appropriate treatment, and improved prognosis of patients with ESCC.

Autophagy is a critical and intricate homeostatic process in cells that is involved in a variety of biological processes (5). When exposed to various external stimuli, such as starvation, hypoxia, and drug, the magnitude of autophagy may increase sharply to provide nutrients and remove harmful substances (6). It suggests that autophagy is subjected to highly orchestrated regulation, including phosphoinositide 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR), p53/damage-regulated autophagy modulator (DRAM), Janus kinase (JAK)–signal transducer and activator of transcription (STAT), RAS, and AMP-activated protein kinase (AMPK)/calcium/calmodulin-dependent protein kinase kinase (CaMKK) signaling pathways; some known signaling pathways regulating critical cell cycle are all related to autophagy (7).

Autophagy is generally regarded as a double-edged sword in tumors (8). It may have the opposite effect depending on the tumor type, clinical stage, genetic background, or treatment, which either suppresses or promotes tumor development (9). In general, autophagy can prevent carcinogenesis by removing carcinogenic protein substrates, misfolded proteins, and damaged organelles (8). However, in established cancer, autophagy can meet the needs of tumor growth by recycling macromolecules and organelles (10). At present, autophagy has been gradually used in the diagnosis and treatment of tumors in some studies. Its inhibitor chloroquine and hydroxychloroquine have been used in clinical treatment (7). These drugs alone or in combination have been used in clinical trials of some tumors, including melanoma, colorectal cancer, myeloma, and renal cell carcinoma. The results show that autophagy inhibitors have certain therapeutic potential (11–13). However, although autophagy has been found to be associated with chemotherapy resistance in esophageal squamous cell lines (14, 15), its roles and clinical value have not been tested in patients with ESCC. Thus, it is of great significance to find suitable molecular biomarkers with autophagy as the core for prognosis prediction and treatment of ESCC.

In this study, we aimed to explore autophagy-related genes (ARGs) involved in ESCC progression. Gene expression data from public databases Gene Expression Omnibus (GEO) were used to classify subtypes of ESCC and established prognosis risk model based on ARGs. The relationships between the molecular subtypes and prognosis and clinical characteristics of ESCC patients were further evaluated. The three-gene prognostic risk model constructed with the differentially expressed ARGs among ESCC can better evaluate the prognosis of ESCC samples. Furthermore, TCGA gene expression data set was used to further verify the well-performance of the prognostic risk model.

MATERIALS AND METHODS

Selection of Autophagy-Related Genes

The 222 ARGs were collected from Human Autophagy Database (HADb; http://www.autophagy.lu/clustering/) in March 2019. And 222 ARGs were listed in **Supplementary Table 1**.

Data Acquisition and Processing

The expression data of the GSE53624 dataset and clinical characteristics of ESCC cohorts were obtained from the GEO (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53624). The expression data of TCGA ESCC RNA sequencing (RNA-seq) dataset were downloaded from TCGA website (https://portal.gdc.cancer.gov/) for the validation studies. We used our own ESCC cohort containing 155 ESCC RNA-seq data to screen the differentially expressed ARGs. In this cohort, we performed RNA-seq on fresh tumor specimens and matched adjacent normal tissues from 155 ESCC patients recruited from Shanxi province, China. In addition, we also verified the differentially expressed genes (DEGs) in 53 pairs of ESCC from the Oncomine database (https://www.oncomine.org).

RNA Sequencing and Gene Expression Analysis

Total RNA was extracted from frozen samples using the TRIzol reagent (Life Technologies, Carlsbad, CA, USA), and DNA was digested by DNase I following the instructions of the manufacturer. RNA quantity and quality were evaluated by NanoDrop spectrophotometer (Thermo Scientific, USA). Here, 1% gel electrophoresis was used to determine the RNA integrity. Enriched mRNA with Oligo (dT) were broken into fragments for the preparation of cDNA libraries. The cDNA libraries were quality inspection qualified with the Agilent 2100 Bioanalyzer and ABI Step One Plus Real-Time PCR System, then sequenced on Illumina HiSeq X Ten.

Abbreviations: EIF2AK2, Eukaryotic Translation Initiation Factor 2 Alpha Kinase 2; BIRC5, Baculoviral IAP Repeat Containing 5; HSP90AB1, Heat Shock Protein 90 Alpha Family Class B Member 1; BID, BH3 Interacting Domain Death Agonist; GAA, Alpha Glucosidase; TP63, Tumor Protein P63; ITGB4, Integrin Subunit Beta 4; ITGB3, Integrin Subunit Beta 3; ATIC, 5-Aminoimidazole-4-Carboxamide Ribonucleotide Formyltransferase/IMP Cyclohydrolase; PELP1, Proline, Glutamate And Leucine Rich Protein 1; DDIT3, DNA Damage Inducible Transcript 3; PIK3R4, Phosphoinositide-3-Kinase Regulatory Subunit 4; SPNS1, Sphingolipid Transporter 1; CAPNS1, Calpain Small Subunit 1; FOXO3, Forkhead Box O3; SESN2, Sestrin 2; PARK2, Parkin RBR E3 Ubiquitin Protein Ligase; GNAI3, G Protein Subunit Alpha I3; SH3GLB1, SH3 Domain Containing GRB2 Like, Endophilin B1; ATG9B, Autophagy Related 9B; ULK3, Unc-51 Like Kinase 3; PINK1, PTEN Induced Kinase 1; ERO1L, Endoplasmic Reticulum Oxidoreductase 1 Alpha; CHMP2B, Charged Multivesicular Body Protein 2B; TP53INP2, Tumor Protein P53 Inducible Nuclear Protein 2; RAB11A, Member RAS Oncogene Family; NRG2, Neuregulin 2; ERBB2, Erb-B2 Receptor Tyrosine Kinase 2; MAPK3, Mitogen-Activated Protein Kinase 3; RAB5A, Member RAS Oncogene Family; HSPB8, Heat Shock Protein Family B (Small) Member 8.

Over 50M raw reads were sequenced for each sample. Raw reads were trimmed by Skewer (v0.2.2) (16) to remove adapter sequences and then aligned against reference genome (GRCh37/hg19) by STAR (v2.4.2a) (17). RSEM (1.2.29) (18) was used to perform expression abundance quantification based on the uniquely mapped reads. Gene annotation GENCODE v19 was used in the above process.

Screening of Differentially Expressed Autophagy-Related Genes

Student's *t*-test, receiver operating characteristic (ROC), and GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r) were used to screen the DEGs between ESCC and normal tissue. Genes with area under the ROC curve (AUC) $\geq$0.85, $q < 0.0001$, $|\log2(FC)| \geq 0.5$ were selected as the significantly differentially expressed ARGs.

First-Round Validation

We used the 155-ESCC dataset to verify the differentially expressed identified ARGs. EdgeR package in R statistical software was applied to estimate differentially expressed ARGs between ESCC and normal samples ($q < 0.0001$ and log2(FC) $\geq$0.5 or $\leq$-0.5).

Second-Round Validation

The first-round validated ARGs were further verified using the Oncomine database (https://www.oncomine.org/resource/main.html). Very strict thresholds were applied, $P \leq 0.0001$, log2(FC) $\geq$0.5 or $\leq$-0.5.

Cluster to Identify Subtypes

We used 34 identified differentially ARGs to cluster analysis. Ward.D2 algorithm was used to cluster the ESCC samples. And then we used pheatmap of R to draw cluster heatmap, annotated by clinical features, including Age, Stage, Lymph node metastasis, Location, Drinking, Smoking, and Gender.

Construction of a Prognostic Gene Signature Based on Autophagy-Related Genes

Univariate Cox regression analyses were performed to select the ARGs whose expression profiles were significantly associated with ESCC patient's overall survival (OS) ($P < 0.1$). And then we further used the random survival forest algorithm to rank the importance of prognostic ARGs. R package random survival forest was used to screen the prognostic genes. We set the number of Monte Carlo iterations to 100 and the number of steps forward to 5 and identified the genes whose relative importance as characteristic genes was >0.3. Finally, we carried out a multivariate Cox regression analysis and constructed a risk scoring model:

$$\text{Risk Score} = \sum_{k-1}^{n} \left(Exp_k \ * \ e_k^{HR} \right)$$

N is the number of prognostic ARGs, Exp_k is the expression value of the ARGs, and e_k^{HR} is the estimated regression coefficient of genes in the multivariate Cox regression analysis.

Functional Enrichment Analysis

We performed a series of gene functional enrichment analyses with DEGs, including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). The Database for Annotation, Visualization, and Integrated Discovery [DAVID (19); https://david.ncifcrf.gov/] was used to identify enriched GO and KEGG terms. And we used GOplot package of R to visualize the results of enrichment analysis.

Development of the Nomogram

Significant factors of univariate analysis (Age, Stage, and Risk score) were used to construct a nomogram by the survival and the rms package for R. And we used the concordance index (C-index) to assess the model performance for predicting prognosis. Following that, calibration curves were plotted to evaluate the concordance between actual and predicted survival.

Statistical Analysis

All statistics were executed using the R software (Version 4.0.2; https://www.R-project.org) and SPSS software (Version 22.0; https://www.ibm.com/analytics/spss-statistics-software). Student's *t*-test was used to compare the expression between tumor and normal samples. Fisher exact test was used to check the association of risk scores with clinical characteristics. Kaplan–Meier (KM) curves were plotted and a log-rank test and univariate Cox proportional hazard regression analysis were used to check the significant difference in OS. Univariate and multivariate Cox proportional hazard regression analysis was also performed to assess the association between risk score or clinical characteristics and OS. The ROC analysis was used to examine the sensitivity and specificity. An AUC served as an indicator of prognostic accuracy. A $P < 0.05$ or 0.1 was set as statistically significant.

RESULTS

Differentially Expressed Autophagy-Related Genes in Esophageal Squamous Cell Carcinoma

We used ESCC dataset GSE53624, which contains 222 ARGs (**Supplementary Table 1**, collected from HADb) from 119 ESCC and paired normal esophageal tissues to determine the differentially expressed ARGs. The overall flowchart of this study is shown in **Figure 1A**. Expression of 42 ARGs was found to more effectively discriminate ESCC from normal esophagus with AUC $\geq$ 0.85, $q < 0.0001$, $|\log2(FC)| \geq$ 0.5 (**Supplementary Table 2**), including 17 upregulated ARGs (**Figure 1B**) and 25 downregulated ARGs (**Figure 1C**).

Analysis and Validation of the Differentially Expressed Autophagy-Related Genes in the 155 Esophageal Squamous Cell Carcinoma Dataset and Oncomine Dataset

We then validated the 42 differentially expressed ARGs in an ESCC RNA-seq dataset, which contains 155 ESCC and paired normal esophagus. Here, 34 overlapping ARGs showed the significant differential expression with $q < 0.0001$ and

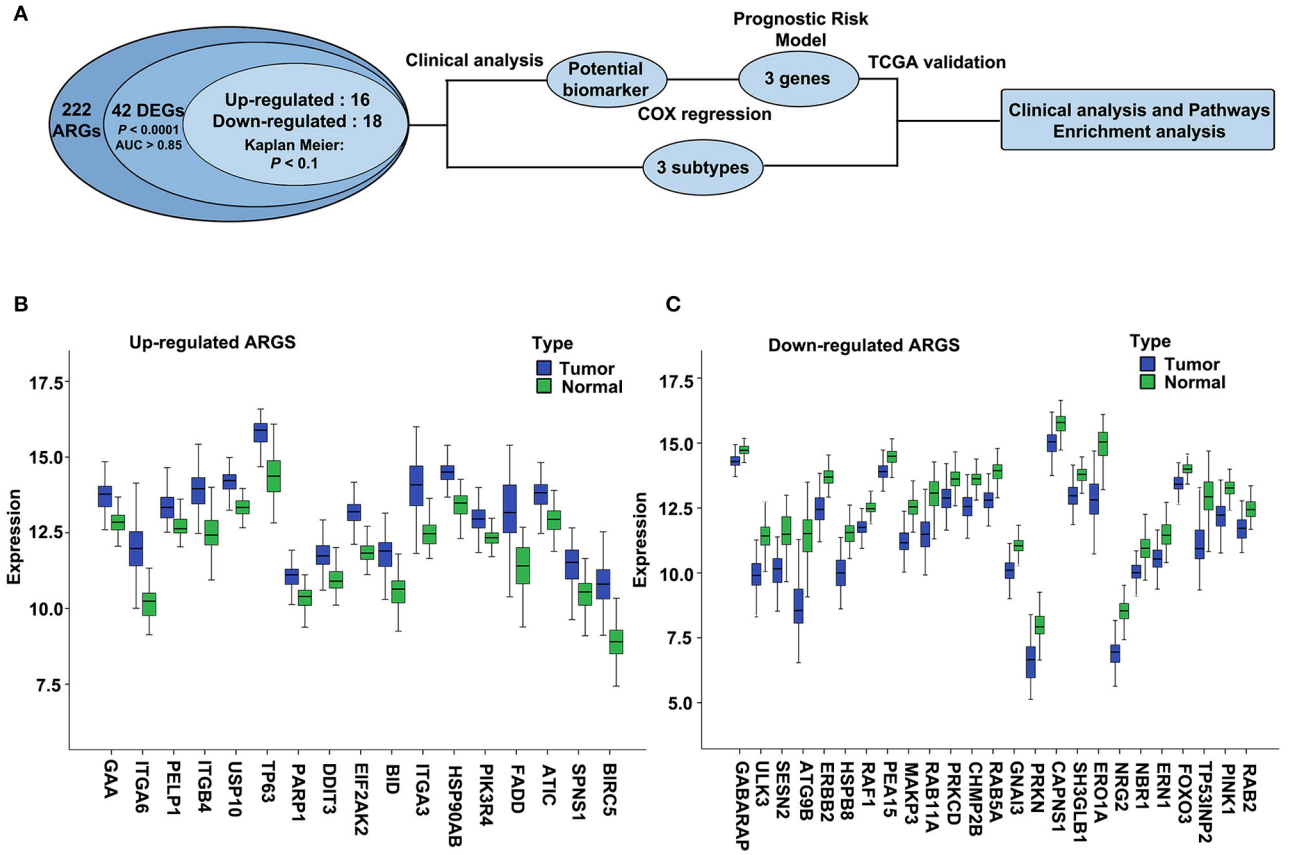


FIGURE 1 | The overall flowchart and significantly differentially expressed autophagy-related genes (ARGs) in esophageal squamous cell carcinoma (ESCC). **(A)** Procedure for the selection and validation of the prognostic risk model in ESCC. DEGs, differentially expressed genes. **(B,C)** The expression patterns of 42 ARGs in ESCC and paired normal samples. Blue box represents tumor, and green box represents normal. **(B)** Upregulated ARGs. **(C)** Downregulated ARGs.

log2(FC) $\geq$0.5 or $\leq$-0.5 (**Supplementary Table 3**), and the trend of ARG differential expression in the two groups was consistent, including 16 upregulated genes [EIF2AK2, BIRC5, HSP90AB1, BID, integrin alpha-6 (ITGA6), GAA, TP63, ITGB4, poly (ADP-ribose) polymerase 1 (PARP1), ITGA3, ATIC, Fas-associated death domain (FADD), PELP1, DDIT3, PIK3R4, SPNS1] and 18 downregulated genes (CAPNS1, FOXO3, SESN2, PARK2, GNAI3, SH3GLB1, ATG9B, ULK3, PINK1, ERO1L, CHMP2B, TP53INP2, RAB11A, NRG2, ERBB2, MAPK3, RAB5A, and HSPB8). We further verified the differential expression trend of these genes in ESCCs of the Oncomine database (**Supplementary Table 3**).

Functional Annotation of the 34 Differentially Expressed Autophagy-Related Genes

Functional enrichment analysis of the 34 differentially expressed ARGs offered the biological understanding of these genes. According to the results of DAVID, the top enriched GO terms for cellular components were cytosol, membrane, mitochondrion, protein complex, cytoplasmic vesicle, integrin complex, late endosome, extracellular exosome, cell–cell adherens junction, cytoplasm, and autophagosome. For the molecular function, genes were mostly enriched in terms of protein binding, identical protein binding, and cadherin binding involved in cell–cell adhesion (**Supplementary Figure 1**). KEGG pathways enrichment analysis for the 34 differentially expressed ARGs showed that these genes were notably associated with pathways in cancer, focal adhesion, and PI3K–AKT signaling pathway (**Supplementary Figure 2A**). The heatmap of the relationship between ARGs and pathways was also displayed (**Supplementary Figure 2B**), including the focal adhesion and PI3K–AKT signaling pathway, which is consistent with previous studies (20–27).

Molecular Typing Based on Autophagy-Related Genes

Molecular subtypes were identified using the cluster method Ward.D2 based on 34 selected differentially expressed ARGs, and the optimal clustering number of 3 was selected (**Figure 2A**). We analyzed the prognosis of these three groups. The results showed that Cluster1 had a relatively better survival followed by Cluster3, whereas Cluster2 had the worst prognosis (**Figure 2B**; $P = 0.02$). The relationships between the subtypes and clinicopathological parameters (Age, Gender, Smoking, Drinking, Location, Grade, Stage, and Lymph node metastasis) of ESCC patients were summarized in **Supplementary Table 4**. We observed significant correlations between subtypes and Drinking ($P = 0.01$), Location

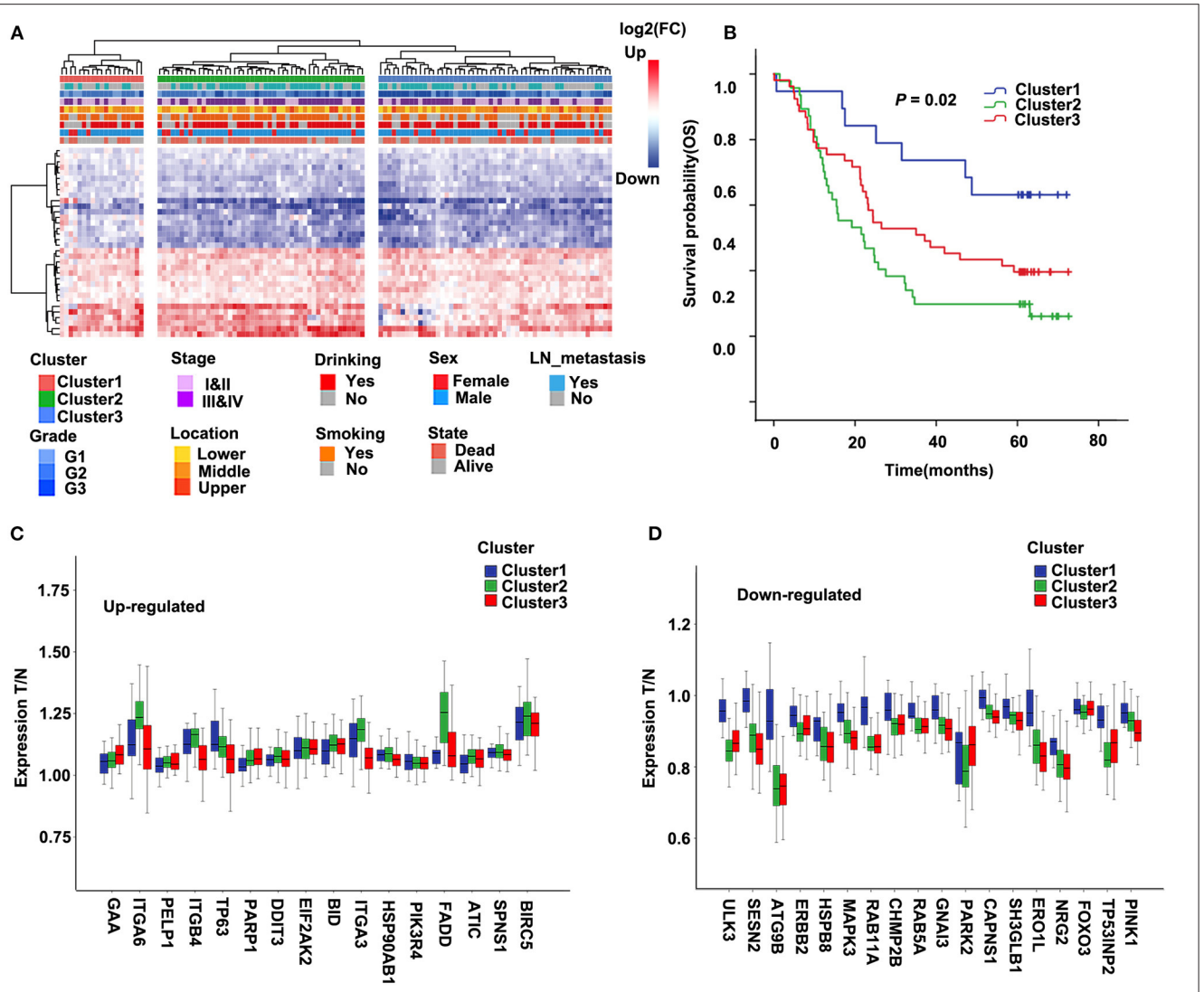


FIGURE 2 | Cluster of differentially expressed autophagy-related genes (ARGs) and Kaplan–Meier (KM) survival plot in the discovery cohort (GSE53624). **(A)** The heatmap of 34 ARGs. Left: Annotation of cluster and clinical features. The color represents logFC of differential expression. LN, lymph node. **(B)** Survival curve and Kaplan–Meier analysis of esophageal squamous cell carcinoma (ESCC) patients by the cluster. OS, overall survival. **(C,D)** The ratio of tumor to normal expression of up- and down-ARGs in Cluster1–3; the ordinate shows the ratio of tumor to normal after normalization of expression.

($P = 0.092$), Grade ($P = 0.009$), Stage ($P = 0.01$), and Lymph node metastasis ($P = 0.076$). Compared to Cluster1, patients in Cluster2 and Cluster3 had higher grade and stage. In Cluster3, the proportion of patients with alcohol drinking was relatively higher. The tumor location of Cluster1 and Cluster3 was more in the middle, while Cluster2 was more in the middle and lower sections of the esophagus.

Further, we compared the differentially expressed ARGs among these three groups. For most genes of 34 ARGs, the degree of upregulated ARGs in Cluster2 was significantly higher than the other two groups, and the degree of downregulated ARGs in Cluster2 and Cluster3 was significantly higher than Cluster1 (**Figures 2C,D**). This result may suggest that the changes of autophagy activities are related to the prognosis of ESCC patients.

Construction of a Prognostic Risk Model Based on These 34 Autophagy-Related Genes

To identify a prognostic risk model, we analyzed the relationship between the expression of 34 ARGs and the prognosis of ESCC patients in the discovery cohort GSE53624 and selected 17 ARGs with significant *P*-value of univariate Cox regression (**Figure 3A**) as candidate genes. We used random forests for feature selection. The relationship between error rate and number of taxonomic trees was used to reveal genes with relative importance >0.3 as the final model (**Figures 3B,C**). We identified three genes, FADD, PARP1, and ITGA6 in this model (**Table 1**, **Supplementary Figure 3**). The important order of the out-of-bag scores for the three genes is displayed in **Figure 3C**. A

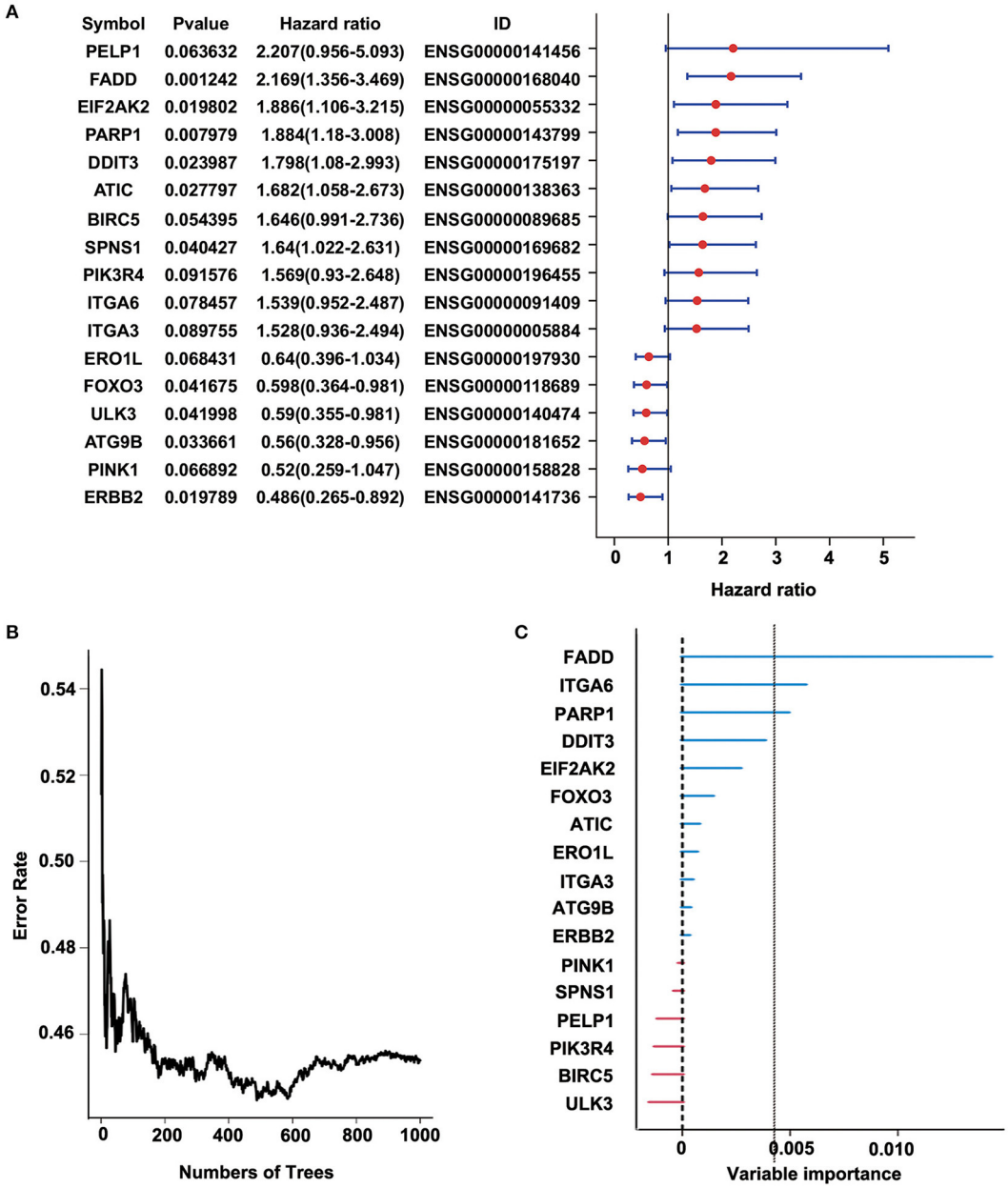


FIGURE 3 | Random forest analysis of prognosis-related autophagy-related genes (ARGs) in the discovery cohort (GSE53624). **(A)** Forest plot of ARGs with esophageal squamous cell carcinoma (ESCC) survival, univariate Cox regression. **(B)** Relationship between the error rate and the number of classification trees. **(C)** Out-of-bag importance values for the predictors.

TABLE 1 | Three genes significantly associated with overall survival in the discovery cohort (GSE53624).

	Symbol	HR	*P*	Importance	Relative imp
ENSG00000168040	FADD	2.169	0.001242	0.0143	1
ENSG00000091409	ITGA6	1.539	0.078457	0.0057	0.3987
ENSG00000143799	PARP1	1.884	0.007979	0.0049	0.3438

FADD, Fas-associated death domain; HR, hazard ratio; Imp, importance; ITGA6, integrin alpha-6; PARP1, poly (ADP-ribose) polymerase 1.

three-gene prognostic risk model was established by multivariate COX regression analysis. The equation is as follows:

$$\text{Risk Score} = 0.184 * \text{expFADD} + 0.562 * \text{expPARP1} + 0.199 * \text{expITGA6}$$

The risk score of each sample was calculated, ROC curve was constructed according to the value of risk and survival of patients, and the samples were divided into high-risk group and low-risk group with the maximum of Youden index. The prognosis

of the high-risk and low-risk groups were significantly different ($P = 5.162E-8$; **Figure 4B**). High expression of FADD, PARP1, and ITGA6 was associated with high risk (**Figures 4A,C**).

Furthermore, we analyzed the prognosis of risk groups and their relationship with cluster, which was constructed by differentially expressed ARGs (**Supplementary Table 5**). The results show that high-risk patients were enriched in Cluster2 and Cluster3, and the high-risk patients accounted for the highest proportion in Cluster2, which had the worst prognosis ($P = 6.511E-07$; **Supplementary Table 5**, **Supplementary Figure 4**). There was no correlation between other clinical factors and risk groups (**Supplementary Table 5**).

Verification of the Robustness of the Three-Gene Signature Model in The Cancer Genome Atlas Esophageal Squamous Cell Carcinoma Dataset

To verify the robustness of the three-gene signature model, we calculated a risk score for each sample in another validation cohort TCGA ESCC dataset. We used the same method to divide the 95 samples into high-risk and low-risk groups in discovery cohort GSE53624. The prognosis of the low-risk group was significantly better than that of the high-risk group ($P = 0.052$; **Figure 5B**, **Supplementary Table 6**). As shown in **Figure 5**, TCGA data revealed that the relationship between the expression of the three genes and risk score is also consistent with the GSE53624. Thus, the three-gene signature model we constructed was effective to predict prognosis for ESCC patients.

Risk Model and Clinical Characteristic Analysis

To assess the independence of the three-gene signature model in clinical application, we used univariate and multivariate Cox regression to analyze hazard ratio (HR), 95% confidence interval (CI), and *P*-values. We systematically analyzed the clinical information from the patients as recorded in ESCC, including their Age, Gender, Location, Smoking, Drinking, Grade, Stage, as well as our three-gene signature (**Figure 6**, **Supplementary Table 7**). In ESCC, univariate Cox regression analysis revealed that the risk score group (HR = 3.617, 95% CI = 2.212–5.914, $P = 3.008E-07$), Age ($P = 0.024$), Stage (HR = 2.19, 95% CI = 1.339–3.582, $P = 0.002$), and Lymph node metastasis (HR = 2.159, 95% CI = 1.319–3.534, $P = 0.002$) had clinical independence. And the corresponding multivariate Cox regression analysis found that the Risk score group (HR = 2.955, 95% CI = 1.761–4.961, $P = 4.100E-05$), Age ($P = 0.03$), and Stage (HR = 1.849, 95% CI = 1.096–3.12, $P = 0.021$) had clinical independence. Importantly, the validation data (TCGA cohort) also confirmed these findings (HR = 1.971, 95% CI = 0.982–3.955, $P = 0.056$ for univariate Cox regression analysis; HR = 2.67, 95% CI = 1.25–5.704, $P = 0.011$ for multivariate Cox regression analysis; **Supplementary Table 8**, **Supplementary Figure 5A**), suggesting that our three-gene signature model may serve as an independent prognostic index for clinical application.

Then, we constructed a nomogram model, as shown in **Figure 7A**. The univariate analysis was performed among nine variables to verify the prognostic variables with the data from the discovery cohort. Of the nine variables, a total of four variables were prognostic predictors for OS (including Age, Stage, Lymph node metastasis, and Risk score; $P < 0.05$). Three significant factors including Age, Stage (it is associated with Lymph node metastasis), and Risk score in the univariable analysis were enrolled into the multivariable analysis based on the Cox regression. A nomogram that incorporated the mentioned three prognostic factors was established. The prediction accuracy of the nomogram was assessed by C-index, and the results showed that the C-index was 0.713. To read the nomogram, draw a vertical line up to the top row of points to specify points for each variable. Then, the total points for a patient can be added up, and one can obtain the probability of 1-, 3-, and 5-year OS by drawing a vertical line from the total points row. **Figure 7B** showed the 1-, 3-, and 5-year nomogram model and the ideal model, and the results showed that the nomogram model was basically consistent with those of the ideal model. The nomogram was validated in the validation cohort, and 1- and 3-year calibration curves were presented in **Supplementary Figure 5B**. These results indicated that the accuracy of our model is relatively high.

Analysis of Pathway Differences Enriched in the High-Risk and Low-Risk Groups

Pathway enrichment analysis of the DEGs in the high-risk group and the low-risk group showed that Metabolic pathways, Pathways in cancer, Protein digestion and absorption, Human papillomavirus infection, ECM–receptor interaction, and Cell cycle were enriched in both high-risk and low-risk groups. And results revealed that the high-risk may be related to the activity of PI3K–AKT signaling pathway and calcium signaling pathway, etc. (**Figure 8**). Besides, the activity of DNA replication and Fatty acid degradation may be related to low-risk (**Figure 8**). The suppression of the PI3K/AKT/mTOR signaling pathway can induce autophagy, which in turn saves tumor cells from the harm of epidermal growth factor receptor (EGFR)-tyrosine kinase inhibitors (TKIs) (20). Therefore, the PI3K–AKT pathway inhibitors may have potential targeting effect on patients in the high-risk group.

DISCUSSION

Cancers are highly heterogeneous diseases in that survival times vary substantially among patients with similar TNM stages. With the diagnosis and treatment at earlier stage, traditional clinicopathological indicators such as Tumor size, TNM stage, and Vascular invasion have proven inadequate for predicting individual prognosis (28). Since autophagy may play an important role in the development, progression, and therapeutic response of ESCC individually, the screening of prognostic molecular markers based on ARGs may reflect the biological characteristics of ESCC, which is of great significance

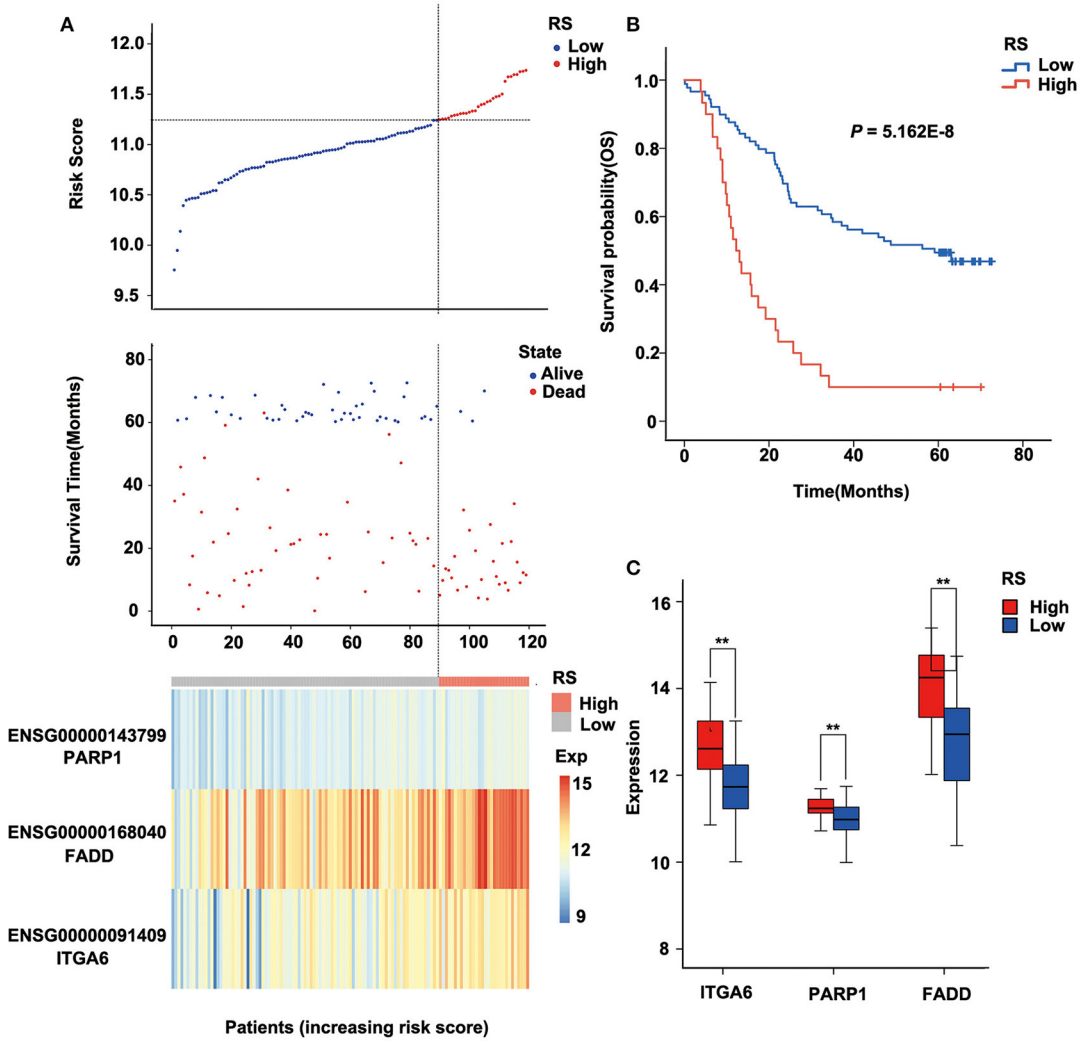


FIGURE 4 | Relation between the three-gene signature and cancer risk in the discovery cohort (GSE53624). **(A)** Risk score, survival time, survival state, and expression of the three ARGs in esophageal squamous cell carcinoma (ESCC) patients of GSE53624. **(B)** Kaplan–Meier analysis of ESCC patients grouped by risk score; log-rank Mantel–Cox test was used to compare survival curves. **(C)** Box plot of expression of three ARGs grouped by risk score, and the independent-sample Student t-test was used as the comparison method between the high- and low-risk groups. RS, risk score; Exp, expression. $^{**}P < 0.01$.

for individualized prevention and treatment. To capture the genes necessary for ESCC from the perspective of autophagy, we screened ARGs and identified key prognostic ARGs, all of which may provide additional potential therapeutic targets. We further used the complementary value of molecular and clinical features and showed that combined analysis can provide a more accurate estimation of OS in ESCC. This comprehensive study of two factors contributes to our new understanding of ESCC biology and depicts potential therapeutic interventions.

In recent research, polygenic prognosis prediction models have been highlighted in clinical practice. For example, Oncotype DX, which provides a breast cancer recurrence score based on 21 genes (29–31), and Coloprin, which provides a colon cancer recurrence score based on 18 genes (32–34). These studies have shown that polygenic prognosis prediction models based on gene expression profiles are efficacious and promising to diagnosis, appropriate treatment, and improved prognosis of patients with cancer. Furthermore, Tian et al. (35) identified a six-gene signature, Zhao et al. (36) identified a three-gene signature, and Wang et al. (28) identified a six-gene signature. These signatures are proofs that the model composed of a small number of genes still has a high prediction efficiency of prognosis. In addition, there are some studies that identified signatures based on differentially expressed ARGs, 22-gene signature in non-small-cell lung cancer (NSCLC; 8) and three-gene signature in bladder cancer [BC; (37)]. These showed us that the model screened based on specific functions also has good efficiency and has good clinical application prospect. Based on these conditions, we screened three ARGs and constructed a polygenic prognosis prediction model and verified its predictive ability.

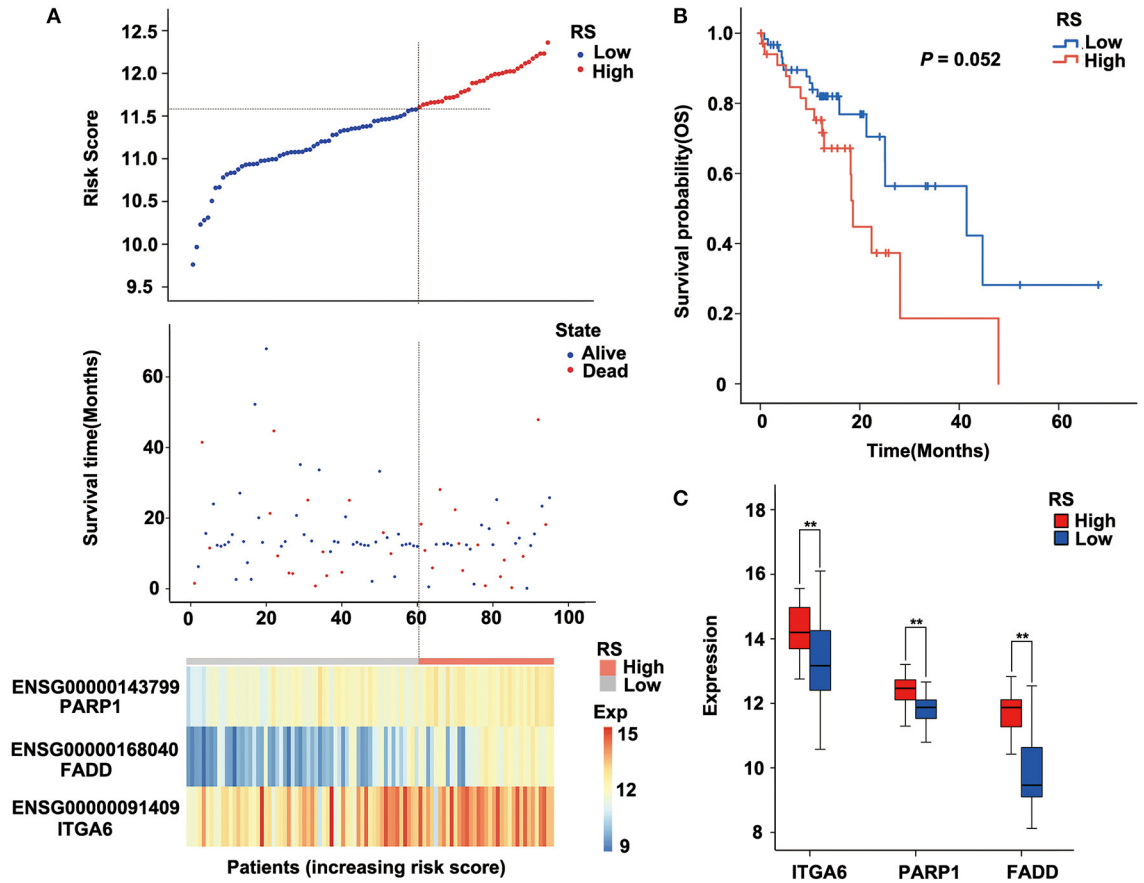


FIGURE 5 | Performance of the three-gene signature model in the validation cohort [The Cancer Genome Atlas (TCGA) data]. **(A)** Risk score, survival time, survival state, and expression of the three ARGs in esophageal squamous cell carcinoma (ESCC) patients of TCGA. **(B)** Kaplan–Meier analysis of ESCC patients grouped by risk score; log-rank Mantel–Cox test was used to compare survival curves. **(C)** Box plot of expression of three ARGs grouped by risk score, and the independent-sample Student *t*-test was used as the comparison method between the high- and low-risk groups. RS, risk score; Exp, expression. $^{**}P < 0.01$.

The three genes in our signature include PARP1, ITGA6, FADD as risk factors. PARP1 is a 113-kDa nuclear polymerase that modifies substrates (38). At present, it has been shown that PARP1 plays a role in the repair of DNA damage (39–41). PARP1-mediated autophagy is a key pathway for TKI resistance in NSCLC cells that participates in the resistance to TKIs (42). PARP1 may be an independent prognostic marker in ESCC, and PARP1 inhibition can induce cell cycle arrest at the G2/M phase through the ATM–Chk2–CDC25C pathway (38). ITGA6 is a member of the integrins family. Many integrins contribute to tumor progression, and ITGA6 has been implicated in breast cancer progression (43–45). In ESCC, it has been reported that expression of ITGA6 is highly upregulated and plays an important role in the proliferation and invasion (46). FADD is an adaptor molecule that interacts with various cell surface receptors and mediates cell apoptotic signals (47). In recent studies, FADD has been used as a potential autophagy-related prognostic marker in lung squamous cell carcinoma and head and neck squamous cell carcinoma (48, 49). The copy number amplification and upregulation of FADD were also found in ESCC, and its expression was significantly correlated with the survival of ESCC (50). Based on the PARP1–ITGA6–FADD three-gene model, ESCC patients were divided into high-risk group and low-risk group. Compared with single-gene prognosis analysis, this grouping method has more significant difference in prognosis. In addition, through KEGG enrichment analysis, we found that PI3K–AKT pathway was significantly enriched in the high-risk group. It has been proven that PI3K/AKT/mTOR-mediated autophagy played pivotal roles in the occurrence, development, and drug resistance of tumors. Autophagy mediated by PI3K–AKT–mTOR pathway can improve the drug sensitivity of tumor cells and avoid drug resistance (51). Therefore, ESCC patients in this high-risk group may benefit more from the targeted drugs. Through targeting PI3K–AKT–mTOR-mediated autophagy, many drugs can more accurately and specifically regulate autophagy activity of tumor cells, so as to achieve better antitumor therapeutic efficacy.

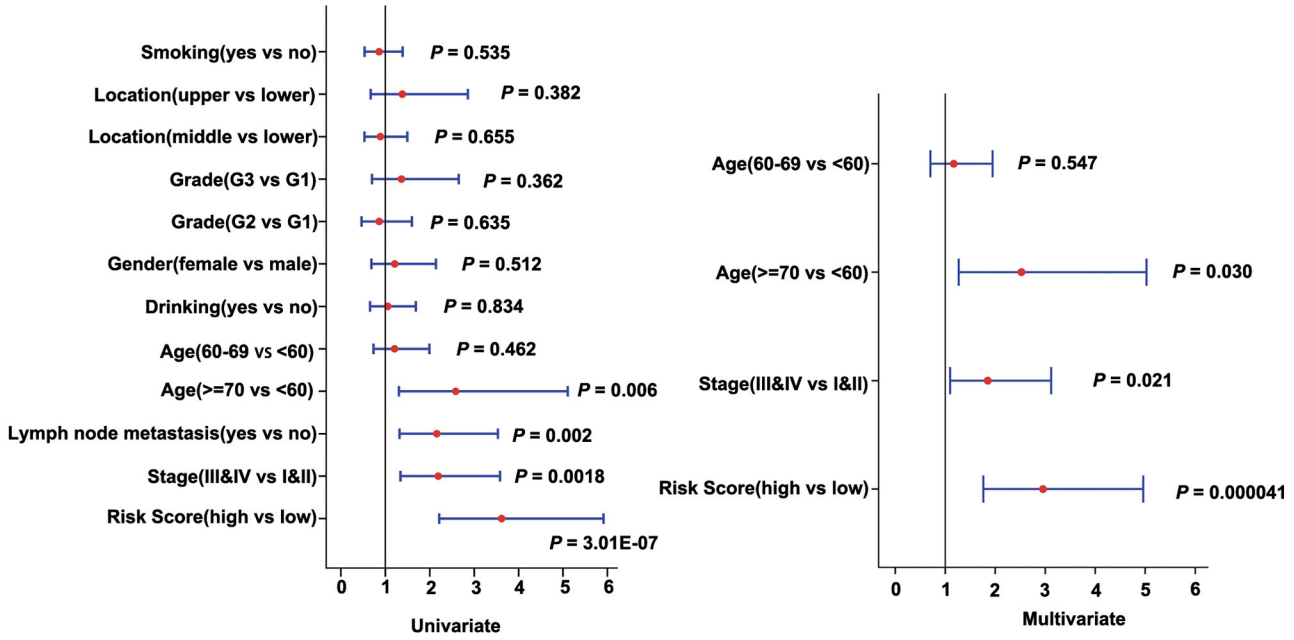


FIGURE 6 | Forest plot of univariate and multivariate Cox regression analyses in the discovery cohort (GSE53624).

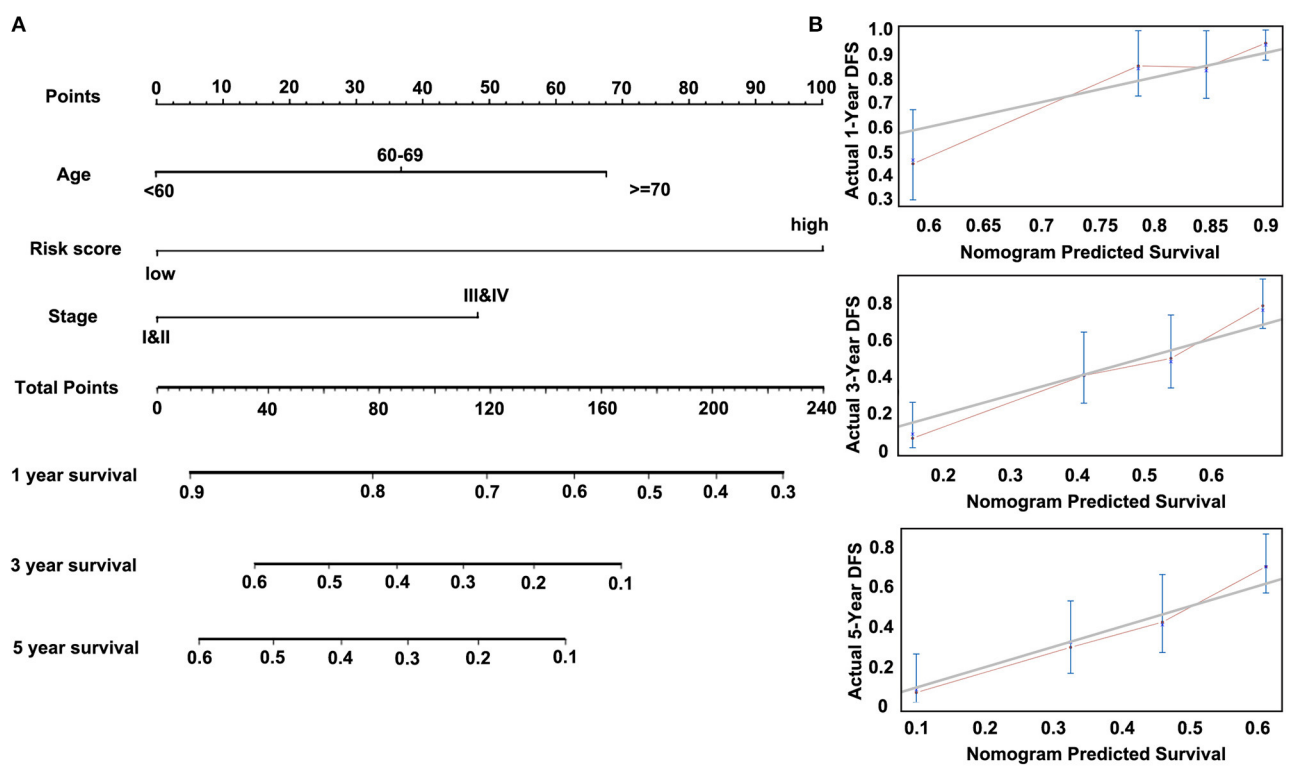


FIGURE 7 | Nomogram of Cox regression model in the discovery cohort (GSE53624). **(A)** The nomogram for predicting overall survival (OS). **(B)** The calibration plots for predicting 1-, 3-, and 5-year OS.

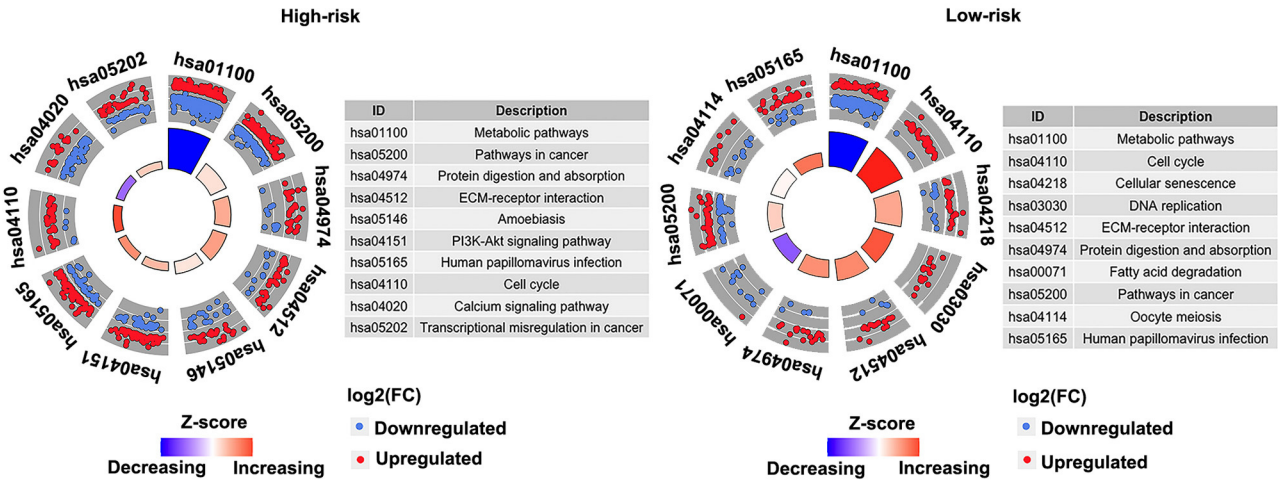


FIGURE 8 | Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in the high- and low-risk groups in the discovery cohort (GSE53624). The outer circle shows a scatter plot for each term of the log2(FC) of the assigned genes. Red circles display upregulation, and the blue ones display downregulation. Left, high-risk. Right, low-risk.

Lastly, we developed a nomogram to predict individuals' clinical outcomes. A nomogram is a stable and reliable tool to quantitatively measure risk on an individual basis by combining delineated risk factors, which has been used for autophagy prognoses (8). A nomogram generates a statistical predictive model presented in a graph, conferring points to each factor such as Stage, Grade, Age, and Gender in the clinical setting. By integrating all the factors, the model provides a predictive assessment for individuals. Apart from traditional clinicopathological features, the risk score based on genes can also be included in the predictive nomogram to better predict clinical results (52–54). Mo et al. (53) built a nomogram to predict survival in colorectal cancer with the inclusion of a prognostic score calculated from autophagy genes. Liu et al. (8) built a nomogram in non-small-cell lung cancer including a 22-autophagy gene signature that can well predict 3- and 5-year survival possibilities. In many cases, the combination of autophagy genes and prognostic factors has better prognosis than using a single factor. Moreover, we also used a calibration curve, the nomogram adopting both the gene signature and conventional prognostic factors that can accurately predict 3- and 5-year survival probabilities.

Although we have identified potential candidate genes and constructed a prognostic model using bioinformatics technology with ESCC samples, our study has several limitations. First, due to the lack of large public ESCC transcriptomic data, the sample size included in this study was not enough, which may affect the efficacy of our prognostic model. Second, although we have verified our findings in different cohorts, it would be better to confirm these results *via* independent experiments, such as immunohistochemistry in another cohort. Therefore, further genetic and experimental studies with larger samples and experimental validation are needed.

In conclusion, we divided ESCC patients into three clusters based on ARGs, and these clusters were related to stage and prognosis. Furthermore, we identified a prognostic three-autophagy gene signature base on GEO and TCGA ESCC cohorts. This three-gene model was an independent predictor of prognosis. And we used gene signature and clinicopathological features to build a nomogram that can accurately predict a 1-, 3-, and 5-year survival probability for individual ESCC patients. This finding suggests that the three-ARG signature may help facilitate personalized treatment.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

HC, YZ, LW, and YW collated and analyzed the data. ND, CC, and LZ completed the writing and repair of the manuscript. WZ and YC designed and guided the subject. All authors reviewed and approved the final manuscript.

FUNDING

This work was supported by the fund of Guangdong Basic and Applied Basic Research Foundation (2019B030302012), China Postdoctoral Science Foundation (2019M663023 and 2020M682840), the National Natural Science Foundation of China (81972613, 81802780), San-ming Project of Medicine in Shenzhen (SZSM201812088), Natural Science Foundation of Guangdong Province (2020A15150143), Major Program of Shenzhen Bay Laboratory (S201101004), Shenzhen Science and Technology Innovation Commission Project (ZDSYS20190902092855097), Shenzhen Key Project of Science and Technology (JCYJ20200109120425045), and Shenzhen Project of Science and Technology (JCYJ20190813094203600).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.650891/full#supplementary-material

REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* (2015) 65:87–108. doi: 10.3322/caac.21262
2. Pennathur A, Gibson MK, Jobe BA, Luketich JD. Oesophageal carcinoma. *Lancet.* (2013) 381:400–12. doi: 10.1016/S0140-6736(12)60643-6
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. *CA Cancer J Clin.* (2016) 66:115–32. doi: 10.3322/caac.21338
4. Hsu PK, Chen HS, Wang BY, Wu SC, Liu CY, Shih CH, et al. Hospital type- and volume-outcome relationships in esophageal cancer patients receiving non-surgical treatments. *World J Gastroenterol.* (2015) 21:1234–42. doi: 10.3748/wjg.v21.i4.1234
5. Mathew R, Karantza-Wadsworth V, White E. Role of autophagy in cancer. *Nat Rev Cancer.* (2007) 7:961–7. doi: 10.1038/nrc2254
6. Marinkovic M, Sprung M, Buljubasic M, Novak I. Autophagy modulation in cancer: current knowledge on action and therapy. *Oxid Med Cell Longev.* (2018) 2018:8023821. doi: 10.1155/2018/8023821
7. Levy JMM, Towers CG, Thorburn A. Targeting autophagy in cancer. *Nat Rev Cancer.* (2017) 17:528–42. doi: 10.1038/nrc.2017.53
8. Liu Y, Wu L, Ao H, Zhao M, Leng X, Liu M, et al. Prognostic implications of autophagy-associated gene signatures in non-small cell lung cancer. *Aging.* (2019) 11:11440–62. doi: 10.18632/aging.102544
9. White E. The role for autophagy in cancer. *J Clin Invest.* (2015) 125:42–6. doi: 10.1172/JCI73941
10. Guo JY, White E. Autophagy, metabolism, and cancer. *Cold Spring Harb Symp Quant Biol.* (2016) 81:73–8. doi: 10.1101/sqb.2016.81.030981
11. Mahalingam D, Mita M, Sarantopoulos J, Wood L, Amaravadi RK, Davis LE, et al. Combined autophagy and HDAC inhibition: a phase I safety, tolerability, pharmacokinetic, and pharmacodynamic analysis of hydroxychloroquine in combination with the HDAC inhibitor vorinostat in patients with advanced solid tumors. *Autophagy.* (2014) 10:1403–14. doi: 10.4161/auto.29231
12. Rangwala R, Leone R, Chang YC, Fecher LA, Schuchter LM, Kramer A, et al. Phase I trial of hydroxychloroquine with dose-intense temozolomide in patients with advanced solid tumors and melanoma. *Autophagy.* (2014) 10:1369–79. doi: 10.4161/auto.29118
13. Vogl DT, Stadtmauer EA, Tan KS, Heitjan DF, Davis LE, Pontiggia L, et al. Combined autophagy and proteasome inhibition: a phase 1 trial of hydroxychloroquine and bortezomib in patients with relapsed/refractory myeloma. *Autophagy.* (2014) 10:1380–90. doi: 10.4161/auto.29264
14. Kijima T, Nakagawa H, Shimonosono M, Chandramouleeswaran PM, Hara T, Sahu V, et al. Three-Dimensional organoids reveal therapy resistance of esophageal and oropharyngeal squamous cell carcinoma cells. *Cell Mol Gastroenterol Hepatol.* (2019) 7:73–91. doi: 10.1016/j.jcmgh.2018.09.003
15. Saxena R, Klochkova A, Murray MG, Kabir MF, Samad S, Beccari T, et al. Roles for autophagy in esophageal carcinogenesis: implications for improving patient outcomes. *Cancers.* (2019) 11:1697. doi: 10.3390/cancers11111697
16. Jiang H, Lei R, Ding SW, Zhu S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics.* (2014) 15:182. doi: 10.1186/1471-2105-15-182
17. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* (2013) 29:15–21. doi: 10.1093/bioinformatics/bts635
18. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* (2011) 12:323. doi: 10.1186/1471-2105-12-323
19. Jiao X, Sherman BT, Huang da W, Stephens R, Baseler MW, Lane HC, et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics.* (2012) 28:1805–6. doi: 10.1093/bioinformatics/bts251
20. Han W, Pan H, Chen Y, Sun J, Wang Y, Li J, et al. EGFR tyrosine kinase inhibitors activate autophagy as a cytoprotective response in human lung cancer cells. *PLoS ONE.* (2011) 6:e18691. doi: 10.1371/journal.pone.0018691
21. Sandilands E, Serrels B, McEwan DG, Morton JP, Macagno JP, McLeod K, et al. Autophagic targeting of Src promotes cancer cell survival following reduced FAK signalling. *Nat Cell Biol.* (2011) 14:51–60. doi: 10.1038/ncb2386
22. Sui X, Kong N, Zhu M, Wang X, Lou F, Han W, et al. Cotargeting EGFR and autophagy signaling: a novel therapeutic strategy for non-small-cell lung cancer. *Mol Clin Oncol.* (2014) 2:8–12. doi: 10.3892/mco.2013.187
23. Kenific CM, Stehbens SJ, Goldsmith J, Leidal AM, Faure N, Ye J, et al. NBR1 enables autophagy-dependent focal adhesion turnover. *J Cell Biol.* (2016) 212:577–90. doi: 10.1083/jcb.201503075
24. Sharifi MN, Mowers EE, Drake LE, Collier C, Chen H, Zamora M, et al. Autophagy promotes focal adhesion disassembly and cell motility of metastatic tumor cells through the direct interaction of paxillin with LC3. *Cell Rep.* (2016) 15:1660–72. doi: 10.1016/j.celrep.2016.04.065
25. Henson E, Chen Y, Gibson S. EGFR family members' regulation of autophagy is at a crossroads of cell survival and death in cancer. *Cancers.* (2017) 9:27. doi: 10.3390/cancers9040027
26. Dower CM, Wills CA, Frisch SM, Wang HG. Mechanisms and context underlying the role of autophagy in cancer metastasis. *Autophagy.* (2018) 14:1110–28. doi: 10.1080/15548627.2018.1450020
27. Wang X, Li W, Zhang N, Zheng X, Jing Z. Opportunities and challenges of co-targeting epidermal growth factor receptor and autophagy signaling in non-small cell lung cancer. *Oncol Lett.* (2019) 18:499–506. doi: 10.3892/ol.2019.10372
28. Wang J, Chen X, Tian Y, Zhu G, Qin Y, Chen X, et al. Six-gene signature for predicting survival in patients with head and neck squamous cell carcinoma. *Aging.* (2020) 12:767–83. doi: 10.18632/aging.102655
29. Bhutiani N, Egger ME, Ajkay N, Scoggins CR, Martin, 2nd RC, McMasters KM. Multigene signature panels and breast cancer therapy: patterns of use and impact on clinical decision making. *J Am Coll Surg.* (2018) 226:406–12.e1. doi: 10.1016/j.jamcollsurg.2017.12.043
30. Siow ZR, De Boer RH, Lindeman GJ, Mann GB. Spotlight on the utility of the oncotype DX((R)) breast cancer assay. *Int J Womens Health.* (2018) 10:89–100. doi: 10.2147/IJWH.S124520
31. Wang SY, Dang W, Richman I, Mougalian SS, Evans SB, Gross CP. Cost-Effectiveness analyses of the 21-gene assay in breast cancer: systematic review and critical appraisal. *J Clin Oncol.* (2018) 36:1619–27. doi: 10.1200/JCO.2017.76.5941
32. Tan IB, Tan P. Genetics: an 18-gene signature (ColoPrint(R)) for colon cancer prognosis. *Nat Rev Clin Oncol.* (2011) 8:131–3. doi: 10.1038/nrclinonc.2010.229
33. Maak M, Simon I, Nitsche U, Roepman P, Snel M, Glas AM, et al. Independent validation of a prognostic genomic signature (ColoPrint)

for patients with stage II colon cancer. *Ann Surg.* (2013) 257:1053–8. doi: 10.1097/SLA.0b013e31827c1180
34. Kopetz S, Tabernero J, Rosenberg R, Jiang ZQ, Moreno V, Bachleitner-Hofmann T, et al. Genomic classifier ColoPrint predicts recurrence in stage II colorectal cancer patients more accurately than clinical factors. *Oncologist.* (2015) 20:127–33. doi: 10.1634/theoncologist.2014-0325
35. Tian S, Meng G, Zhang W. A six-mRNA prognostic model to predict survival in head and neck squamous cell carcinoma. *Cancer Manag Res.* (2019) 11:131–42. doi: 10.2147/CMAR.S185875
36. Zhao X, Sun S, Zeng X, Cui L. Expression profiles analysis identifies a novel three-mRNA signature to predict overall survival in oral squamous cell carcinoma. *Am J Cancer Res.* (2018) 8:450–61.
37. Wang SS, Chen G, Li SH, Pang JS, Cai KT, Yan HB, et al. Identification and validation of an individualized autophagy-clinical prognostic index in bladder cancer patients. *Onco Targets Ther.* (2019) 12:3695–712. doi: 10.2147/OTT.S197676
38. Yamamoto M, Yamasaki M, Tsukao Y, Tanaka K, Miyazaki Y, Makino T, et al. Poly (ADP-ribose) polymerase-1 inhibition decreases proliferation through G2/M arrest in esophageal squamous cell carcinoma. *Oncol Lett.* (2017) 14:1581–7. doi: 10.3892/ol.2017.6334
39. Juarez-Salinas H, Sims JL, Jacobson MK. Poly(ADP-ribose) levels in carcinogen-treated cells. *Nature.* (1979) 282:740–1. doi: 10.1038/282740a0
40. Durkacz BW, Omidiji O, Gray DA, Shall S. (ADP-ribose)n participates in DNA excision repair. *Nature.* (1980) 283:593–6. doi: 10.1038/283593a0
41. Rouleau M, Patel A, Hendzel MJ, Kaufmann SH, Poirier GG. PARP inhibition: PARP1 and beyond. *Nat Rev Cancer.* (2010) 10:293–301. doi: 10.1038/nrc2812
42. Zhang Z, Lian X, Xie W, Quan J, Liao M, Wu Y, et al. Role of PARP1-mediated autophagy in EGFR-TKI resistance in non-small cell lung cancer. *Sci Rep.* (2020) 10:20924. doi: 10.1038/s41598-020-77908-z
43. Friedrichs K, Ruiz P, Franke F, Gille I, Terpe HJ, Imhof BA. High expression level of alpha 6 integrin in human breast carcinoma is correlated with reduced survival. *Cancer Res.* (1995) 55:901–6.
44. Wewer UM, Shaw LM, Albrechtsen R, Mercurio AM. The integrin alpha 6 beta 1 promotes the survival of metastatic human breast carcinoma cells in mice. *Am J Pathol.* (1997) 151:1191–8.
45. Mukhopadhyay R, Theriault RL, Price JE. Increased levels of alpha6 integrins are associated with the metastatic phenotype of human breast cancer cells. *Clin Exp Metastasis.* (1999) 17:325–32. doi: 10.1023/A:1006659230585
46. Ma G, Jing C, Huang F, Li X, Cao X, Liu Z. Integrin alpha6 promotes esophageal cancer metastasis and is targeted by miR-92b. *Oncotarget.* (2017) 8:6681–90. doi: 10.18632/oncotarget.14259
47. Hartmann A, Mouatt-Prigent A, Faucheux BA, Agid Y, Hirsch EC. FADD: A link between TNF family receptors and caspases in Parkinson's disease. *Neurology.* (2002) 58:308–10. doi: 10.1212/WNL.58.2.308
48. Li C, Wu ZH, Yuan K. Autophagy-Related signature for head and neck squamous cell carcinoma. *Dis Markers.* (2020) 2020:8899337. doi: 10.1155/2020/8899337
49. Zhu J, Wang M, Hu D. Development of an autophagy-related gene prognostic signature in lung adenocarcinoma lung squamous cell carcinoma. *PeerJ.* (2020) 8:e8288. doi: 10.7717/peerj.8288
50. Qin HD, Liao XY, Chen YB, Huang SY, Xue WQ, Li FF, et al. Genomic characterization of esophageal squamous cell carcinoma reveals critical genes underlying tumorigenesis and poor prognosis. *Am J Hum Genet.* (2016) 98:709–27. doi: 10.1016/j.ajhg.2016.02.021
51. Xu Z, Han X, Ou D, Liu T, Li Z, Jiang G, et al. Targeting PI3K/AKT/mTOR-mediated autophagy for tumor therapy. *Appl Microbiol Biotechnol.* (2020) 104:575–87. doi: 10.1007/s00253-019-10257-8
52. Xiong Y, Wang R, Peng L, You W, Wei J, Zhang S, et al. An integrated lncRNA, microRNA and mRNA signature to improve prognosis prediction of colorectal cancer. *Oncotarget.* (2017) 8:85463–78. doi: 10.18632/oncotarget.20013
53. Mo S, Dai W, Xiang W, Li Y, Feng Y, Zhang L, et al. Prognostic and predictive value of an autophagy-related signature for early relapse in stages I-III colon cancer. *Carcinogenesis.* (2019) 40:861–70. doi: 10.1093/carcin/bgz031
54. Wang Y, Zhang Q, Gao Z, Xin S, Zhao Y, Zhang K, et al. A novel 4-gene signature for overall survival prediction in lung adenocarcinoma patients with lymph node metastasis. *Cancer Cell Int.* (2019) 19:100. doi: 10.1186/s12935-019-0822-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 04 August 2021
doi: 10.3389/fonc.2021.722417

IL6 and BCL3 Expression Are Potential Biomarkers in Esophageal Squamous Cell Carcinoma

Sheila Coelho Soares-Lima[1†], Isabela Martins Gonzaga[1†], Diego Camuzi[1], Pedro Nicolau-Neto[1], Raissa Vieira da Silva[1], Simone Guaraldi[1,2], Maria Aparecida Ferreira[2], Hector Hernandez-Vargas[3], Zdenko Herceg[4] and Luis Felipe Ribeiro Pinto[1]*

[1] Programa de Carcinogênese Molecular, Instituto Nacional de Câncer (INCA), Rio de Janeiro, Brazil, [2] Seção de Endoscopia, Instituto Nacional de Câncer (INCA), Rio de Janeiro, Brazil, [3] Cancer Research Centre of Lyon (CRCL), Inserm U 1052, CNRS UMR 5286, Centre Léon Bérard, Université de Lyon, Lyon, France, [4] Epigenetics Group, Mechanisms of Carcinogenesis Section, International Agency for Research on Cancer (IARC), Lyon, France

OPEN ACCESS

Edited by:
Die Wang,
Hudson Institute of Medical Research, Australia

Reviewed by:
Shutao Zheng,
Xinjiang Medical University, China
Pengzhou Kong,
Shanxi Medical University, China

***Correspondence:**
Luis Felipe Ribeiro Pinto
lfrpinto@inca.gov.br

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *08 June 2021*
Accepted: *22 July 2021*
Published: *04 August 2021*

Citation:
Soares-Lima SC, Gonzaga IM, Camuzi D, Nicolau-Neto P, Vieira da Silva R, Guaraldi S, Ferreira MA, Hernandez-Vargas H, Herceg Z and Ribeiro Pinto LF (2021) IL6 and BCL3 Expression Are Potential Biomarkers in Esophageal Squamous Cell Carcinoma. Front. Oncol. 11:722417. doi: 10.3389/fonc.2021.722417

Esophageal squamous cell carcinoma (ESCC) ranks among the most lethal tumors worldwide, as a consequence of late detection and poor treatment response, evidencing the need for diagnosis anticipation and new therapeutic targets. First, we investigated the IL6 gene and protein expression in the esophagus of individuals without esophageal disorders (healthy), ESCC, and non-tumoral surrounding tissue (NTST). Our results showed that IL6 mRNA and protein expression is upregulated in tumor cells relative to NTST. In the TCGA dataset, we identified a set of genes whose expression was correlated with *IL6* mRNA levels, including the antiapoptotic gene *BCL3*. By using an immortalized esophageal cell line, we confirmed that IL6 was capable of inducing *BCL3* expression in esophageal cells. BCL3 mRNA and protein are overexpressed in ESCC and NTST compared to healthy esophagus, and *BCL3* mRNA could distinguish the morphologically normal samples (healthy and NTST) with 100% sensitivity and 95.12% specificity. The spatial intratumoral heterogeneity of both *IL6* and *BCL3* expression was evaluated, corroborating *IL6* upregulation throughout the tumor, while tumor and NTST showed a consistent increase of *BCL3* expression relative to the healthy esophagus. Our study shows that IL6 overexpression seems to be a key event in ESCC carcinogenesis, contributing to ESCC through a homogeneous antiapoptotic signalling *via* BCL3 overexpression, thus suggesting anti-IL6 therapies to be further considered for ESCC treatment. Finally, our data support the use of *BCL3* mRNA expression as a potential biomarker for ESCC detection.

Keywords: esophageal squamous cell carcinoma (ESCC), BCL3, IL6, biomarker, therapeutic target, diagnosis

INTRODUCTION

Esophageal squamous cell carcinoma (ESCC) accounts for the majority of esophageal tumors around the world, with a significant impact on public health since this cancer is the seventh most common type of cancer worldwide, affecting mainly low and middle-income countries. Furthermore, ESCC presents a poor prognosis, with an overall survival between 10 and 20%,

mainly due to late-stage diagnosis (1, 2). Therefore, the development of biomarkers that can help ESCC early diagnosis is extremely important. Chronic irritation and inflammation of the esophageal epithelium have been correlated with an increased risk of ESCC development. Complementary, genome-wide DNA methylation profiling of esophageal squamous cell carcinoma carried out by our group showed that genes involved in the inflammatory response are commonly altered by DNA methylation (3, 4). Among them, the hypomethylation of interleukin-6 (*IL6*) promoter suggests an upregulation of this cytokine in ESCC, corroborating previous studies (5–7).

IL6 is a pleiotropic cytokine first described for its role in the induction of B cell maturation into antibody-producing cells (8). IL6 binds to the IL6 receptor that can be found anchored in the cell membrane (mbIL-6R) or in its soluble form (sIL-6R). The classical signaling is triggered by IL6 binding to mbIL-6R and the consequent dimerization of gp130. This leads to the activation of cytoplasmic tyrosine kinases that, in turn, activate the JAK/STAT axis. Alternatively, in the so-called trans-signaling, IL6 binds to sIL-6R and activates pro-inflammatory pathways, even in cells that do not express mbIL-6R. Among IL6 transcriptional targets, many genes involved in the different cancer hallmarks have been described (9). In ESCC, for example, a role for IL6 in inhibiting apoptosis through *MCL1* induction has been described (5). But, other transcriptional targets and the putative systemic effects of IL6 have not been further explored in this tumor type.

Therefore, this study aimed to evaluate *IL6* dysregulation in ESCC and to identify potential targets that could contribute to tumor phenotypes. With this, we showed *BCL3* overexpression, already detected in the non-transformed esophageal epithelium of ESCC patients, that presents the potential to be used as a detection biomarker of the cancerization field in the esophagus. Furthermore, the IL6 signaling pathway might represent an alternative therapeutic target for these patients.

MATERIAL AND METHODS

Human Samples

A total of 77 patients with a histologically confirmed diagnosis of ESCC at Instituto Nacional de Câncer (INCA, Rio de Janeiro, Brazil), Hospital Universitário Pedro Ernesto (HUPE, Rio de Janeiro, Brazil) and Hospital das Clínicas de Porto Alegre (HCPA, Porto Alegre, Brazil) were included in this study. Frozen tissue samples were obtained from tumor and tumor-surrounding mucosa (histologically normal tissue, collected 5 cm from tumor border), and, at the moment of the biopsy, patients had not undergone any chemotherapy or radiotherapy treatments. Frozen tissue samples were also obtained from individuals without esophageal disorders (n = 68) submitted to routine endoscopic examination, not related to cancer or esophageal disorders, at HUPE. Each individual donated three biopsies, one from each third of the esophagus (superior, middle, and inferior). All individuals signed informed consent, and information was obtained from medical records and/or using a standardized questionnaire, including data on tobacco smoking and alcohol drinking. Individuals were classified as ever smokers when they smoked at least one cigarette per day for over one year. Similarly, individuals were classified as ever drinkers when they drank alcoholic beverages at least twice a week for over one year (patient's characteristics are shown in **Table 1**).

For the analysis of intratumor heterogeneity, biopsies were collected from five patients submitted to endoscopy at INCA. Two fragments (profound and superficial) were collected from three different regions of the tumor mass: proximal, medial, and distal areas. Also, two biopsies of adjacent non-tumor tissue were collected 5 cm from the tumor border whenever possible, from the proximal and distal esophagus (**Figure 3A**). The study proposal was approved by the Ethics Committees of the institutions involved and was carried out according to the Helsinki Declaration.

mRNA Expression Analysis

RNeasy Micro Kit (Qiagen) was used to extract total RNA from matched tumor and surrounding tissues from 39 ESCC patients and esophageal mucosa from 46 individuals without cancer. Total RNA from cell lines was extracted using the TRIzol® reagent (Invitrogen). Five hundred nanograms of total RNA were used in reverse transcription (RT) reactions using SuperScript II Reverse Transcriptase, according to the manufacturer's instructions (Life Technologies). The following primers were used in quantitative PCR (qPCR): *BCL3* forward - 5' CGGAGCCTTACTGCCTTTGT 3', *BCL3* reverse - 5' GCCATGGCGATGTCAGCAGA 3', *IL6* forward - 5' GACCGAAGGCGCTTGTGGA 3', *IL6* reverse - 5' CTCATTCTGCCCTCGAGCC 3', *GAPDH* forward - 5' CAACAGCCTCAAGATCATCAGCAA 3', *GAPDH* reverse - 5' AGTGATGGCATGGACTGTGGTCAT 3', *HPRT1* forward - 5' CATTGTAGCCCTCTGTGTGC 3' and *HPRT1* reverse - 5' CACTATTTCTATTCAGTGCTTTGATGT 3'. All reactions were performed in a Rotor-Gene Q system (Qiagen) and consisted of 5.0 µL of QuantiFast SYBR Green PCR Mix 2X (Qiagen), 10 pmol of each primer and 1 µL of cDNA diluted 10 times in sterile deionized water, in a final volume of 10 µL. The thermal cycling program consisted of an initial denaturation for 5 min at 95°C, followed by 40 cycles of 5 s at 95°C and 10 s at 60°C. At the end of the cycling, dissociation curves were added to inspect the formation of nonspecific products and contamination. All analyses were done in triplicates, and the mean was used for further calculations. *BCL3* and *IL6* mRNA expression were calculated by the ΔCt method, using *GAPDH* or *HPRT1* as the housekeeping gene. The data was presented as $2^{-\Delta Ct}$. Undetermined Ct values for the amplification of target genes were set to 40 to allow further comparisons.

Immunohistochemistry

Immunohistochemistry was performed on paraffin sections of 38 ESCC cases, including eight cases with non-tumor surrounding tissue, and 22 healthy controls. For antigen retrieval, sections were incubated in a pressure cooker while submerged in citrate buffer, pH 6.0 for BCL3 staining and EDTA buffer, pH 8.0 for IL6 detection. Sections with 3 µm were then incubated in 3% hydrogen peroxide for 20 min and Protein Block solution for

TABLE 1 | Characteristics of the individuals included in the study.

	Individuals without esophageal disorders*	ESCC patients*
Total number	68	77
Age		
Median (min-max)	57 (18-85)	59 (39-77)
Gender		
Female	43 (69%)	20 (26%)
Male	19 (31%)	57 (74%)
Smoking Status		
Never	38 (62%)	1 (2%)
Ever	23 (38%)	45 (98%)
Drinking Status		
Never	29 (47%)	6 (14%)
Ever	33 (53%)	36 (86%)
Esophageal tumor central location		
Proximal	NA	8 (10%)
Middle	NA	59 (71%)
Distal	NA	16 (19%)
Tumor Differentiation		
in situ	NA	1 (1%)
Well	NA	0 (0%)
Moderately	NA	55 (76%)
Poorly	NA	16 (22%)
Tumor Stage		
I-II	NA	12 (29%)
III-IV	NA	34 (71%)
Type of analysis performed		
mRNA expression	41 (45%)	39 (51%)
Protein expression	22 (24%)	38 (49%)

**Numbers may vary due to missing data.*
NA, not applicable.

30 min (Dako®, Denmark) before the incubation with the primary antibody against BCL3 (Abcam® – ab49470) or IL6 (Abcam® – ab6672), overnight at 4°C. Detection and staining were performed with the Novolink™ Polymer Detection System (Leica Biosystems, UK). Sections were counterstained with Harris' hematoxylin. FFPE healthy tonsil was used as a positive control of BCL3 expression and lymph node as a positive control to IL6 detection. In the negative control, the primary antibody was replaced with the antibody diluent solution.

Cell Line and IL6 Treatment

HET-1A, a normal esophagus epithelial cell line immortalized with SV40 large T antigen, was purchased from ATCC and grown in bronchial epithelial cell growth medium (BEGM) containing all supplements provided by the manufacturer (Lonza) in flasks pre-coated with a mixture of 0.01 mg/mL fibronectin, 0.03 mg/mL bovine collagen type I and 0.01 mg/mL bovine serum albumin dissolved in the culture medium, at 37°C under 5% CO_2. Experiments were performed after the third passage after thawing. The MycoSensor qPCR Assay Kit (Agilent) was used for Mycoplasma testing, which is performed every three months as a laboratory routine.

Recombinant human IL6 was purchased from Peprotech (USA) and eluted according to the manufacturer's instructions. HET-1A cell line was treated with different concentrations of IL6 (10, 20, and 100 ng/mL) for different periods (30 min, 24, and 48 hours). Each experiment was performed three times in triplicates.

In Silico Analyses

The public data portal cBioPortal for Cancer Genomics (10, 11) was used to retrieve gene expression data from ESCC samples (n = 95). For this, the Firehose Legacy project was assessed, and Spearman correlation rho, p-values, and q-values were obtained for *IL6* mRNA levels compared to all genes from the genome. Only those correlations with q-values < 0.05 were considered statistically significant.

To evaluate whether the genes whose expression was significantly correlated with that of *IL6* were dysregulated in ESCC, previously gene expression microarray datasets (Affymetrix Human Exon 1.0 ST Array platform) generated by the group were reanalyzed (deposited in the Gene Expression Omnibus database, accession GSE75241). Differences between tumors and non-tumor surrounding tissues, as well as non-tumor surrounding tissues and healthy esophagus, were considered statistically significant when $|\text{fold-change}| > 1.5$ and FDR < 0.05.

The over-representation analysis was performed in the WEB-based Gene Set Analysis Toolkit (WebGestalt) using the Kyoto Encyclopedia of Genes and Genomes (KEGG) as a functional database. Only pathways with FDR < 0.05 were considered significantly enriched.

Statistical Analyses

All statistical analyses were performed using the GraphPad Prism 5 software (GraphPad Software, USA) or R environment. When comparing two groups, unpaired t-test or Mann-Whitney were

used for unpaired samples while paired t-test or Wilcoxon signed-rank tests were applied for paired samples. For determining significant differences between more than two groups, we have applied One-Way ANOVA or Kruskal Wallis test and Tukey's post-test or Dunn's post-test, respectively.

Receiver operating characteristic (ROC) curves were used to determine whether the proposed biomarkers were able to distinguish the sample groups.

Univariate survival analyses were carried out with the Kaplan–Meier method and Log-rank test. Gene expression cut-offs were determined according to the best performing threshold (12), being $2.53x10^{-2}$ *GAPDH* units for *IL6* expression and $1.24x10^{-2}$ *GAPDH* units for *BCL3* expression. Variables with $p < 0.2$ were selected for multivariate analysis. Finally, Cox regression was applied with the stepwise forward method (13). The 'survival' package was used.

In all cases, p values were considered statistically significant when less than 0.05.

RESULTS

Characterization of the Individuals Without Esophageal Disorders and ESCC Patients Included in This Study

Among the individuals without esophageal disorders included in this study, most were female (69%), never smokers (62%), and ever drinkers (53%) (**Table 1**). This group presented a median age of 57 years (18–85), similar to what was observed in the group of ESCC patients, with a median age of 59 years (39-77). However, ESCC patients were mostly male (74%), ever smokers (98%), and ever drinkers (86%). Regarding tumors' characteristics, most were located at the middle third of the esophagus (71%), were moderately differentiated (76%), and were mostly diagnosed in stages III and IV (71%) (**Table 1**).

IL6 Is Upregulated in ESCC, Produced by Tumor Cells and It Is Likely to Depict Transcriptional Programs

We have previously shown *IL6* promoter hypomethylation in ESCC (3) and, therefore, we aimed to investigate *IL6* expression in the groups of samples included in the present work. In **Figure 1A**, we show that ESCC presents higher *IL6* mRNA levels in comparison with histologically normal esophageal samples, either from the same ESCC patients (tumor-surrounding samples, $p < 0.0001$), or from individuals without esophageal disorders ($p < 0.0001$). Furthermore, *IL6* expression was undetectable in 100% of samples from individuals without esophageal disorders, in 60% of tumor-surrounding samples from ESCC patients, but in only one (2.7%) ESCC sample. Using a Receiver Operator Characteristic (ROC) curve, we showed that *IL6* expression was able to distinguish normal-appearing surrounding tissue from tumors with an accuracy of 93%, with 95% sensitivity and 89% specificity ($p < 0.0001$, Cut-off = $5.7x10^{-4}$ *GAPDH* relative-units, **Figure 1B**). Although *IL6* mRNA levels were not associated with clinical and tumor features (**Supplementary Table 1** and **Supplementary Figure 1**), patients with high tumor *IL6* expression showed a median overall survival (OS) of 6.97 months in comparison with a median OS of 15.5 months among patients with low *IL6* expression (**Figure 1C**). Cox regression analysis further showed that high *IL6* expression is a predictor of poorer prognosis independent of tumor stage (HR = 4.57, 95% CI= 1.54-13.55, p = 0.006; **Table 2**).

Since IL6 is a cytokine that can be produced either by immune or tumor cells (14), we decided to analyze which cells were responsible for the high *IL6* expression detected in ESCC. So, we assessed the correlation between the percentage of different immune populations and *IL6* expression in ESCC using the TCGA database. This analysis revealed that there were no significant correlations between monocytes, lymphocytes, neutrophil infiltrates, or stromal cells and *IL6* expression (**Figure 1D**). Necrosis was also not correlated with *IL6* expression (**Figure 1D**). By contrast, IL6 immunohistochemistry performed in our sample set showed positive staining only in tumor cells, with some tumor areas presenting a strong cellular membrane and cytoplasmic, and intercellular staining (probably as a result of IL6 secretion) (**Figure 1E**). Complementary, we could not detect IL6 immunostaining in the esophagus from individuals without esophageal disorders, while diffuse staining was observed in tumor-surrounding esophageal mucosa (**Figure 1E**).

Next, by using TCGA database, we identified a significant correlation between the expression of *IL6* and other 215 genes in ESCC (**Supplementary Table 2**), with 161 (75%) of them presenting a positive and 54 (25%) presenting an inverse correlation (q-value < 0.05). Among the positively correlated genes, 77 (48%) were differentially expressed in ESCC when compared to the respective tumor-surrounding esophageal mucosa (|fold-change| > 1.5 and FDR < 0.05), being 76 upregulated in tumors (**Figure 1F**). When considering the 54 inversely correlated genes, 26 (48%) were downregulated in ESCC compared to the tumor-surrounding esophageal mucosa (**Figure 1F**). An over-representation analysis, including all differentially expressed genes using the KEGG database, showed that the "TNF signaling" and "*Staphylococcus aureus* infection" pathways were enriched (FDR < 0.05).

Considering the same set of 161 genes whose expression was positively correlated with *IL6* mRNA levels, six (4%) were differentially expressed in tumor-surrounding esophagus relative to esophageal mucosa from individuals without esophageal disorders (|fold-change| > 1.5 and FDR < 0.05), with only one gene found to be downregulated among the former (*ZFAND5*). In contrast, five were upregulated (*BCL3*, highlighted in the Figure, *TIMP1*, *IFITM1*, *PMP22*, and *NINJ1*) (**Figure 1F**).

So, we decided to evaluate *BCL3* expression in our samples. **Figure 2A** shows that *BLC3* was expressed at much higher levels in tumors (~22-fold, $p < 0.0001$) and in normal tumor surrounding tissue (~25-fold, $p < 0.0001$), when compared to normal esophageal samples from patients without esophageal disorders. *BCL3* expression was not associated with etiological or clinical-pathological variables analyzed (**Supplementary Table 1**

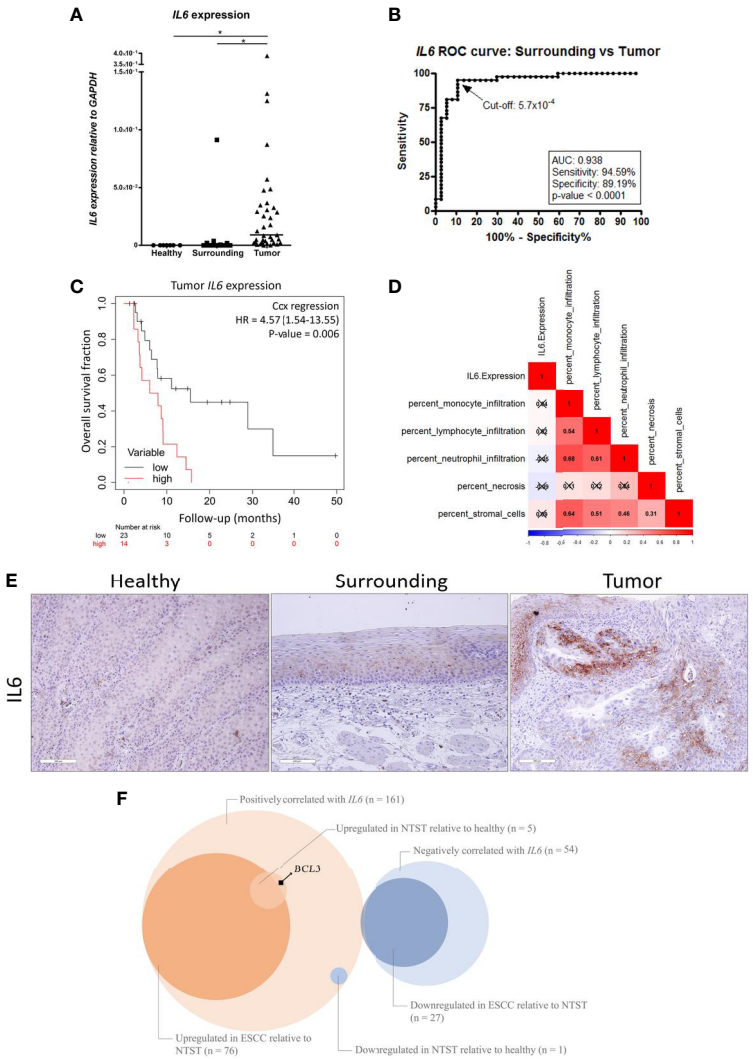


FIGURE 1 | IL6 dysregulation in ESCC. **(A)** Dot-plot representing *IL6* mRNA expression normalized with *GAPDH* in esophageal mucosa from individuals without esophageal disorders (Healthy), histologically normal tumor-surrounding tissues from ESCC patients (Surrounding), and ESCC (Tumor). **(B)** Receiver Operating Characteristic (ROC) curve showing the sensitivity and specificity of *IL6* mRNA expression normalized with *GAPDH* to distinguish Surrounding and Tumor samples from ESCC patients. IL6 expression cut-off of 5.7x10-4 showed the highest accuracy (93%), sensitivity (95%) and specificity (89%) ($p < 0.0001$). **(C)** Kaplan-Meier curve showing the overall survival of ESCC patients according to *IL6* expression. High expression was defined as ≥ $2.53x10^{-2}$ *GAPDH* units, according to the best performing threshold. Tumor stage was used for adjustment. **(D)** Correlation matrix between *IL6* expression and the percentage of different cellular populations in ESCC samples from the TCGA database. Red represents positive correlations, and blue represents inverse correlations. Boxes marked with an X depict non-significant correlations ($p \geq 0.05$). **(E)** Representative FFPE slides from immunohistochemistry performed with IL6 antibody in esophageal mucosa from individuals without esophageal disorders (Healthy), histologically normal tumor-surrounding tissues from ESCC patients (Surrounding), and ESCC (Tumor). Sections were counterstained with Harris' hematoxylin. All images are shown in 20X magnification and 300 μm scale bars are shown. **(F)** Venn diagram showing the number of genes whose mRNA expression was positively or negatively correlated with *IL6* mRNA expression (q-value < 0.05) in ESCC samples from TCGA database, number of genes upregulated or downregulated in ESCC relative to non-tumor surrounding tissue (NSTS) (|fold-change| > 1.5 and FDR < 0.05) in our dataset, and number of genes upregulated or downregulated in NTST relative to healthy esophageal mucosa (|fold-change| > 1.5 and FDR < 0.05) in our dataset. *BCL3* is highlighted as a gene whose mRNA expression was positively correlated with *IL6* expression in ESCC, and that was found upregulated in NTST relative to the healthy esophagus. *$p < 0.05$.

and **Supplementary Figure 1**). BCL3 protein expression was assessed by immunohistochemistry, confirming its lack of expression in esophageal epithelium from individuals without esophageal disorders. At the same time, both ESCC and tumor-surrounding mucosa showed positive nuclear staining (**Figure 2B**). We also evaluated the efficiency of *BCL3* expression to distinguish esophageal epithelium from individuals without esophageal disorders from the histologically normal mucosa from ESCC patients using a ROC curve. With a cut-off of $1.3x10^{-3}$, *BCL3* expression relative to *GAPDH* distinguished histologically normal samples from individuals without esophageal disorders from those with

TABLE 2 | Overall survival analyses.

Feature	Category	Univariate Analysis			Multivariate Analysis		
		HR	95% CI	p-value	HR	95% CI	p-value
Age (years)	< 60 *vs* ≥ 60	0.91	0.42 - 1.98	0.8			
Tumor stage	III & IV *vs* I & II	2.59	0.71 - 9.44	0.1	3.51	0.90 - 13.67	0.069
Histologic grade	G3 *vs* G2	0.46	0.1 - 2.0	0.3			
IL6 expression*	High *vs* Low	3.19	1.38 - 7.36	**0.004**	4.57	1.54 - 13.55	**0.006**
BCL3 expression*	High *vs* Low	1.48	0.68 – 3.21	0.31			

**IL6 and BCL3 expression cut-offs were determined according to the best performing threshold, being 2.53x10^{-2} GAPDH units and 1.24 x10^{-2} GAPDH units, respectively.*
HR, hazard ratio; CI, confidence interval.
Bold numbers represent p < 0.05.

ESCC, with an accuracy of 96% (100% sensitivity and 95.12% specificity, p < 0.0001, **Figure 2C**). *BCL3* expression in the surrounding tissue was positively correlated with *IL6* expression in the tumors (Spearman r = 0.4135, p = 0.0122; **Figure 2D**).

To verify whether IL6 could modulate *BCL3* expression in the esophagus, we used a non-transformed esophageal cell line (HET-1A). We showed that this cytokine could significantly induce *BCL3* mRNA expression in all tested concentrations and time intervals, except for the treatment with 10 ng/mL for 24 hours (**Figure 2E**).

Spatial Intratumoral Analysis for *IL6* and *BCL3* Expression

In order to evaluate the spatial intratumoral pattern of *IL6* and *BCL3* expression in ESCC, we used three to six tumor biopsies from different regions of the tumoral mass and two fragments of non-tumoral surrounding tissue, from five patients with ESCC (**Figure 3A**). Corroborating the results presented in **Figure 1A**, *IL6* overexpression was observed in all tumoral fragments of all patients, compared to their respective non-tumoral surrounding tissues (**Figure 3B**). Only in two tumor fragments from Patient 5, *IL6* expression did not surpass the previously established cut-off for differentiating tumor and non-tumor surrounding tissue (using the ROC curve cut-off = 5.7x10^{-4} *GAPDH* relative-units) (**Figure 3B**). Spatial intratumoral heterogeneity was observed for *IL6* expression within the tumor mass in four out of five patients. In addition, using the previous cut-off of 1.3x10^{-3} *GAPDH* relative-units defined to distinguish morphologically normal tissues from healthy individuals and ESCC patients, *BCL3* expression exceeded this cut-off in every analyzed fragment of tumoral and non-tumoral tissues (**Figure 3C**).

DISCUSSION

IL6 is a pleiotropic cytokine upregulated in different tumor types, and it can depict several transcriptional programs. Here we propose IL6 produced by ESCC tumors cells might induce the expression of several genes and downregulate a number of others in a juxtacrine fashion, but also induce the expression of genes involved in key cancer-associated pathways in the non-tumor surrounding tissue. One of these targets might be *BCL3*, a recognized oncogene for hematological malignancies, whose alterations have been associated with tumor progression and poor prognosis (15, 16). More recently, the role of *BCL3* in solid tumors started to be elucidated, and its overexpression, observed in breast and nasopharyngeal tumors, seems to be induced by NF-κB activation and EBV latent genes, respectively (17–19). Here, we show for the first time *BCL3* upregulation in esophageal squamous cell carcinoma and its high potential as a diagnostic biomarker, since it precedes the first histopathological alterations.

A key aspect of esophageal carcinogenesis is the development of synchronic tumors that has been attributed to field cancerization. This concept was first proposed by Slaughter and colleagues (20) to explain a similar phenomenon observed in oral carcinogenesis and assumes that the entire organ, in this case, the entire esophagus could present patches of premalignant cells that, although may not show any morphological alteration, already carry molecular alterations that predispose to neoplastic transformation. In this context, *BCL3* overexpression in the normal-appearing tumor-surrounding tissue observed in this study could represent one of these predisposing molecular alterations. It has been proposed before by us and other authors that epigenetic alterations, specifically alterations of DNA methylation, could be the first disturbance to occur during the formation of the cancerization field in the esophagus (3, 4, 21). Recently, we have shown the great potential of *TFF1* expression downregulation, one of the genes epigenetically deregulated, as an early ESCC detection biomarker, since it takes place in the tumor-surrounding tissue from ESCC patients, preceding the first morphological and genetic alterations (22).

Along with these findings, we demonstrated that *BCL3* expression is able to distinguish healthy esophagus from non-tumor tissue adjacent to ESCC with high accuracy, sensitivity, and specificity and could help to identify the patches of premalignant esophageal cells. Moreover, using the proposed cut-off for mRNA analysis, *BCL3* overexpression was consistent in non-tumor adjacent tissues and tumors in the spatial intratumoral analysis, showing its robust potential as a biomarker of diagnosis. This could be particularly useful in patients with head and neck cancer since these individuals show a relatively high incidence of second primary tumors in the esophagus, impacting their overall survival (23). Another important point to discuss is the currently available methods for ESCC early diagnosis. Although the impact of downstaging through screening on patient mortality is clear, there is no globally accepted method. Endoscopy following lugol staining (chromoendoscopy) is one of the most commonly used approaches

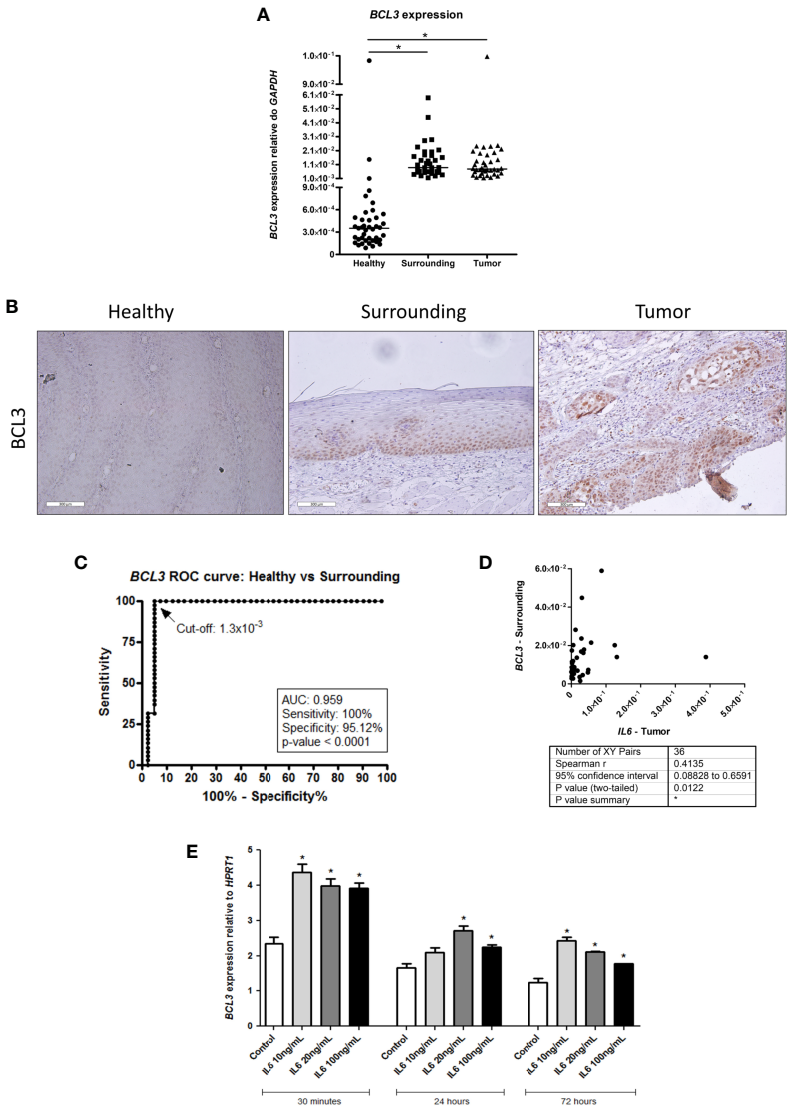


FIGURE 2 | BCL3 expression in healthy esophagus and ESCC. **(A)** Dot-plot representing *BCL3* mRNA expression normalized with *GAPDH* in esophageal mucosa from individuals without esophageal disorders (Healthy), histologically normal tumor-surrounding tissues from ESCC patients (Surrounding), and ESCC (Tumor). **(B)** Representative FFPE slides from immunohistochemistry performed with BCL3 antibody in esophageal mucosa from individuals without esophageal disorders (Healthy), histologically normal tumor-surrounding tissues from ESCC patients (Surrounding), and ESCC (Tumor). Sections were counterstained with Harris' hematoxylin. All images are shown in 20X magnification and 300 µm scale bars are shown. **(C)** Receiver Operating Characteristic (ROC) curve showing the sensitivity and specificity of *BCL3* mRNA expression normalized with *GAPDH* to distinguish Healthy from Surrounding. *BCL3* expression cut-off of 1.3x10-3 showed the highest accuracy (95.89%), sensitivity (100%) and specificity (95.12%) ($p < 0.0001$). **(D)** Correlation analysis between IL6 expression in tumor samples (X-axis) and *BCL3* expression in the non-tumor surrounding tissue (Y-axis) from the same patient. Each dot represents an ESCC patient. **(E)** Bar graphs showing *BCL3* mRNA expression normalized with *HPRT1* in HET-1A esophageal cells after IL6 treatment in different doses for different periods of time. *$p < 0.05$.

and shows high sensitivity, but low specificity (24). Therefore, molecular alterations that are detectable before morphological changes and are maintained during tumor development (independently of tumor stage) can be useful in helping to identify those individuals at risk, potentially increasing specificity of the screening methods.

In the present study, we also observed a positive correlation between *IL6* expression in tumors and *BCL3* expression in tumor-adjacent samples of ESCC patients, suggesting a possible dysregulation of this axis in the early stages of esophageal carcinogenesis. It has been shown before that IL6 is capable of inducing *BCL3* transcription in multiple myeloma cells *via* STAT3 binding to an enhancer in the *BCL3* gene body (24). Since *IL6* overexpression is only observed in ESCC samples, while *BCL3* higher mRNA levels are detected both in tumor-surrounding mucosa and esophageal tumors, paracrine IL6 signaling could be responsible for *BCL3* induction in tumor-adjacent tissue. In agreement with this hypothesis, we have

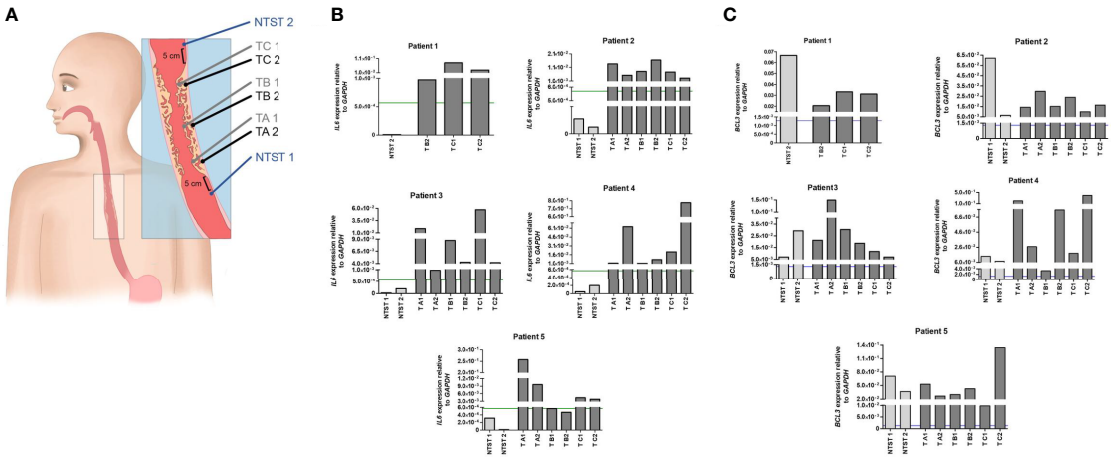


FIGURE 3 | Intratumor heterogeneity of *IL6* and *BCL3* expression in ESCC. **(A)** Schematic representation of sample collection by endoscopy for the evaluation of intratumor heterogeneity. From each patient, a total of eight biopsies was collected when possible. Non-tumor surrounding tissue samples were collected 5 cm below (NTST 1) and above (NIST 2) the tumor border. ESCC samples were collected from each third of the tumor mass (lower third, TA; middle third, TB; and upper third, TC), being one superficial (1) and another profound (2), the latter collected by a biopsy-on-biopsy scheme. **(B)** Bar graphs representing *IL6* expression normalized with *GAPDH* in the different non-tumor surrounding tissue (NTST) and tumor biopsies from five patients diagnosed with ESCC (Patient 1-5). The green line represents the *IL6* expression cut-off that showed the highest accuracy to differentiate the tumor tissue from the non-tumor surrounding tissue, as shown in **Figure 1B** ($5.7x10^{-4}$). **(C)** Bar graphs representing *BCL3* expression normalized with *GAPDH* in the different non-tumor surrounding tissue (NTST) and tumor biopsies from five patients diagnosed with ESCC (Patient 1-5). The blue line represents the *BCL3* expression cut-off that showed the highest accuracy to differentiate the non-tumor surrounding tissue from the esophageal mucosa from individuals without esophageal disorders, as shown in **Figure 2C** ($1.3x10^{-3}$).

shown that this cytokine can induce *BCL3* expression in immortalized esophageal cells. However, this inflammatory pathway may not be the only mechanism involved in *BCL3* regulation in ESCC tissues, and further studies should be performed to confirm this hypothesis.

BCL3 has been shown to contribute to carcinogenesis through different mechanisms and mediators. Although most studies have shown that BCL3 promotes cell survival and proliferation of different cell types for its role as a key regulator of NF-κB signaling (25, 26), recent works have shown that BCL3 contribution to cancer goes beyond this role in these two cancer hallmarks. So, in mouse embryonic stem cell (mESCs), BCL3 seems to regulate proliferation and pluripotency by influencing Nanog transcription (27). A similar effect has been observed in colorectal cancer cells, in which BCL3 acts as a co-activator of β-catenin/TCF-mediated transcriptional activity and induces the expression of stemness markers such as *LGR5* (Leucine-Rich Repeat Containing G Protein-Coupled Receptor 5) and *ASCL2* (Achaete-Scute Family BHLH Transcription Factor 2) (28). Other BCL3 transcriptional targets include *PD-L1* (Programmed Cell Death 1 Ligand 1) and *CAII* (Carbonic Anhydrase II) that mediate cell proliferation and resistance to alkylating agents in ovarian cancer and gliomas, respectively (29, 30). Therefore, *BCL3* overexpression before esophageal transformation may contribute not only to apoptosis inhibition and induced proliferation capacity, but also to the acquisition of stemness phenotypes. The maintenance of its high expression levels in tumors should also be further investigated in terms of immune escape and resistance to therapy.

Our study showed a consistent *IL6* overexpression in all ESCC patients, and the ability to distinguish the surrounding mucosa from tumoral tissue with high sensitivity and specificity. *IL6* upregulation in esophageal cancer has been shown before by other authors, and it has been proposed that serum levels of this cytokine could be used as a diagnostic biomarker in ESCC, with higher efficiency than classic tumor markers (carcinoembryonic and squamous cell cancer antigens), and could also predict overall and disease-free survival (31–33). Our data corroborated the association of high *IL6* expression with a poor prognosis in ESCC. Also, it has been proposed that IL6 could be a therapeutic target in ESCC. Currently, two drugs targeting IL6 signaling are clinically registered, siltuximab, an anti-IL6 mAb, and tocilizumab, an anti-IL6R mAb. Although the FDA approved the treatment of patients with multicentric Castleman disease with siltuximab (34), the studies in cancer patients are still scarce.

Early diagnosis of esophageal squamous cell carcinoma has not improved in recent years, indicating not only a lack of knowledge on the mechanisms of early alterations and development of ESCC, but also the difficulty in identifying and validating robust biomarkers of early tumor detection. According to data from us and other authors (3, 4, 35–37), inflammation seems to be an essential hallmark of ESCC and may provide essential biomarkers to anticipate diagnosis, predict prognosis, and may represent new therapeutic targets.

In conclusion, our study shows the consistent *IL6* overexpression in ESCC cells, suggesting its potential use as a therapeutic target for ESCC, and the paracrine induction of the expression of the antiapoptotic gene *BCL3*, revealing alterations that may contribute to ESCC development. Furthermore, our study highlights a set of data to support the use of *BCL3* mRNA expression as a biomarker of ESCC detection, suggesting further

studies should be performed to corroborate these findings in high-risk groups for ESCC development, as head and neck cancer patients.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/geo/, GSE75241 https://www.cbioportal.org/study/summary?id=esca_tcga, Esophageal Carcinoma (TCGA, Firehose Legacy).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comissão Nacional de Ética em Pesquisa - CAAE 0086.0.007.000-11 Comitê de Ética em Pesquisa do Instituto Nacional de Câncer - CEP n° 116/11. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conceptualization, ZH, HH-V, LR, and SS-L. Methodology, IG, SG, HH-V, and SS-L. Validation, IG, PN-N, and RV. Formal analysis, IG, DC, PN-N, RV, and SS-L. Investigation, IG, DC, MF, SG, and RV. Resources, ZH, LP, and SS-L. Data curation, IG, DC, and SS-L. Writing—original draft preparation, IG, DC, and SS-L. Writing—review and editing, IG, DC, PN-N, RV, MF, SG, HH-V, ZH, LR, and SS-L. Visualization, IG, DC, RV, HH-V, ZH, LR, and SS-L. Supervision, LR and SS-L. Project administration, LR and SS-L. Funding acquisition, LR, ZH, and SS-L. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) grant E-26/010.001856/2015 and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant 407992/2016-2.

ACKNOWLEDGMENTS

The authors would like to thank the Tumor National Bank from INCA (BNT/INCA) for all the support with sample processing and the Endoscopy and Pathology Sections from INCA for all the support with samples collection and selection. The authors also thank the graphical designer Luiza Morena Moraes for the intratumor heterogeneity biopsy collection illustration.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.722417/full#supplementary-material

Supplementary Figure 1 | *IL6* and *BCL3* expression association with tumor staging and impact on overall survival. Dot plots representing *IL6* and *BCL3* expression relative to *GAPDH* according to T **(A)**, N **(B)** and M **(C)** from TNM staging system.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin* (2018) 68 (6):394–424. doi: 10.3322/caac.21492
2. Rustgi AK, El-Serag HB. Esophageal Carcinoma. *New Engl J Med* (2014) 371 (26):2499–509. doi: 10.1056/NEJMra1314530
3. Lima SCS, Hernández-Vargas H, Simão T, Durand G, Kruel CDP, Calvez-Kelm FL, et al. Identification of a DNA Methylome Signature of Esophageal Squamous Cell Carcinoma and Potential Epigenetic Biomarkers. *Epigenetics* (2011) 6(10):1217–27. doi: 10.4161/epi.6.10.17199
4. Talukdar FR, Soares Lima SC, Khoueiry R, Laskar RS, Cuenin C, Sorroche BP, et al. Genome-Wide DNA Methylation Profiling of Esophageal Squamous Cell Carcinoma From Global High-Incidence Regions Identifies Crucial Genes and Potential Cancer Markers. *Cancer Res* (2021) 81(10):2612–24. doi: 10.1158/0008-5472.CAN-20-3445
5. Leu C-M, Wong F-H, Chang C, Huang S-F, Hu C. Interleukin-6 Acts as an Antiapoptotic Factor in Human Esophageal Carcinoma Cells Through the Activation of Both STAT3 and Mitogen-Activated Protein Kinase Pathways. *Oncogene* (2003) 22(49):7809–18. doi: 10.1038/sj.onc.1207084
6. Karakasheva TA, Lin EW, Tang Q, Qiao E, Waldron TJ, Soni M, et al. IL-6 Mediates Cross-Talk Between Tumor Cells and Activated Fibroblasts in the Tumor Microenvironment. *Cancer Res* (2018) 78(17):4957–70. doi: 10.1158/0008-5472.CAN-17-2268
7. Yang P-W, Huang P-M, Yong L-S, Chang Y-H, Wu C-W, Hua K-T, et al. Circulating Interleukin-6 Is Associated With Prognosis and Genetic Polymorphisms of MIR608 in Patients With Esophageal Squamous Cell Carcinoma. *Ann Surg Oncol* (2018) 25(8):2449–56. doi: 10.1245/s10434-018-6532-4
8. Hirano T, Taga T, Nakano N, Yasukawa K, Kashiwamura S, Shimizu K, et al. Purification to Homogeneity and Characterization of Human B-Cell Differentiation Factor (BCDF or BSFp-2). *Proc Natl Acad Sci USA* (1985) 82(16):5490–4. doi: 10.1073/pnas.82.16.5490
9. Kumari N, Dwarakanath BS, Das A, Bhatt AN. Role of Interleukin-6 in Cancer Progression and Therapeutic Resistance. *Tumour Biol: J Int Soc Oncodevelopmental Biol Med* (2016) 37(9):11553–72. doi: 10.1007/s13277-016-5098-7
10. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* (2012) 2(5):401–4. doi: 10.1158/2159-8290.CD-12-0095
11. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the Cbioportal. *Sci Signaling* (2013) 6(269):pl1. doi: 10.1126/scisignal.2004088
12. Lanczky A, Gyorffy B. Kmplot.Com: A Survival Analysis Tool Tailored for Medical Research. *J Med Internet Res* (2021) 23(7):e27633. doi: 10.2196/27633
13. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival Analysis Part III: Multivariate Data Analysis—Choosing a Model and Assessing Its Adequacy and Fit. *Br J Cancer* (2003) 89:605–11. doi: 10.1038/sj.bjc.6601120

14. Tanaka T, Narazaki M, Kishimoto T. IL-6 in Inflammation, Immunity, and Disease. *Cold Spring Harbor Perspect Biol* (2014) 6(10):a016295. doi: 10.1101/cshperspect.a016295
15. De Keersmaecker K, Michaux L, Bosly A, Graux C, Ferreiro JF, Vandenberghe P, et al. Rearrangement of NOTCH1 or BCL3 Can Independently Trigger Progression of CLL. *Blood* (2012) 119(16):3864–6. doi: 10.1182/blood-2011-10-388124
16. Au WY, Horsman DE, Ohno H, Klasa RJ, Gascoyne RD. Bcl-3/IgH Translocation (14;19)(Q32;Q13) in Non-Hodgkin's Lymphomas. *Leukemia Lymphoma* (2002) 43(4):813–6. doi: 10.1080/10428190290016935
17. Cogswell PC, Guttridge DC, Funkhouser WK, Baldwin AS. Selective Activation of NF-Kappa B Subunits in Human Breast Cancer: Potential Roles for NF-Kappa B2/P52 and for Bcl-3. *Oncogene* (2000) 19(9):1123–31. doi: 10.1038/sj.onc.1203412
18. Thornburg NJ, Pathmanathan R, Raab-Traub N. Activation of Nuclear Factor-KappaB P50 Homodimer/Bcl-3 Complexes in Nasopharyngeal Carcinoma. *Cancer Res* (2003) 63(23):8293–301.
19. Chung GT-Y, Lou WP-K, Chow C, To K-F, Choy K-W, Leung AW-C, et al. Constitutive Activation of Distinct NF-κB Signals in EBV-Associated Nasopharyngeal Carcinoma. *J Pathol* (2013) 231(3):311–22. doi: 10.1002/path.4239
20. Slaughter DP, Southwick HW, Smejkal W. Field Cancerization in Oral Stratified Squamous Epithelium; Clinical Implications of Multicentric Origin. *Cancer* (1953) 6(5):963–8. doi: 10.1002/1097-0142(195309)6:5<963::aid-cncr2820060515>3.0.co;2-q
21. Lee Y-C, Wang H-P, Wang C-P, Ko J-Y, Lee J-M, Chiu H-M, et al. Revisit of Field Cancerization in Squamous Cell Carcinoma of Upper Aerodigestive Tract: Better Risk Assessment With Epigenetic Markers. *Cancer Prev Res* (2011) 4(12):1982–92. doi: 10.1158/1940-6207.CAPR-11-0096
22. Gonzaga IM, Soares Lima SC, Nicolau MC, Nicolau-Neto P, da Costa NM, de Almeida Simão T, et al. TFF1 Hypermethylation and Decreased Expression in Esophageal Squamous Cell Carcinoma and Histologically Normal Tumor Surrounding Esophageal Cells. *Clin Epigenet* (2017) 9(1):130. doi: 10.1186/s13148-017-0429-0
23. Lampri ES, Chondrogiannis G, Ioachim E, Varouktsi A, Mitselou A, Galani A, et al. Biomarkers of Head and Neck Cancer, Tools or a Gordian Knot? *Int J Clin Exp Med* (2015) 8(7):10340–5724. doi: 10.4253/wjge.v9.i9.438
24. Maldonado V, Melendez-Zajgla J. Role of Bcl-3 in Solid Tumors. *Mol Cancer* (2011) 10(1):152. doi: 10.1186/1476-4598-10-152
25. Brocke-Heidrich K, Ge B, Cvijic H, Pfeifer G, Löffler D, Henze C, et al. BCL3 Is Induced by IL-6 *via* Stat3 Binding to Intronic Enhancer HS4 and Represses Its Own Transcription. *Oncogene* (2006) 25(55):55. doi: 10.1038/sj.onc.1209711
26. Franzoso G, Bours V, Park S, Tomita-Yamaguchi M, Kelly K, Siebenlist U. The Candidate Oncoprotein Bcl-3 Is an Antagonist of P50/NF-Kappa B-Mediated Inhibition. *Nature* (1992) 359(6393):339–42. doi: 10.1038/359339a0
27. Kang S, Yun J, Kim DY, Jung SY, Kim YJ, Park JH, et al. Adequate Concentration of B Cell Leukemia/Lymphoma 3 (Bcl3) is Required for Pluripotency and Self-Renewal of Mouse Embryonic Stem Cells *via* Downregulation of Nanog Transcription. *BMB Rep* (2018) 51(2):92–7. doi: 10.5483/bmbrep.2018.51.2.219
28. Legge DN, Shephard AP, Collard TJ, Greenhough A, Chambers AC, Clarkson RW, et al. BCL-3 Promotes a Cancer Stem Cell Phenotype by Enhancing β-Catenin Signalling in Colorectal Tumour Cells. *Dis Models Mech* (2019) 12 (3):dmm037697. doi: 10.1242/dmm.037697
29. Zou Y, Uddin MM, Padmanabhan S, Zhu Y, Bu P, Vancura A, et al. The Proto-Oncogene Bcl3 Induces Immune Checkpoint PD-L1 Expression, Mediating Proliferation of Ovarian Cancer Cells. *J Biol Chem* (2018) 293 (40):15483–96. doi: 10.1074/jbc.RA118.004084
30. Wu L, Bernal GM, Cahill KE, Pytel P, Fitzpatrick CA, Mashek H, et al. BCL3 Expression Promotes Resistance to Alkylating Chemotherapy in Gliomas. *Sci Trans Med* (2018) 10(448):eaar2238. doi: 10.1126/scitranslmed.aar2238
31. Wang L-S, Chow K-C, Wu C-W. Expression and Up-Regulation of Interleukin-6 in Oesophageal Carcinoma Cells by N-Sodium Butyrate. *Br J Cancer* (1999) 80(10):10. doi: 10.1038/sj.bjc.6690571
32. Fujiwara H, Suchi K, Okamura S, Okamura H, Umehara S, Todo M, et al. Elevated Serum CRP Levels After Induction Chemoradiotherapy Reflect Poor Treatment Response in Association With IL-6 in Serum and Local Tumor Site in Patients With Advanced Esophageal Cancer. *J Surg Oncol* (2011) 103 (1):62–8. doi: 10.1002/jso.21751
33. Łukaszewicz-Zając M, Mroczko B, Kozłowski M, Niklíński J, Laudański J, Szmitkowski M. Higher Importance of Interleukin 6 Than Classic Tumor Markers (Carcinoembryonic Antigen and Squamous Cell Cancer Antigen) in the Diagnosis of Esophageal Cancer Patients. *Dis Esophagus: Off J Int Soc Dis Esophagus* (2012) 25(3):242–9. doi: 10.1111/j.1442-2050.2011.01242.x
34. Deisseroth A, Ko C-W, Nie L, Zirkelbach JF, Zhao L, Bullock J, et al. FDA Approval: Siltuximab for the Treatment of Patients With Multicentric Castleman Disease. *Clin Cancer Research: Off J Am Assoc Cancer Res* (2015) 21(5):950–4. doi: 10.1158/1078-0432.CCR-14-1678
35. Maghsudlu M, Farashahi Yazd E. Heat-Induced Inflammation and its Role in Esophageal Cancer. *J Digestive Dis* (2017) 18(8):431–44. doi: 10.1111/1751-2980.12511
36. Diakowska D. Cytokines Association With Clinical and Pathological Changes in Esophageal Squamous Cell Carcinoma. *Dis Markers* (2013) 35(6):883–93. doi: 10.1155/2013/302862
37. Rapozo DCM, Blanco TCM, Reis BB, Gonzaga IM, Valverde P, Canetti C, et al. Recurrent Acute Thermal Lesion Induces Esophageal Hyperproliferative Premalignant Lesions in Mice Esophagus. *Exp Mol Pathol* (2016) 100(2):325–31. doi: 10.1016/j.yexmp.2016.02.005

Author Disclaimer: Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

REVIEW
published: 30 August 2021
doi: 10.3389/fonc.2021.732702

Small Extracellular Vesicles in the Development, Diagnosis, and Possible Therapeutic Application of Esophageal Squamous Cell Carcinoma

Zheng Zhao†, Shuyue Yang†, Anni Zhou, Xiao Li, Rui Fang, Shutian Zhang, Guiping Zhao and Peng Li**

Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Die Wang,
Hudson Institute of Medical Research, Australia

Reviewed by:
Jafar Rezaie,
Urmia University of Medical Sciences, Iran
Katie Meehan,
The Chinese University of Hong Kong, China
Suna Zhou,
Wenzhou Medical University, China

***Correspondence:**
Peng Li
lipeng@ccum.edu.cn
Guiping Zhao
zhaoguiping@ccmu.edu.cn

†These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *29 June 2021*
Accepted: *09 August 2021*
Published: *30 August 2021*

Citation:
Zhao Z, Yang S, Zhou A, Li X, Fang R, Zhang S, Zhao G and Li P (2021) Small Extracellular Vesicles in the Development, Diagnosis, and Possible Therapeutic Application of Esophageal Squamous Cell Carcinoma. Front. Oncol. 11:732702. doi: 10.3389/fonc.2021.732702

Esophageal squamous cell carcinoma (ESCC) persists among the most lethal and broad-spreading malignancies in China. The exosome is a kind of extracellular vesicle (EV) from about 30 to 200 nm in diameter, contributing to the transfer of specific functional molecules, such as metabolites, proteins, lipids, and nucleic acids. The paramount role of exosomes in the formation and development of ESCC, which relies on promoting intercellular communication in the tumor microenvironment (TME), is manifested with immense amounts. Tumor-derived exosomes (TDEs) participate in most hallmarks of ESCC, including tumorigenesis, invasion, angiogenesis, immunologic escape, metastasis, radioresistance, and chemoresistance. Published reports have delineated that exosome-encapsulated cargos like miRNAs may have utility in the diagnosis, as prognostic biomarkers, and in the treatment of ESCC. This review summarizes the function of exosomes in the neoplasia, progression, and metastasis of ESCC, which improves our understanding of the etiology and pathogenesis of ESCC, and presents a promising target for early diagnostics in ESCC. However, recent studies of exosomes in the treatment of ESCC are sparse. Thus, we introduce the advances in exosome-based methods and indicate the possible applications for ESCC therapy in the future.

Keywords: esophageal squamous cell carcinoma, exosome, exosomal RNAs, biomarker, diagnosis

INTRODUCTION

Esophageal cancer ranks the seventh most prevalent malignancy and the sixth-highest cancer-related mortality globally (1). Esophageal cancer is broadly divided into esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (2). The incidence of esophageal cancer in men is approximately three to four times that in women and varies among countries (3). The highest rates of ESCC are found in Eastern Asia, where ESCC accounts for more than 90% of esophageal cancers (4). Despite significant progress in diagnosis and treatment, ESCC is often identified late, which leads to delayed treatment. The reason for this challenge can be attributable to multiple aspects: in the early stage, ESCC is characterized by a lack of specific symptoms and definitive

diagnosis; coming to the advanced stage, ESCC can exhibit considerable metastatic potential and strong resistance to conventional treatment. Given the above, ESCC is commonly presented with a poor prognosis, as the 5-year survival rate of late-stage ESCC is approximately 10–20% (5). Hence, more effort is urgently needed to reveal the mechanisms of tumorigenesis and increase the early diagnosis rate of ESCC.

In the last decade, there has been a steep increase in the investigations focusing on extracellular vesicles' (EVs') physiological and pathological functions, referring to multiple subtypes of cell-released, membranous structures (6–8). In particular, exosome is among the most studied and deliberated population of EVs in the rapidly growing number of publications. Notably, MISEV2018 guidelines have endorsed that the term "exosome" should be applied strictly to an EV of endosomal origin owing to the difficulties to confirm such an origin after an EV has left the cell (8, 9). According to the conventional description of exosome (from about 30 to 200 nm in diameter) (10), it would be more appropriate to nominate "exosome" as "small extracellular vesicle (sEV)", as the guidelines propose. However, considering the number of studies published before the criteria were issued, we decide to preserve "exosome", referring to "small extracellular vesicle of endosomal origin", in this review to help readers adapt to the new standard.

With molecular heterogeneity, exosomes encapsulate diverse bioactive molecules, ranging from nucleic acid (including DNA and RNA) to proteins, lipids, and other metabolites (11). Exosomes have mediated a new paradigm of intercellular communication *via* the transfer of bioactive molecules from donor cells to recipient cells, and they function in both normal physiology and acquired pathological activities, such as reproduction, immune responses, metabolic and cardiovascular diseases, ischemic diseases, neurodegeneration, and malignant tumors (12–17). In the process of tumorigenesis, exosomes can participate in the formation of the tumor microenvironment (TME), the proliferation of cancer cells, angiogenesis, metastasis, therapy resistance, and many other physiological and pathological processes (18). The amounts and cargos of exosomes derived from the same cell can dramatically vary from different conditions, and the heterogeneity of exosomal cargos has been recognized among different individuals (19–21). Considering the homogeneity between exosomes and donor cells, these cargos of the tumor-derived exosomes (TDEs) carry cancer-related information and allow them to fulfill diagnostic functions, serving as tumor biomarkers of ESCC that can be detected in early-stage cancer (22, 23). Moreover, the specialty of exosomes in delivering diverse and specific functional cargos into recipient cells has accelerated their clinical application in the therapy of patients with malignant tumors or other diseases (24–26). Ongoing studies and trials have proven that exosomes can be engineered to carry specific lipids, proteins, and other chemotherapeutic agents to targeted cells or organs and facilitate the treatment of several diseases (27–30).

The primary objectives of this review are to introduce the biology of exosomes, summarize the function of exosome-carried cargos in the initiation and development of ESCC, and discuss the potential clinical applications in both the early diagnosis and treatment of ESCC (**Figure 1**).

BIOLOGY OF EXOSOME (sEVs)

Exosomes, which should have been called sEVs following MISEV2018 guidelines, are a kind of lipid bilayer-encapsulated, nanosized vesicles that are enriched in specific DNA, RNA, lipids, proteins, and bioactive compounds (10).

Generally, the biogenesis and releasing of exosomes involve the double invagination of the plasma membrane and the sequential generation of multivesicular bodies (MVBs) and

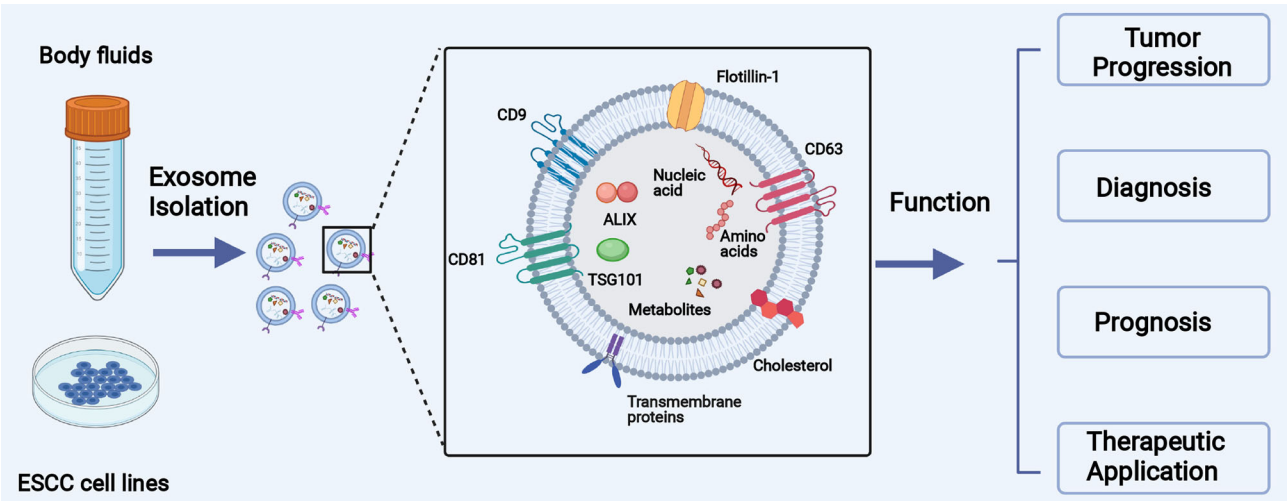


FIGURE 1 | The main objective of this review is to introduce the biology of exosomes, summarize the function of exosomal cargos in the initiation and development of ESCC, and collect the clinical application of exosomes in the diagnosis and treatment of ESCC.

intraluminal vesicles (ILVs) (11) (**Figure 2**). The first invagination of the plasma membrane forms a cup-shaped structure, containing cell-surface proteins and soluble proteins from the extracellular milieu. Then, the cup-shaped plasma membrane buds in the inner side of the cell, which gives rise to an early-sorting endosome (ESE) and, sometimes, may directly fuse with a preexisting ESE. Meanwhile, the endoplasmic reticulum (ER), mitochondria, and trans-Golgi network (TGN) also engage with the formation of the ESEs. Furthermore, the ESEs can also blend into the ER and TGN, possibly interpreting how the extracellular and cell-surface ingredients enter them (31–36). Afterward, ESEs form late-sorting endosomes (LSEs) and subsequently give rise to MVBs (also named as multivesicular endosomes). MVBs come into being with the inward invagination of the endosomal limiting membrane, the double invagination of the plasma membrane exactly, and they will be released as intraluminal vesicles (ILVs) after fusion with the plasma membrane. During the process, cytoplasmic constituents can enter the newly forming ILVs, leading to further changes in the future exosomal cargos (32, 37). Some proteins, including endosomal sorting complexes required for transport proteins (ESCRT), soluble N-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs), apoptosis-linked gene 2-interacting protein X (ALIX), tumor susceptibility gene 101 (TSG101), Rab GTPases, CD9, CD63, and CD8 1, play a critical part in the origin and biogenesis of exosomes, and some are regarded as markers of exosomes (38, 39). After being secreted into the extracellular milieu, exosomes are delivered and identified by the targeted recipient cells. As a result, they alter the phenotype and biological response of these recipient cells (40). The mechanism of exosome uptake is complex; the fate of the exosomal contents and the molecular alterations induced in recipient cells add complexity to the cell–cell crosstalk (41). When docking the recipient cell, exosomes can activate signaling pathways by straightly interacting with the receptors on the cell surface, directly fusing with the plasma membrane, or getting internalized (42). Firstly, the interaction between exosomes and extracellular

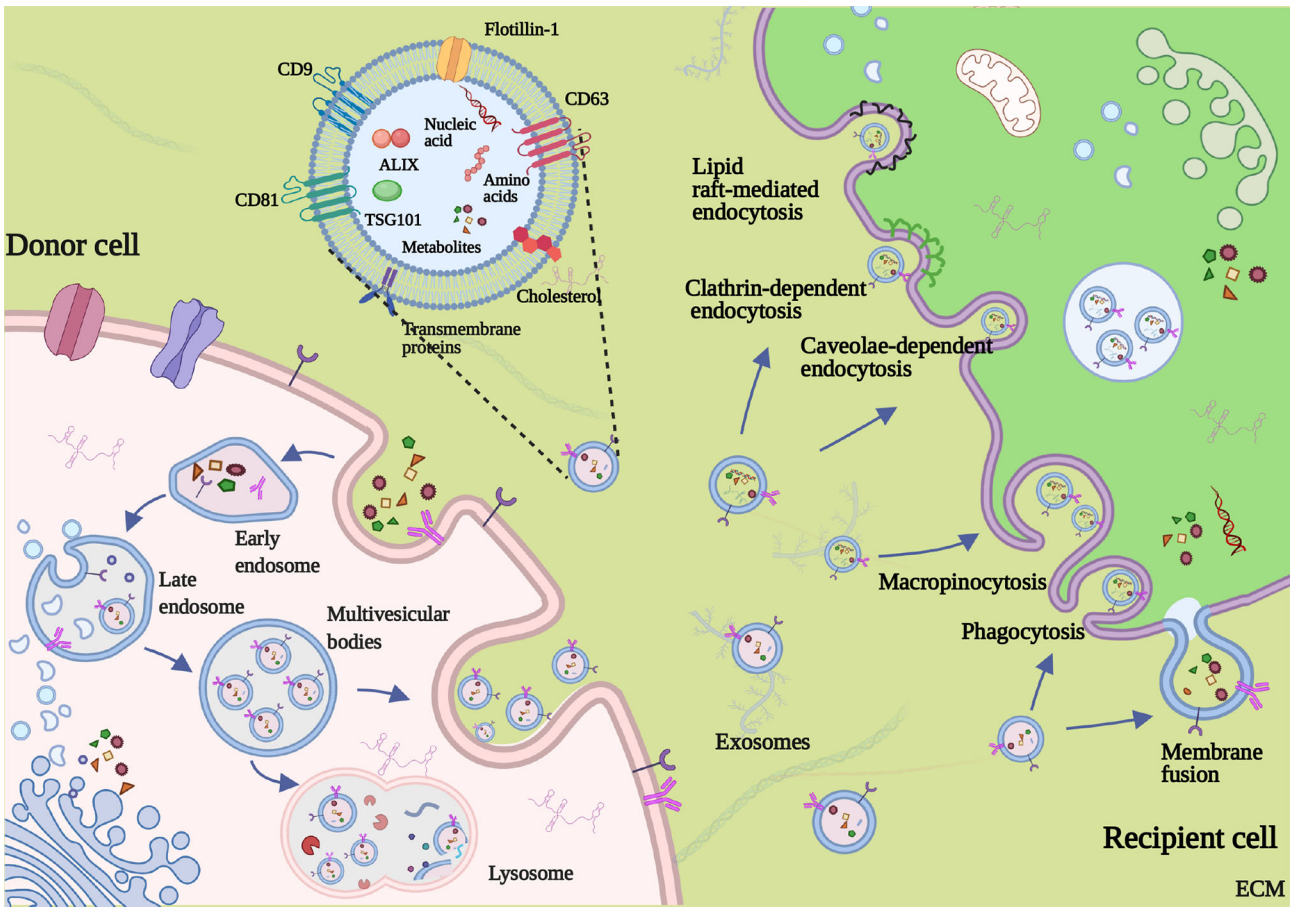


FIGURE 2 | Formation, secretion, and uptake of exosome. Invagination of the plasma membrane forms the early endosome. Then, inward invagination of endosomes gives birth to the formation of multivesicular bodies (MVBs) containing intraluminal vesicles (ILVs). Exosomes are eventually released by fusing of MVBs to the plasma membrane and the exocytosis of ILVs. The mechanism of exosome uptake includes direct fusing with the plasma membrane, macropinocytosis, phagocytosis, caveolin-mediated, lipid raft-mediated endocytosis, and clathrin-dependent endocytosis.

receptors has been reported to exist in mediating immunomodulatory. For example, Tkach et al. showed that exosomes secreted by dendritic cells (DCs) could carry MHC–peptide complexes and bind Toll-like receptor ligands on the bacterial surface, which induced the activation of bystander DCs and T lymphocytes (43, 44). Besides, the families of SNAREs and Rab proteins were reported to mediate the fusion with the plasma membrane and release exosomal cargos (45). Moreover, as representative of internalization, the mode of clathrin-mediated endocytosis has been demonstrated in multiple cell types, such as gastric epithelial cells, colon tumor cells, and cardiomyocytes (46–48). Different exosomal uptake modes may be attributed to the properties of the exosome that shuttles cargos and the metabolic status of recipient cells, but the precise regulating mechanism deserves additional in-depth exploration.

ROLES OF EXOSOME IN THE INITIATION AND DEVELOPMENT OF ESCC

Exosome-related research has focused on the exosome's ability to efficiently transfer an array of selected cargos to recipient cells (49). Studies about the function of exosomes in malignant tumors have developed substantially compared with studies in other fields, and increasing evidence supports exosome-mediated intercellular crosstalk in the TME (50, 51). Accumulating evidence has revealed that exosomes are involved in many features of malignant tumors, including neoplasia, progression, metastasis, angiogenesis, and drug resistance (52–55). In recent years, research involving the correlation between exosomes and ESCC has increased rapidly and yielded valuable information about the function of exosomes in ESCC progression. Here, we summarize the biological function of exosomes that shuttle cargos in the initiation and development of ESCC (**Figure 3**), as shown in **Table 1**. Understanding the function of exosomes and how to use exosomes in ESCC cells to transfer nanoparticles in cell–cell communication are topics at the forefront of oncobiology and may open new avenues for ESCC treatment.

Tumor Microenvironment (TME)

The occurrence of ESCC is a result of a continuous accumulation of mutations in esophageal cells and oncogenic alteration in the TME (75). The TME involves blood vessels, the extracellular matrix (ECM), cytokines, and stromal cells and is indispensable

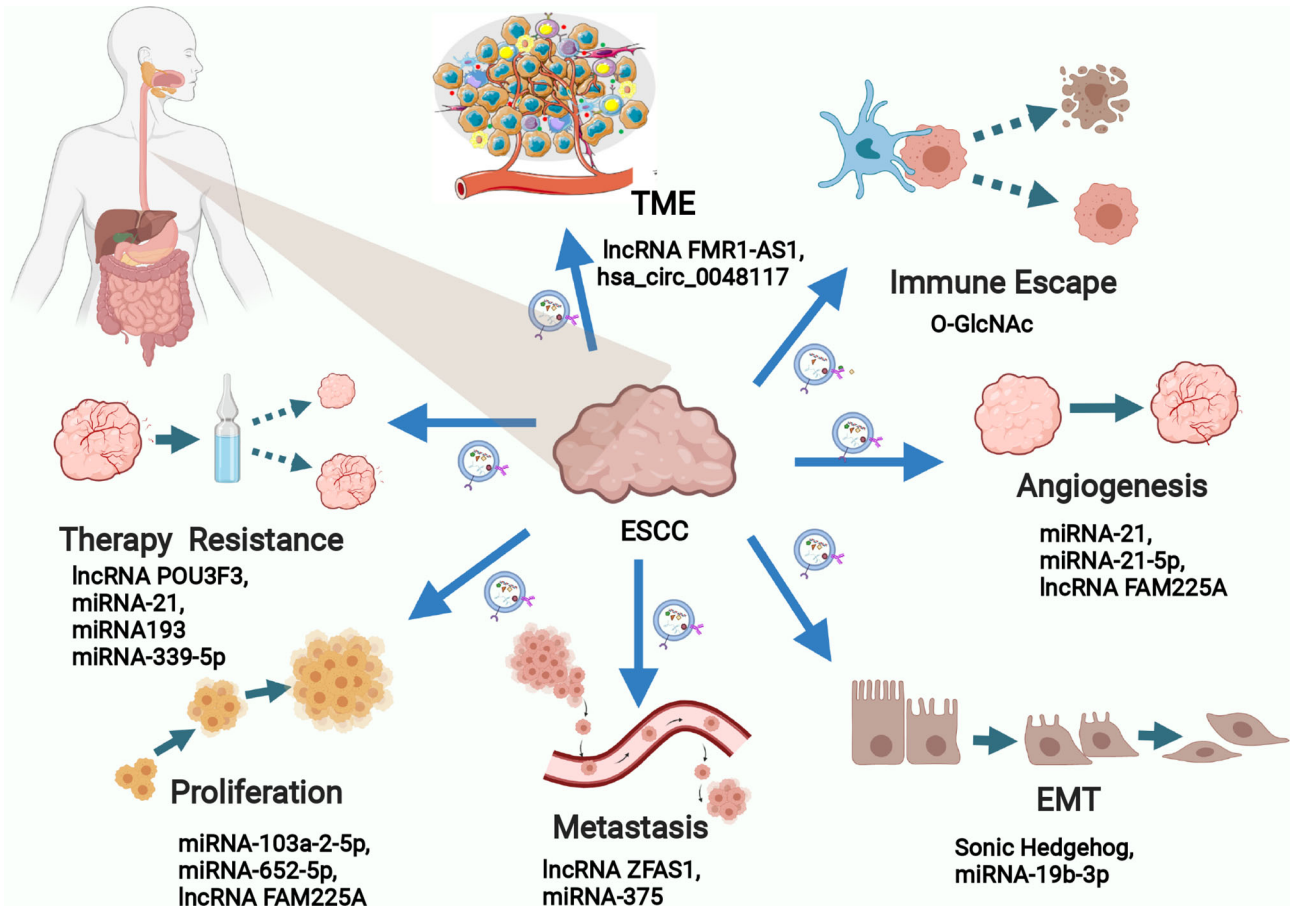


FIGURE 3 | The biological function of exosomal cargos in the development of ESCC.

TABLE 1 | Roles of exosome in the initiation and development of ESCC.

Type	Molecule	Function	Signaling/Target	Ref.
miRNA	miRNA-19b-3p	Reduce apoptosis rate, and promote migration and invasion	PTEN	(56)
	miRNA-103a-2-5p	Promote proliferation and migration	CDH11 gene and NR3C1 gene	(57)
	miRNA-652-5p	Inhibit proliferation and metastasis	PARG and VEGF pathways	(58)
	miRNA-21-5p	Promote angiogenesis	PDCD4 and PTEN/Akt pathway	(59)
	miRNA-21	Promote angiogenesis	SPRY1	(60)
		Promote chemotherapy resistance	PDCD4	(61)
	miRNA-375	Promote apoptosis, and inhibit proliferation, invasion, migration	ENAH	(62)
	miRNA-193	Promote chemotherapy resistance	VEGF and Jak-STAT pathways	(63)
	miRNA-339-5p	Enhance radiosensitivity	Cdc25A	(64)
lncRNA	FMR1-AS1	Maintain TME, promote proliferation, invasion, and inhibit apoptosis	TLR7/NFκB/c-Myc pathway	(65)
	ZFAS1	Promote proliferation, migration, invasion, and inhibit apoptosis	miRNA-124/STAT3 axis	(66)
	FAM225A	Promote apoptosis, and inhibit proliferation, migration, and invasion	miRNA-206/NETO2/FOXP1	(67)
	UCA1	Inhibit proliferation, invasion and migration	miRNA-613	(68)
	PART1	Promote chemotherapy resistance	miRNA-129/Bcl-2 pathway	(69)
	POU3F3	Promote proliferation and chemotherapy resistance	IL-6	(70)
circRNA	has_circ_0048117	Promote invasion and migration	miRNA-140/M2 macrophage	(71)
Others	Sonic Hedgehog	Promote proliferation and migration	Hedgehog pathway	(72)
	O-GlcNAc transferase	Promote the immune escape	PD-1	(73)
	HMGB1	Promote the immune escape	PD-1 positive TAMs	(74)

in tumorigenesis because it provides necessary conditions for tumor growth and manipulates the interaction between cancer cells and their surroundings (76). Exosomes act as an essential role in the formation and reprogramming of the TME, as has been widely documented in many cancer types (77–79). In ESCC, Li et al. found that exosomal FMR1 antisense RNA 1 (FMR1-AS1) could remodel the TME in ESCC (65). The study confirmed that the FMR1-AS1 exosomes were secreted from cancer stem-like cells (CSCs) of ESCC, which transferred stemness phenotypes to recipient non-CSCs in the TME through the mechanism of activating toll-like receptor 7-nuclear factor κ B (TLR7-NFκB) signaling, upregulating the c-Myc level in recipient cells. Another example of exosomal cargos promoting the TME formation involves circ-0048117-rich exosomes derived from hypoxic ESCC cells, promoting M2 macrophage polarization to alter the components in the microenvironment (71). Researchers indicated that hypoxic exosomes modulated the TME in ESCC *via* the transformation of endothelial cell phenotypes and transcriptomes, which enhance angiogenesis and metastasis. Moreover, the exchange of exosomes in cancer cells and the stroma is bidirectional, and cancer-associated fibroblasts (CAFs) can also secrete exosomal Sonic Hedgehog to promote the generation of TME in ESCC (72). These findings suggest that exosomes can play an essential role in the formation, remolding, and normal function of the TME and that novel therapies targeting the TME may be a new approach to cancer treatment.

Proliferation and Apoptosis

The progression of ESCC results from rapid growth and expansion of cancer cells, which may incur tumor survival and defiance to therapy. Exosomes can influence the growth of ESCC by mediating the apoptosis, cycle, and proliferation rate of ESCC cells (80). Molecular profiling has indicated that exosomal miRNA-19b-3p from EC9706 cells targets *PTEN*, a well-known tumor suppressor gene, to regulate the apoptosis of ESCC (56). A similar study elucidated that exosomal lncRNA ZNFX1 antisense RNA 1 (ZFAS1) derived from EC109 cells regulates ESCC proliferation, apoptosis, and migration *via* targeting the miRNA-124/STAT3 signaling pathway (66). Cancer cell-derived exosomes can regulate the ratio of G1-phase cells and influence the cycle and migration ability of ESCC cells (81). Some researchers have demonstrated that the proliferation and apoptosis of ESCC cells are modulated by several other exosomal cargos, including miRNA-103a-2-5p, miRNA-652-5p, lncRNA Family with sequence similarity 225 member A (FAM225A), and lncRNA urothelial cancer-associated 1 (UCA1) (57, 58, 67, 68). Generally, these exosomal contents can work by mediating the expression of proliferation- or apoptosis-related proteins and triggering a subsequent signaling pathway. Importantly, these studies also indicated that the identified exosomal RNAs and other exosomal contents can facilitate the proliferation ability of tumor cells alone and may concurrently alter the potential for migration, angiogenesis, and metastasis. Above all, the activation of cell proliferation cannot entirely be attributed to the expansion and development of ESCC; instead, it results from several steps, including migration and metastasis, angiogenesis, immune response, and therapy resistance. The role of exosomes in these steps of tumor development is explored in the following sections.

Angiogenesis

Angiogenesis, a critical phase during neoplasia, migration, and metastasis, is a multistep formation of neovascularization through which cancer cells obtain sufficient oxygen, nutrition, and energy (82). Reports have illustrated that some exosomes could play a role in inducing angiogenesis in many cancer types (83). Exosomes can deliver numerous pro-angiogenic bioactive substances, including vascular endothelial growth factor (VEGF), miRNAs, or other bioactive mediators. Published data suggest that exosomal cargos accelerate angiogenesis by suppressing the expression of anti-angiogenesis genes and promoting the

expression of pro-angiogenic genes (84, 85). For example, compared with normal exosomes, hypoxic exosomes have played a unique role in facilitating the aggressive behavior of human umbilical vein endothelial cells (HUVECs) both *in vitro* and *in vivo*, and HUVECs exposed to hypoxic exosomes induce enhanced proliferation, metastatic dissemination, and vessel formation ability in ESCC (86). Consistent with that study, Zhuang showed that ESCC cell-derived exosomal miRNA-21 potentiates the angiogenesis ability of HUVEC by targeting sprouty RTK signaling antagonist 1 (SPRY1) in ESCC (60). Moreover, exosomal miRNA-21-5p has been shown to significantly promote the angiogenesis of targeted cells *via* the activation of programmed cell death 4 and downgrading of the signaling pathway or the PTEN/Akt signal pathway in ESCC (59). Similar findings have suggested that exosomal lncRNA FAM225A accelerates ESCC angiogenesis by binding to miRNA-206 and promoting NETO2 and FOXP1 expression (67). Given the pivotal role of angiogenesis in ESCC development and progression, exosome-related research may provide a new avenue to counteract these mechanisms of progression in ESCC, and these discoveries will become even more promising if they are linked to antitumor vascular drugs.

Epithelial-Mesenchymal Transition and Metastasis

Metastasis is a critical step in tumor growth, and it remains a paramount threshold for cancer treatment and the chief cause of cancer mortality (87). Metastasis is a complicated and intricate process involving several steps such as epithelial-mesenchymal transition (EMT) of cancer cells, migration and infiltration into surrounding tissues, intravascular transport, and recognition and establishment in distant tissues (88, 89). Intercellular communication by delivering exosomes from primary tumor cells to the local microenvironment or distant organs is crucial for the phenotypic change and biological aggressive behavior of cancer cells, forming a pre-metastatic niche, and attachment and implantation to distant organs (90). Esophageal cancer cell-derived exosomes can modulate gene expression of recipient cancer cells, leading to an increased risk of invasion and metastasis. For example, the Sonic Hedgehog (SHH) signaling pathway, which has played important roles during development and in cancer (91, 92), can drive tumorigenesis and progression of ESCC. One study showed that exosomal Sonic Hedgehog derived from cancer-associated fibroblasts (CAFs) could increase the activation of N-cadherin and Vimentin in EC109 cell lines and consequently promoted the growth and migration abilities of ESCC (72). Similarly, exosomes derived by infiltrating T cells from irradiated esophageal carcinoma can incur the EMT in ESCC and facilitate metastasis (93). Conversely, human umbilical cord mesenchymal stem cells can suppress enabled homolog (ENAH) expression and decrease the invasion and migration ability of ESCC *via* the exosomal delivery of miRNA-375 (62). Exosomes play a pivotal role in forming a premetastatic niche in distant organs in ESCC, like in gastric cancer and breast cancer (94). This role may be attributed partly to the predisposition of early lymphatic metastasis in ESCC; metastatic dissemination in the liver or lung is relatively rare (95). In conclusion, exosomal cargos can exert pro-tumorigenic effects in most steps of ESCC metastasis, thus promoting the metastatic potential of ESCC. The studies discussed here may offer new insights that help the researcher understand the function of exosomes in metastatic ESCC and uncover exosome-based therapies that may curb cancer metastasis.

Immune Response and Therapy Resistance

The elimination of tumor cells relies heavily on the immune system *in vivo* and exogenous therapies, such as drugs or irradiation (96). The immune system is an intricate network that can guard the body by monitoring, recognizing, and eliminating foreign invaders, such as bacteria, parasites, and endogenous antigens like cancer cells (97). These days, when we talk about the relationship between immune response and tumors, programmed cell death protein 1 (PD-1) is one of the monumental works that are closely associated with it indeed (98). A recent study suggested that O-linked β-N-acetylglucosamine (O-GlcNAc) transferase from stem cells of ESCC can upregulate PD-1 in CD8+ T cells and promote cancer immunosuppression (73). Furthermore, exosomes isolated from serum, plasma, urine, or other body fluids of patients with ESCC, as well as from ESCC cell lines, can reduce B-cell proliferation and induce an increase in interleukin-10 positive regulatory B cells and a high level of PD-1 regulatory B cells (99). Besides, another research demonstrated that exosomal High Mobility Group Box 1 (HMGB1) obtained from ESCC could successfully trigger clonal expansion of PD1 positive tumor-associated macrophages (TAMs), which thereby created conditions for the development of ESCC (74). These findings above contribute to understanding exosomal functions in the immune response and illustrate exosomes' therapeutic application that promotes antitumor immune responses.

Currently, chemotherapy is regarded as the most effective therapy for ESCC after surgery, and tumor recurrence can be attributed mainly to chemotherapy resistance (100, 101). Tumors can achieve drug resistance in many ways, including *via* information exchange by exosomes (102). A recent study indicated that exosomes carrying lncRNA prostate androgen-regulated transcript 1 (PART1) derived from Gefitinib-resistant cells confer cisplatin resistance in ESCC (69). Furthermore, several studies have elucidated that many other exosome-shuttled cargos, such as lncRNA POU class 3 homeobox 3 (POU3F3), miRNA-21, and miRNA-193, are involved in Cisplatin resistance in ESCC (61, 63, 70). Apart from roles in chemoresistance, exosomes reportedly regulate radiation therapy and induce radiation-induced bystander effect (RIBE) (103). Exosomal miRNA-339-5p can mediate the radiosensitivity of ESCC by downregulating cell division cycle 25A (Cdc25A) and can predict outcomes in preoperative radiotherapy (64). These discoveries capture the role of exosomes in therapy resistance and shed light on how engineered exosomes may deliver therapeutic agents for ESCC treatment in the future.

CLINICAL APPLICATION OF EXOSOME IN ESCC

Diagnostic Potential of Exosomal Cargos as Biomarkers for ESCC

ESCC is considered silent cancer because it lacks characteristic manifestations in the early stage. Patients are frequently diagnosed in middle or late stages, delaying the optimal time for treatment and causing a high mortality rate (104). Therefore, it is paramount to find early diagnostic methods to identify patients with ESCC to benefit from early interventions (105). Currently, the gold standard for the diagnosis of ESCC is tissue biopsy under endoscopy, an invasive inspection with correspondingly high costs (106). A non-invasive diagnostic method in early-stage ESCC is urgently needed. Researchers have made great efforts to screen for ESCC biomarkers; possible candidates include circulating tumor cells (CTCs), serum miRNAs, small extracellular vesicles (sEVs), as well as circulating tumor DNA (ctDNA) (107, 108). Exosomal cargos serve as promising tumor biomarkers because they reflect the donor cell and their presence in various biological fluids (109–111). New research into the potential application of exosomes as tumor biomarkers has emerged and yielded valuable information for additional in-depth exploration (112, 113). Growing evidence has confirmed that exosomal RNAs outperform peripheral blood-free RNAs in cancer diagnosis because of several advantages: First, exosomes exist in all biological fluids and are easily accessible compared with plasma. Second, exosomal cargos can be well protected from degradation by enzymes or elimination by the biological barrier. Third, the components of exosomes have high homology with donor cells, which may encourage a higher specificity of exosome-based detection. Last, the concentration of exosomal cargos is higher than the expression of plasma RNAs (114–116). Here, we focus on state-of-the-art exosomal cargos in ESCC.

The distinct expression of exosome-shuttling contents between cancer cells and normal cells supports the application of exosomes as biomarkers for ESCC. Among these exosomal compounds, exosome-carrying miRNAs are most investigated. For example, Zeng corroborated that exosome-shuttled miRNA-19b-3p separated from patients with ESCC is significantly upregulated compared with healthy controls, suggesting that serum exosome-encapsulated miRNA-19b-3p highlights the potential utility of exosomal RNAs in the early detection of ESCC (56). A study from 51 patients with ESCC and 41 with benign diseases showed that plasma exosomal miRNA-21 levels were significantly elevated in ESCC versus benign diseases so that they were suitable to be biomarkers for early diagnosis of ESCC (117). Except for the effect of exosomes on the differential diagnosis of ESCC, lymph node metastasis and TNM grade have been associated with the expression of some exosomal cargos that may serve as independent prognostic factors of ESCC. Lu et al. suggested that tumor cells-derived exosomes markedly upregulated the expression of hsa-circ-0048117 under the condition of hypoxia—a change that may be positively correlated with advanced T and N stages serves as a biomarker for progression (71). Similar studies have revealed that downregulated exosomal miRNA-339-5p and miRNA-652-5p in the serum are related to advanced TNM stages and a higher lymph node metastasis rate (58, 64). Moreover, several studies have indicated that higher serum exosomal miRNA-182, miRNA-766-3p, lncRNA POU3F3, and has-circ-0026611 levels in patients with ESCC are positively related to poor prognosis (70, 118–120). Except for exosomal RNAs, a small amount of exosome-carrying protein has been reported previously. The over-expression of Stathmin-1, regarded as microtubule depolymerization protein, is related to the process of tumor spread, adverse clinical outcomes, and chemoresistance in many types of cancer, especially squamous cell carcinoma, by controlling cell division, proliferation, and migration (121–124). In ESCC, Yan et al. corroborated that the average expression of stathmin-1 elevated in oncogenic exosomes, and the serum stathmin-1 level in patients with ESCC was obviously higher than that of healthy individuals (125). In addition, elevated concentration of stathmin-1 was related to lymphatic metastasis and late staged cancer.

Thus, several exosomal cargos have been identified as biomarkers for ESCC in applying possible diagnosis and potential prognosis, as described in **Table 2**. It is noteworthy that few literature reports address the diagnostic value of these molecules; the diagnostic efficacy of most candidates, apart from few miRNAs like miRNA-21, deserve additional validation. Much work remains before exosomal cargos can be applied as ideal biomarkers of ESCC. In addition, readers should note that some of the articles we referenced were published before the MISEV2018 guidelines were issued, which means the research methods they applied might not meet the standard of the guidelines. For example, few works failed to further validate sEV-specific markers, which is not recognized by the standards of EV isolation protocols in MISEV2018. Therefore, we hope readers accommodate the term "small extracellular vesicle" to replace "exosome",

Potential Application of Exosomes in Treatment of ESCC

In addition to having a diagnostic role in cancer, exosomes have potential use in disease therapy. The characteristic property in transferring selected payloads to recipient cells has translated into potential applications for treating many diseases, including cancer and cardiovascular diseases (134–136). However, recent articles published on the potential application of exosomes in the treatment of ESCC are sparse. Thus, this section mainly summarizes exosome-based therapy in other cancers and introduces the exosomes that may be potential targets for ESCC therapy in the future. Researchers harness engineered exosomes to deliver chemotherapeutic agents, achieving better performance than traditional vectors like liposomes (49). Theoretically, exosomes have the following advantages: First, the membrane structure of an exosome can protect pharmacological agents from degradation. Second, exosomes naturally exist in all biological fluids, and thus they can be well tolerated when introduced into the body. They can efficiently

TABLE 2 | Exosomal cargos as biomarkers for ESCC.

Type	Molecules	Origin	Potential Functions	Ref.
miRNA	miRNA -19b-3p	Serum	Distinguish ESCC patients from healthy individuals	(56)
	miRNA -21	Serum	Predict TNM stage	(117)
	miRNA -652-5p	Serum	Predict TNM stage, lymph node metastasis, and survival rate	(58)
	miRNA -339-5p	Serum	Predict radiotherapy sensitivity and survival rate	(64)
	miRNA -182	Serum	Distinguish ESCC patients from healthy individuals, predict TNM stage and survival rate	(118)
	miRNA -766-3p	Serum	Predict TNM stage and survival rate	(119)
	miRNA -103a-2-5p	Serum	Predict survival rate	(57)
	miRNA -93-5p	Serum	Predict survival rate	(126)
	chr 8-23234-3p, chr 1-17695-5p, chr 8-2743-5p, miRNA-432-5p	Serum	Predict lymph node metastasis	(127)
lncRNA	POU3F3	Serum	Predict chemotherapy sensitivity and survival rate	(70)
	RP5-1092A11.2	Serum	Distinguish ESCC patients from esophagitis patients and from healthy individuals	(128)
	NR_039819 NR_036133 NR_003353 ENST00000442416.1 ENST00000416100.1	Serum	Distinguish ESCC patients from healthy individuals	(129)
	UCA1	Serum	Early diagnosis	(68)
	FMR1-AS1	Serum	Predict survival rate, especially in female ESCC patients	(65)
	POU3F3	Serum	Predict survival rate	(130)
circRNA	hsa-circ-0048117	Serum	Predict TNM stage	(71)
	hsa_circ_0026611	Serum	Predicted lymph node metastasis and survival rate	(120)
	hsa_circ_0001946 hsa_circ_0001946	Serum	Predict recurrence and survival rate	(131)
Other biomarkers	G-NchiRNA	Salivary	Reflect tumor burden, evaluate therapeutic response and predict survival rate	(132)
	uc.189	Serum	Evaluate lymph node metastasis	(133)
	Stathmin-1	Serum	Differentiate patients with ESCC from healthy individuals, and be associated with lymph node metastasis and advanced cancer stage	(125)

penetrate biological barriers and deliver targeted cargo with minimal immune clearance. Third, some exosomes have receptor-targeting features resulting from the heterogeneity of exosomal surface proteins, enabling targeted delivery of therapeutic agents for cancer. Last, because they are shed by all cells as part of their normal physiology, exosomes may induce less toxicity and minimize other adverse reactions even with repeated injection (137–140). Therefore, exosomes may have a bright future as nanocarriers for cancer treatment.

Currently, researchers are committed to designing exosomes to encapsulate therapeutic agents and conducting studies that yield valuable information about the application of exosomes for the administration of diseases. For example, gene-engineered exosome-thermosensitive liposomes can block the CD47 immune checkpoint and improve the macrophage-mediated elimination of cancer cells (30). Pan et al. indicated that urinary exosome-based engineered nanovectors could help deliver targeted homologous treatment in prostate cancer and may exemplify a novel, efficient, and facile therapy strategy (29). So far, evidence about the function of engineered exosomes in cancer therapy is primarily from cancers like gastric cancer or prostate cancer, not from ESCC; relevant data in ESCC are, at best, sparse. However, despite the many unanswered questions about their clinical application, exosomes show great potential to facilitate ESCC therapy. In addition, accumulating research has suggested that engineered exosomes *in vitro* can play a paramount role in different experimental settings, but more exploration is needed before these findings translate into clinical practice (141). Also, research must guarantee the homogeneity of exosomes by standardizing protocols for exosome isolation, preparation, and route of administration (8). Currently, inefficient isolation methods of exosomes cannot provide sufficient exosomes to meet cancer therapy requirements (142). Identifying ways to prevent exosomes from being taken up by other cells and to drive the engineered vehicles to targeted cells or organs remains a considerable challenge (143–145). Overall, engineered exosomes provide a promising therapeutic option for cancer treatment, although the utility of this strategy in clinical practice requires additional exploration.

CONCLUSIONS AND PERSPECTIVES

As referred to above, substantial evidence has delineated that exosomes and their inclusion, such as DNA, RNA, proteins, lipids, and other biological complexes, significantly affect cellular pathways and mediate pathophysiology behaviors, involving cell growth oncogenesis and tumor differentiation. EVs may act as biomarkers for early ESCC diagnosis, therapeutic monitoring, or prognosis evaluation. The association with exosomal cargos and disease state could be used in diagnostic and prognostic biomarkers for early ESCC, such as miRNA-21 was recognized

as an exosome-derived small RNA superior to traditional tumor markers for early diagnosis of ESCC (117). It should be noted that a panel of miRNAs has delineated a higher sensitivity and specificity compared with a single miRNA, but researchers have to make a considerable effort to screen for miRNA panels that can be used in early diagnosis for ESCC (129, 130). In addition, there are some limitations in the clinical application of circulating exosome-related analysis. The separation and purification technology is mainly used in scientific research but rarely applied in clinical practice. Exosome-testing kits are developing, and some laboratory investigations and clinical studies on ESCC exosomes are ongoing or are forthcoming. For example, a newly developed commercial kit (ExoLutE®) utilizing the principle of size-exclusion spun column improves the efficiency and purity of circulating exosome separation compared to conventional kits (146).

Exosomes have been widely regarded as a promising carrier of anticancer drugs, which were proved in animal studies. If engineering exosomes that carried anticancer agents could be realized, the delivery of therapeutic agents by exosomes would make exosomes ideal vectors for cancer therapy (147). To date, evidence on the function of engineered exosomes in cancer therapy has mainly come from other cancers, such as pancreatic or prostate cancer (29, 148), and data on ESCC have been sparse at best. With the development of science and technology, large-scale clinical studies or trials on ESCC exosomes will undoubtedly be carried out. It is believed that more and more achievements on exosomes will be made and applied in ESCC clinical diagnosis and therapy soon.

AUTHOR CONTRIBUTIONS

Literature review and writing—original draft preparation: ZZ and SY. Writing—review and editing: AZ, XL, RF, and SZ. Supervision and funding acquisition: PL and GZ. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded by the National Natural Science Foundation of China (no. 82070575), Beijing Natural Science Foundation (J180010), and Beijing Municipal Science & Technology Commission (Z191100006619080).

ACKNOWLEDGMENTS

Many important contributions could not be cited due to space constraints, and we apologize to our colleagues for any relevant exclusion. The figures in this manuscript were created with BioRender.com.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
2. Thrift AP. Global Burden and Epidemiology of Barrett Oesophagus and Oesophageal Cancer. *Nat Rev Gastroenterol Hepatol* (2021) 18(6):432–43. doi: 10.1038/s41575-021-00419-3
3. Yang S, Lin S, Li N, Deng Y, Wang M, Xiang D, et al. Burden, Trends, and Risk Factors of Esophageal Cancer in China From 1990 to 2017: An Up-to-Date Overview and Comparison With Those in Japan and South Korea. *J Hematol Oncol* (2020) 13(1):146. doi: 10.1186/s13045-020-00981-4
4. Pakzad R, Mohammadian-Hafshejani A, Khosravi B, Soltani S, Pakzad I, Mohammadian M, et al. The Incidence and Mortality of Esophageal Cancer and Their Relationship to Development in Asia. *Ann Transl Med* (2016) 4 (2):29. doi: 10.3978/j.issn.2305-5839.2016.01.11
5. Qin Q, Ge X, Wang X, Wang L, Li C, Chen J, et al. Stage III Esophageal Squamous Cell Carcinoma Patients With Three-Dimensional Conformal or Intensity-Modulated Radiotherapy: A Multicenter Retrospective Study. *Front Oncol* (2020) 10:580450. doi: 10.3389/fonc.2020.580450
6. Stein JM, Luzio JP. Ectocytosis Caused by Sublytic Autologous Complement Attack on Human Neutrophils. The Sorting of Endogenous Plasma-Membrane Proteins and Lipids Into Shed Vesicles. *Biochem J* (1991) 274 (Pt 2):381–6. doi: 10.1042/bj2740381
7. Witwer KW, Soekmadji C, Hill AF, Wauben MH, Buzas EI, Di Vizio D, et al. Updating the MISEV Minimal Requirements for Extracellular Vesicle Studies: Building Bridges to Reproducibility. *J Extracell Vesicles* (2017) 6 (1):1396823. doi: 10.1080/20013078.2017.1396823
8. Thery C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, et al. Minimal Information for Studies of Extracellular Vesicles 2018 (MISEV2018): A Position Statement of the International Society for Extracellular Vesicles and Update of the MISEV2014 Guidelines. *J Extracell Vesicles* (2018) 7(1):1535750. doi: 10.1080/20013078.2018.1535750
9. Witwer KW, Thery C. Extracellular Vesicles or Exosomes? On Primacy, Precision, and Popularity Influencing a Choice of Nomenclature. *J Extracell Vesicles* (2019) 8(1):1648167. doi: 10.1080/20013078.2019.1648167
10. Pegtel DM, Gould SJ. Exosomes. *Annu Rev Biochem* (2019) 88:487–514. doi: 10.1146/annurev-biochem-013118-111902
11. Kalluri R, LeBleu VS. The Biology, Function, and Biomedical Applications of Exosomes. *Science* (2020) 367(6478):eaau6977. doi: 10.1126/science.aau6977
12. Sullivan R, Saez F, Girouard J, Frenette G. Role of Exosomes in Sperm Maturation During the Transit Along the Male Reproductive Tract. *Blood Cells Mol Dis* (2005) 35(1):1–10. doi: 10.1016/j.bcmd.2005.03.005
13. Xiao J, Pan Y, Li XH, Yang XY, Feng YL, Tan HH, et al. Cardiac Progenitor Cell-Derived Exosomes Prevent Cardiomyocytes Apoptosis Through Exosomal miR-21 by Targeting PDCD4. *Cell Death Dis* (2016) 7(6):e2277. doi: 10.1038/cddis.2016.181
14. Rajagopal C, Harikumar KB. The Origin and Functions of Exosomes in Cancer. *Front Oncol* (2018) 8:66. doi: 10.3389/fonc.2018.00066
15. Chennakrishnaiah S, Meehan B, D'Asti E, Montermini L, Lee TH, Karatzas N, et al. Leukocytes as a Reservoir of Circulating Oncogenic DNA and Regulatory Targets of Tumor-Derived Extracellular Vesicles. *J Thromb Haemost* (2018) 16(9):1800–13. doi: 10.1111/jth.14222
16. Kurywchak P, Tavormina J, Kalluri R. The Emerging Roles of Exosomes in the Modulation of Immune Responses in Cancer. *Genome Med* (2018) 10 (1):23. doi: 10.1186/s13073-018-0535-4
17. Babaei M, Rezaie J. Application of Stem Cell-Derived Exosomes in Ischemic Diseases: Opportunity and Limitations. *J Transl Med* (2021) 19(1):196. doi: 10.1186/s12967-021-02863-w
18. Xavier CPR, Caires HR, Barbosa MAG, Bergantim R, Guimaraes JE, Vasconcelos MH. The Role of Extracellular Vesicles in the Hallmarks of Cancer and Drug Resistance. *Cells* (2020) 9(5):1141. doi: 10.3390/cells9051141
19. Barile L, Vassalli G. Exosomes: Therapy Delivery Tools and Biomarkers of Diseases. *Pharmacol Ther* (2017) 174:63–78. doi: 10.1016/j.pharmthera.2017.02.020

20. Wang M, Ji S, Shao G, Zhang J, Zhao K, Wang Z, et al. Effect of Exosome Biomarkers for Diagnosis and Prognosis of Breast Cancer Patients. *Clin Transl Oncol* (2018) 20(7):906–11. doi: 10.1007/s12094-017-1805-0
21. Wang H, Lu Z, Zhao X. Tumorigenesis, Diagnosis, and Therapeutic Potential of Exosomes in Liver Cancer. *J Hematol Oncol* (2019) 12(1):133. doi: 10.1186/s13045-019-0806-6
22. Soung YH, Ford S, Zhang V, Chung J. Exosomes in Cancer Diagnostics. *Cancers (Basel)* (2017) 9(1):8. doi: 10.3390/cancers9010008
23. Huang T, Deng CX. Current Progresses of Exosomes as Cancer Diagnostic and Prognostic Biomarkers. *Int J Biol Sci* (2019) 15(1):1–11. doi: 10.7150/ijbs.27796
24. Katakowski M, Buller B, Zheng X, Lu Y, Rogers T, Osobamiro O, et al. Exosomes From Marrow Stromal Cells Expressing miR-146b Inhibit Glioma Growth. *Cancer Lett* (2013) 335(1):201–4. doi: 10.1016/j.canlet.2013.02.019
25. Gehrmann U, Naslund TI, Hiltbrunner S, Larssen P, Gabrielsson S. Harnessing the Exosome-Induced Immune Response for Cancer Immunotherapy. *Semin Cancer Biol* (2014) 28:58–67. doi: 10.1016/j.semcancer.2014.05.003
26. Samanta S, Rajasingh S, Drosos N, Zhou Z, Dawn B, Rajasingh J. Exosomes: New Molecular Targets of Diseases. *Acta Pharmacol Sin* (2018) 39(4):501–13. doi: 10.1038/aps.2017.162
27. Wu P, Zhang B, Ocansey DKW, Xu W, Qian H. Extracellular Vesicles: A Bright Star of Nanomedicine. *Biomaterials* (2021) 269:120467. doi: 10.1016/j.biomaterials.2020.120467
28. You DG, Lim GT, Kwon S, Um W, Oh BH, Song SH, et al. Metabolically Engineered Stem Cell-Derived Exosomes to Regulate Macrophage Heterogeneity in Rheumatoid Arthritis. *Sci Adv* (2021) 7(23):eabe0083. doi: 10.1126/sciadv.abe0083
29. Pan S, Zhang Y, Huang M, Deng Z, Zhang A, Pei L, et al. Urinary Exosomes-Based Engineered Nanovectors for Homologously Targeted Chemo-Chemodynamic Prostate Cancer Therapy *via* Abrogating EGFR/AKT/NF-Kb/IkB Signaling. *Biomaterials* (2021) 275:120946. doi: 10.1016/j.biomaterials.2021.120946
30. Cheng L, Zhang X, Tang J, Lv Q, Liu J. Gene-Engineered Exosomes-Thermosensitive Liposomes Hybrid Nanovesicles by the Blockade of CD47 Signal for Combined Photothermal Therapy and Cancer Immunotherapy. *Biomaterials* (2021) 275:120964. doi: 10.1016/j.biomaterials.2021.120964
31. Kalluri R. The Biology and Function of Exosomes in Cancer. *J Clin Invest* (2016) 126(4):1208–15. doi: 10.1172/JCI81135
32. van Niel G, D'Angelo G, Raposo G. Shedding Light on the Cell Biology of Extracellular Vesicles. *Nat Rev Mol Cell Biol* (2018) 19(4):213–28. doi: 10.1038/nrm.2017.125
33. McAndrews KM, Kalluri R. Mechanisms Associated With Biogenesis of Exosomes in Cancer. *Mol Cancer* (2019) 18(1):52. doi: 10.1186/s12943-019-0963-9
34. Mathieu M, Martin-Jaular L, Lavieu G, Thery C. Specificities of Secretion and Uptake of Exosomes and Other Extracellular Vesicles for Cell-to-Cell Communication. *Nat Cell Biol* (2019) 21(1):9–17. doi: 10.1038/s41556-018-0250-9
35. Willms E, Cabanas C, Mager I, Wood MJA, Vader P. Extracellular Vesicle Heterogeneity: Subpopulations, Isolation Techniques, and Diverse Functions in Cancer Progression. *Front Immunol* (2018) 9:738. doi: 10.3389/fimmu.2018.00738
36. Hessvik NP, Llorente A. Current Knowledge on Exosome Biogenesis and Release. *Cell Mol Life Sci* (2018) 75(2):193–208. doi: 10.1007/s00018-017-2595-9
37. Kahlert C, Kalluri R. Exosomes in Tumor Microenvironment Influence Cancer Progression and Metastasis. *J Mol Med (Berl)* (2013) 91(4):431–7. doi: 10.1007/s00109-013-1020-6
38. Villarroya-Beltri C, Baixauli F, Gutierrez-Vazquez C, Sanchez-Madrid F, Mittelbrunn M. Sorting It Out: Regulation of Exosome Loading. *Semin Cancer Biol* (2014) 28:3–13. doi: 10.1016/j.semcancer.2014.04.009
39. Spencer N, Yeruva L. Role of Bacterial Infections in Extracellular Vesicles Release and Impact on Immune Response. *BioMed J* (2021) 44(2):157–64. doi: 10.1016/j.bj.2020.05.006
40. Yanez-Mo M, Siljander PR, Andreu Z, Zavec AB, Borras FE, Buzas EI, et al. Biological Properties of Extracellular Vesicles and Their Physiological Functions. *J Extracell Vesicles* (2015) 4:27066. doi: 10.3402/jev.v4.27066
41. Maia J, Caja S, Strano Moraes MC, Couto N, Costa-Silva B. Exosome-Based Cell-Cell Communication in the Tumor Microenvironment. *Front Cell Dev Biol* (2018) 6:18. doi: 10.3389/fcell.2018.00018
42. Gurung S, Perocheau D, Touramanidou L, Baruteau J. The Exosome Journey: From Biogenesis to Uptake and Intracellular Signalling. *Cell Commun Signal* (2021) 19(1):47. doi: 10.1186/s12964-021-00730-1
43. Tkach M, Kowal J, Zucchetti AE, Enserink L, Jouve M, Lankar D, et al. Qualitative Differences in T-Cell Activation by Dendritic Cell-Derived Extracellular Vesicle Subtypes. *EMBO J* (2017) 36(20):3012–28. doi: 10.15252/embj.201696003
44. Sobo-Vujanovic A, Munich S, Vujanovic NL. Dendritic-Cell Exosomes Cross-Present Toll-Like Receptor-Ligands and Activate Bystander Dendritic Cells. *Cell Immunol* (2014) 289(1-2):119–27. doi: 10.1016/j.cellimm.2014.03.016
45. Prada I, Meldolesi J. Binding and Fusion of Extracellular Vesicles to the Plasma Membrane of Their Cell Targets. *Int J Mol Sci* (2016) 17(8):1296. doi: 10.3390/ijms17081296
46. Yoon JH, Ashktorab H, Smoot DT, Nam SW, Hur H, Park WS. Uptake and Tumor-Suppressive Pathways of Exosome-Associated GKN1 Protein in Gastric Epithelial Cells. *Gastric Cancer* (2020) 23(5):848–62. doi: 10.1007/s10120-020-01068-2
47. Horibe S, Tanahashi T, Kawauchi S, Murakami Y, Rikitake Y. Mechanism of Recipient Cell-Dependent Differences in Exosome Uptake. *BMC Cancer* (2018) 18(1):47. doi: 10.1186/s12885-017-3958-1
48. Eguchi S, Takefuji M, Sakaguchi T, Ishihama S, Mori Y, Tsuda T, et al. Cardiomyocytes Capture Stem Cell-Derived, Anti-Apoptotic microRNA-214 *via* Clathrin-Mediated Endocytosis in Acute Myocardial Infarction. *J Biol Chem* (2019) 294(31):11665–74. doi: 10.1074/jbc.RA119.007537
49. Antimisiaris SG, Mourtas S, Marazioti A. Exosomes and Exosome-Inspired Vesicles for Targeted Drug Delivery. *Pharmaceutics* (2018) 10(4):218. doi: 10.3390/pharmaceutics10040218
50. He C, Li L, Wang L, Meng W, Hao Y, Zhu G. Exosome-Mediated Cellular Crosstalk Within the Tumor Microenvironment Upon Irradiation. *Cancer Biol Med* (2021) 18(1):21–33. doi: 10.20892/j.issn.2095-3941.2020.0150
51. Yang E, Wang X, Gong Z, Yu M, Wu H, Zhang D. Exosome-Mediated Metabolic Reprogramming: The Emerging Role in Tumor Microenvironment Remodeling and Its Influence on Cancer Progression. *Signal Transduct Target Ther* (2020) 5(1):242. doi: 10.1038/s41392-020-00359-5
52. Qu L, Ding J, Chen C, Wu ZJ, Liu B, Gao Y, et al. Exosome-Transmitted lncARSR Promotes Sunitinib Resistance in Renal Cancer by Acting as a Competing Endogenous RNA. *Cancer Cell* (2016) 29(5):653–68. doi: 10.1016/j.ccell.2016.03.004
53. Plebanek MP, Angeloni NL, Vinokour E, Li J, Henkin A, Martinez-Marin D, et al. Pre-Metastatic Cancer Exosomes Induce Immune Surveillance by Patrolling Monocytes at the Metastatic Niche. *Nat Commun* (2017) 8(1):1319. doi: 10.1038/s41467-017-01433-3
54. Lee HY, Chen CK, Ho CM, Lee SS, Chang CY, Chen KJ, et al. EIF3C-Enhanced Exosome Secretion Promotes Angiogenesis and Tumorigenesis of Human Hepatocellular Carcinoma. *Oncotarget* (2018) 9(17):13193–205. doi: 10.18632/oncotarget.24149
55. Stefanius K, Servage K, de Souza Santos M, Gray HF, Toombs JE, Chimalapati S, et al. Human Pancreatic Cancer Cell Exosomes, But Not Human Normal Cell Exosomes, Act as an Initiator in Cell Transformation. *Elife* (2019) 8:e40226. doi: 10.7554/eLife.40226
56. Zeng Q, Zhu Z, Song L, He Z. Transferred by Exosomes-Derived MiR-19b-3p Targets PTEN to Regulate Esophageal Cancer Cell Apoptosis, Migration and Invasion. *Biosci Rep* (2020) 40(11):BSR20201858. doi: 10.1042/BSR20201858
57. Gao DC, Hou B, Zhou D, Liu QX, Zhang K, Lu X, et al. Tumor-Derived Exosomal miR-103a-2-5p Facilitates Esophageal Squamous Cell Carcinoma Cell Proliferation and Migration. *Eur Rev Med Pharmacol Sci* (2020) 24(11):6097–110. doi: 10.26355/eurrev_202006_21505
58. Gao P, Wang D, Liu M, Chen S, Yang Z, Zhang J, et al. DNA Methylation-Mediated Repression of Exosomal miR-652-5p Expression Promotes Oesophageal Squamous Cell Carcinoma Aggressiveness by Targeting PARG and VEGF Pathways. *PloS Genet* (2020) 16(4):e1008592. doi: 10.1371/journal.pgen.1008592

59. Liao J, Liu R, Yin L, Pu Y. Expression Profiling of Exosomal miRNAs Derived From Human Esophageal Cancer Cells by Solexa High-Throughput Sequencing. *Int J Mol Sci* (2014) 15(9):15530–51. doi: 10.3390/ijms150915530
60. Zhuang H, Wang H, Yang H, Li H. Exosome-Encapsulated MicroRNA-21 From Esophageal Squamous Cell Carcinoma Cells Enhances Angiogenesis of Human Umbilical Venous Endothelial Cells by Targeting Spry1. *Cancer Manag Res* (2020) 12:10651–67. doi: 10.2147/CMAR.S259077
61. Yang YC, Liu GJ, Yuan DF, Li CQ, Xue M, Chen LJ. Influence of Exosome-Derived miR-21on Chemotherapy Resistance of Esophageal Cancer. *Eur Rev Med Pharmacol Sci* (2019) 23(4):1513–9. doi: 10.26355/eurrev_201902_17109
62. He Z, Li W, Zheng T, Liu D, Zhao S. Human Umbilical Cord Mesenchymal Stem Cells-Derived Exosomes Deliver microRNA-375 to Downregulate ENAH and Thus Retard Esophageal Squamous Cell Carcinoma Progression. *J Exp Clin Cancer Res* (2020) 39(1):140. doi: 10.1186/s13046-020-01631-w
63. Shi S, Huang X, Ma X, Zhu X, Zhang Q. Research of the Mechanism on Mirna193 in Exosomes Promotes Cisplatin Resistance in Esophageal Cancer Cells. *PloS One* (2020) 15(5):e0225290. doi: 10.1371/journal.pone.0225290
64. Luo A, Zhou X, Shi X, Zhao Y, Men Y, Chang X, et al. Exosome-Derived miR-339-5p Mediates Radiosensitivity by Targeting Cdc25A in Locally Advanced Esophageal Squamous Cell Carcinoma. *Oncogene* (2019) 38 (25):4990–5006. doi: 10.1038/s41388-019-0771-0
65. Li W, Zhang L, Guo B, Deng J, Wu S, Li F, et al. Exosomal FMR1-AS1 Facilitates Maintaining Cancer Stem-Like Cell Dynamic Equilibrium *via* TLR7/NFkappaB/c-Myc Signaling in Female Esophageal Carcinoma. *Mol Cancer* (2019) 18(1):22. doi: 10.1186/s12943-019-0949-7
66. Li Z, Qin X, Bian W, Li Y, Shan B, Yao Z, et al. Exosomal lncRNA ZFAS1 Regulates Esophageal Squamous Cell Carcinoma Cell Proliferation, Invasion, Migration and Apoptosis *via* microRNA-124/STAT3 Axis. *J Exp Clin Cancer Res* (2019) 38(1):477. doi: 10.1186/s13046-019-1473-8
67. Zhang C, Luo Y, Cao J, Wang X, Miao Z, Shao G. Exosomal lncRNA FAM225A Accelerates Esophageal Squamous Cell Carcinoma Progression and Angiogenesis *via* Sponging miR-206 to Upregulate NETO2 and FOXP1 Expression. *Cancer Med* (2020) 9(22):8600–11. doi: 10.1002/cam4.3463
68. Zhu Z, Wang H, Pang Y, Hu H, Zhang H, Wang W. Exosomal Long non-Coding RNA UCA1 Functions as Growth Inhibitor in Esophageal Cancer. *Aging (Albany NY)* (2020) 12(20):20523–39. doi: 10.18632/aging.103911
69. Kang M, Ren M, Li Y, Fu Y, Deng M, Li C. Exosome-Mediated Transfer of lncRNA PART1 Induces Gefitinib Resistance in Esophageal Squamous Cell Carcinoma *via* Functioning as a Competing Endogenous RNA. *J Exp Clin Cancer Res* (2018) 37(1):171. doi: 10.1186/s13046-018-0845-9
70. Tong Y, Yang L, Yu C, Zhu W, Zhou X, Xiong Y, et al. Tumor-Secreted Exosomal lncRNA POU3F3 Promotes Cisplatin Resistance in ESCC by Inducing Fibroblast Differentiation Into CAFs. *Mol Ther Oncolytics* (2020) 18:1–13. doi: 10.1016/j.omto.2020.05.014
71. Lu Q, Wang X, Zhu J, Fei X, Chen H, Li C. Hypoxic Tumor-Derived Exosomal Circ0048117 Facilitates M2 Macrophage Polarization Acting as miR-140 Sponge in Esophageal Squamous Cell Carcinoma. *Onco Targets Ther* (2020) 13:11883–97. doi: 10.2147/OTT.S284192
72. Zhao G, Li H, Guo Q, Zhou A, Wang X, Li P, et al. Exosomal Sonic Hedgehog Derived From Cancer-Associated Fibroblasts Promotes Proliferation and Migration of Esophageal Squamous Cell Carcinoma. *Cancer Med* (2020) 9(7):2500–13. doi: 10.1002/cam4.2873
73. Yuan Y, Wang L, Ge D, Tan L, Cao B, Fan H, et al. Exosomal O-GlcNAc Transferase From Esophageal Carcinoma Stem Cell Promotes Cancer Immunosuppression Through Up-Regulation of PD-1 in CD8(+) T Cells. *Cancer Lett* (2021) 500:98–106. doi: 10.1016/j.canlet.2020.12.012
74. Li B, Song TN, Wang FR, Yin C, Li Z, Lin JP, et al. Tumor-Derived Exosomal HMGB1 Promotes Esophageal Squamous Cell Carcinoma Progression Through Inducing PD1(+) TAM Expansion. *Oncogenesis* (2019) 8(3):17. doi: 10.1038/s41389-019-0126-2
75. Takeshima H, Ushijima T. Accumulation of Genetic and Epigenetic Alterations in Normal Cells and Cancer Risk. *NPJ Precis Oncol* (2019) 3:7. doi: 10.1038/s41698-019-0079-0
76. Baghban R, Roshangar L, Jahanban-Esfahlan R, Seidi K, Ebrahimi-Kalan A, Jaymand M, et al. Tumor Microenvironment Complexity and Therapeutic Implications at a Glance. *Cell Commun Signal* (2020) 18(1):59. doi: 10.1186/s12964-020-0530-4
77. Hamidi H, Ivaska J. Every Step of the Way: Integrins in Cancer Progression and Metastasis. *Nat Rev Cancer* (2018) 18(9):533–48. doi: 10.1038/s41568-018-0038-z
78. Di Virgilio F, Sarti AC, Falzoni S, De Marchi E, Adinolfi E. Extracellular ATP and P2 Purinergic Signalling in the Tumour Microenvironment. *Nat Rev Cancer* (2018) 18(10):601–18. doi: 10.1038/s41568-018-0037-0
79. Finicle BT, Jayashankar V, Edinger AL. Nutrient Scavenging in Cancer. *Nat Rev Cancer* (2018) 18(10):619–33. doi: 10.1038/s41568-018-0048-x
80. Gurunathan S, Kang MH. Kim JH. A Comprehensive Review on Factors Influences Biogenesis, Functions, Therapeutic and Clinical Implications of Exosomes. *Int J Nanomed* (2021) 16:1281–312. doi: 10.2147/IJN.S291956
81. Matsumoto Y, Kano M, Murakami K, Toyozumi T, Suito H, Takahashi M, et al. Tumor-Derived Exosomes Influence the Cell Cycle and Cell Migration of Human Esophageal Cancer Cell Lines. *Cancer Sci* (2020) 111(12):4348–58. doi: 10.1111/cas.14660
82. Paduch R. The Role of Lymphangiogenesis and Angiogenesis in Tumor Metastasis. *Cell Oncol (Dordr)* (2016) 39(5):397–410. doi: 10.1007/s13402-016-0281-9
83. Ludwig N, Whiteside TL. Potential Roles of Tumor-Derived Exosomes in Angiogenesis. *Expert Opin Ther Targets* (2018) 22(5):409–17. doi: 10.1080/14728222.2018.1464141
84. Wang Z, Dabrosin C, Yin X, Fuster MM, Arreola A, Rathmell WK, et al. Broad Targeting of Angiogenesis for Cancer Prevention and Therapy. *Semin Cancer Biol* (2015) 35 Suppl:S224–S43. doi: 10.1016/j.semcancer.2015.01.001
85. Olejarz W, Kubiak-Tomaszewska G, Chrzanowska A, Lorenc T. Exosomes in Angiogenesis and Anti-Angiogenic Therapy in Cancers. *Int J Mol Sci* (2020) 21(16):5840. doi: 10.3390/ijms21165840
86. Mao Y, Wang Y, Dong L, Zhang Y, Zhang Y, Wang C, et al. Hypoxic Exosomes Facilitate Angiogenesis and Metastasis in Esophageal Squamous Cell Carcinoma Through Altering the Phenotype and Transcriptome of Endothelial Cells. *J Exp Clin Cancer Res* (2019) 38(1):389. doi: 10.1186/s13046-019-1384-8
87. Guan X. Cancer Metastases: Challenges and Opportunities. *Acta Pharm Sin B* (2015) 5(5):402–18. doi: 10.1016/j.apsb.2015.07.005
88. Ganesh K, Massague J. Targeting Metastatic Cancer. *Nat Med* (2021) 27 (1):34–44. doi: 10.1038/s41591-020-01195-4
89. Lambert AW, Weinberg RA. Linking EMT Programmes to Normal and Neoplastic Epithelial Stem Cells. *Nat Rev Cancer* (2021) 21(5):325–38. doi: 10.1038/s41568-021-00332-6
90. Wortzel I, Dror S, Kenific CM, Lyden D. Exosome-Mediated Metastasis: Communication From a Distance. *Dev Cell* (2019) 49(3):347–60. doi: 10.1016/j.devcel.2019.04.011
91. Wu X, Xiao S, Zhang M, Yang L, Zhong J, Li B, et al. A Novel Protein Encoded by Circular SMO RNA Is Essential for Hedgehog Signaling Activation and Glioblastoma Tumorigenicity. *Genome Biol* (2021) 22 (1):33. doi: 10.1186/s13059-020-02250-6
92. Shi X, Zhang Z, Zhan X, Cao M, Satoh T, Akira S, et al. An Epigenetic Switch Induced by Shh Signalling Regulates Gene Activation During Development and Medulloblastoma Growth. *Nat Commun* (2014) 5:5425. doi: 10.1038/ncomms6425
93. Min H, Sun X, Yang X, Zhu H, Liu J, Wang Y, et al. Exosomes Derived From Irradiated Esophageal Carcinoma-Infiltrating T Cells Promote Metastasis by Inducing the Epithelial-Mesenchymal Transition in Esophageal Cancer Cells. *Pathol Oncol Res* (2018) 24(1):11–8. doi: 10.1007/s12253-016-0185-z
94. Wang M, Zhao X, Huang F, Wang L, Huang J, Gong Z, et al. Exosomal Proteins: Key Players Mediating Premetastatic Niche Formation and Clinical Implications (Review). *Int J Oncol* (2021) 58(4):4. doi: 10.3892/ijo.2021.5184
95. Shaheen O, Ghibour A, Alsaid B. Esophageal Cancer Metastases to Unexpected Sites: A Systematic Review. *Gastroenterol Res Pract* (2017) 2017:1657310. doi: 10.1155/2017/1657310
96. Gonzalez H, Hagerling C, Werb Z. Roles of the Immune System in Cancer: From Tumor Initiation to Metastatic Progression. *Genes Dev* (2018) 32(19-20):1267–84. doi: 10.1101/gad.314617.118
97. Ozga AJ, Chow MT, Luster AD. Chemokines and the Immune Response to Cancer. *Immunity* (2021) 54(5):859–74. doi: 10.1016/j.immuni.2021.01.012
98. Syn NL, Teng MWL, Mok TSK, Soo RA. De-Novo and Acquired Resistance to Immune Checkpoint Targeting. *Lancet Oncol* (2017) 18(12):e731–e41. doi: 10.1016/S1470-2045(17)30607-1

99. Mao Y, Wang Y, Dong L, Zhang Q, Wang C, Zhang Y, et al. Circulating Exosomes From Esophageal Squamous Cell Carcinoma Mediate the Generation of B10 and PD-1(High) Breg Cells. *Cancer Sci* (2019) 110 (9):2700–10. doi: 10.1111/cas.14122
100. Goodman KA, Ou FS, Hall NC, Bekaii-Saab T, Fruth B, Twohy E, et al. Randomized Phase II Study of PET Response-Adapted Combined Modality Therapy for Esophageal Cancer: Mature Results of the CALGB 80803 (Alliance) Trial. *J Clin Oncol* (2021) JCO2003611. doi: 10.1200/JCO.20.03611
101. Hulshof M, Geijsen ED, Rozema T, Oppedijk V, Buijsen J, Neelis KJ, et al. Randomized Study on Dose Escalation in Definitive Chemoradiation for Patients With Locally Advanced Esophageal Cancer (ARTDECO Study). *J Clin Oncol* (2021) JCO2003697. doi: 10.1200/JCO.20.03697
102. Li S, Yi M, Dong B, Jiao Y, Luo S, Wu K. The Roles of Exosomes in Cancer Drug Resistance and Its Therapeutic Application. *Clin Transl Med* (2020) 10 (8):e257. doi: 10.1002/ctm2.257
103. Jella KK, Rani S, O'Driscoll L, McClean B, Byrne HJ, Lyng FM. Exosomes Are Involved in Mediating Radiation Induced Bystander Signaling in Human Keratinocyte Cells. *Radiat Res* (2014) 181(2):138–45. doi: 10.1667/RR13337.1
104. Liang H, Fan JH, Qiao YL. Epidemiology, Etiology, and Prevention of Esophageal Squamous Cell Carcinoma in China. *Cancer Biol Med* (2017) 14(1):33–41. doi: 10.20892/j.issn.2095-3941.2016.0093
105. Global Burden of Disease Cancer C, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* (2017) 3(4):524–48. doi: 10.1001/jamaoncol.2016.5688
106. Mannath J, Ragunath K. Role of Endoscopy in Early Oesophageal Cancer. *Nat Rev Gastroenterol Hepatol* (2016) 13(12):720–30. doi: 10.1038/nrgastro.2016.148
107. Chen X, Gole J, Gore A, He Q, Lu M, Min J, et al. Non-Invasive Early Detection of Cancer Four Years Before Conventional Diagnosis Using a Blood Test. *Nat Commun* (2020) 11(1):3475. doi: 10.1038/s41467-020-17316-z
108. Grady WM, Yu M, Markowitz SD. Epigenetic Alterations in the Gastrointestinal Tract: Current and Emerging Use for Biomarkers of Cancer. *Gastroenterology* (2021) 160(3):690–709. doi: 10.1053/j.gastro.2020.09.058
109. Mori MA, Ludwig RG, Garcia-Martin R, Brandao BB, Kahn CR. Extracellular miRNAs: From Biomarkers to Mediators of Physiology and Disease. *Cell Metab* (2019) 30(4):656–73. doi: 10.1016/j.cmet.2019.07.011
110. O'Brien K, Breyne K, Ughetto S, Laurent LC, Breakefield XO. RNA Delivery by Extracellular Vesicles in Mammalian Cells and Its Applications. *Nat Rev Mol Cell Biol* (2020) 21(10):585–606. doi: 10.1038/s41580-020-0251-y
111. Thietart S, Rautou PE. Extracellular Vesicles as Biomarkers in Liver Diseases: A Clinician's Point of View. *J Hepatol* (2020) 73(6):1507–25. doi: 10.1016/j.jhep.2020.07.014
112. van der Pol Y, Mouliere F. Toward the Early Detection of Cancer by Decoding the Epigenetic and Environmental Fingerprints of Cell-Free DNA. *Cancer Cell* (2019) 36(4):350–68. doi: 10.1016/j.ccell.2019.09.003
113. Hoshino A, Kim HS, Bojmar L, Gyan KE, Cioffi M, Hernandez J, et al. Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers. *Cell* (2020) 182(4):1044–61.e18. doi: 10.1016/j.cell.2020.07.009
114. Thind A, Wilson C. Exosomal miRNAs as Cancer Biomarkers and Therapeutic Targets. *J Extracell Vesicles* (2016) 5:31292. doi: 10.3402/jev.v5.31292
115. Salehi M, Sharifi M. Exosomal miRNAs as Novel Cancer Biomarkers: Challenges and Opportunities. *J Cell Physiol* (2018) 233(9):6370–80. doi: 10.1002/jcp.26481
116. Fitts CA, Ji N, Li Y, Tan C. Exploiting Exosomes in Cancer Liquid Biopsies and Drug Delivery. *Adv Healthc Mater* (2019) 8(6):e1801268. doi: 10.1002/adhm.201801268
117. Tanaka Y, Kamohara H, Kinoshita K, Kurashige J, Ishimoto T, Iwatsuki M, et al. Clinical Impact of Serum Exosomal microRNA-21 as a Clinical Biomarker in Human Esophageal Squamous Cell Carcinoma. *Cancer* (2013) 119(6):1159–67. doi: 10.1002/cncr.27895
118. Qiu ML, Li X, Lin JB, Luo RG, Liu B, Feng Z. Serum Exosomal miR-182 Upregulation Predicts Unfavorable Prognosis of Esophageal Squamous Cell Carcinoma. *Eur Rev Med Pharmacol Sci* (2020) 24(10):5412–8. doi: 10.26355/eurrev_202005_21325
119. Liu S, Lin Z, Zheng Z, Rao W, Lin Y, Chen H, et al. Serum Exosomal microRNA-766-3p Expression Is Associated With Poor Prognosis of Esophageal Squamous Cell Carcinoma. *Cancer Sci* (2020) 111(10):3881–92. doi: 10.1111/cas.14550
120. Liu S, Lin Z, Rao W, Zheng J, Xie Q, Lin Y, et al. Upregulated Expression of Serum Exosomal Hsa_Circ_0026611 Is Associated With Lymph Node Metastasis and Poor Prognosis of Esophageal Squamous Cell Carcinoma. *J Cancer* (2021) 12(3):918–26. doi: 10.7150/jca.50548
121. Ni PZ, He JZ, Wu ZY, Ji X, Chen LQ, Xu XE, et al. Overexpression of Stathmin 1 Correlates With Poor Prognosis and Promotes Cell Migration and Proliferation in Oesophageal Squamous Cell Carcinoma. *Oncol Rep* (2017) 38(6):3608–18. doi: 10.3892/or.2017.6039
122. Ma HL, Jin SF, Ju WT, Fu Y, Tu YY, Wang LZ, et al. Stathmin Is Overexpressed and Regulated by Mutant P53 in Oral Squamous Cell Carcinoma. *J Exp Clin Cancer Res* (2017) 36(1):109. doi: 10.1186/s13046-017-0575-4
123. Rong B, Nan Y, Liu H, Gao W. Increased Stathmin Correlates With Advanced Stage and Poor Survival of Non-Small Cell Lung Cancer. *Cancer Biomark* (2017) 19(1):35–43. doi: 10.3233/CBM-160239
124. Li M, Yang J, Zhou W, Ren Y, Wang X, Chen H, et al. Activation of an AKT/FOXM1/STMN1 Pathway Drives Resistance to Tyrosine Kinase Inhibitors in Lung Cancer. *Br J Cancer* (2017) 117(7):974–83. doi: 10.1038/bjc.2017.292
125. Yan L, Dong X, Gao J, Liu F, Zhou L, Sun Y, et al. A Novel Rapid Quantitative Method Reveals Stathmin-1 as a Promising Marker for Esophageal Squamous Cell Carcinoma. *Cancer Med* (2018) 7(5):1802–13. doi: 10.1002/cam4.1449
126. Liu MX, Liao J, Xie M, Gao ZK, Wang XH, Zhang Y, et al. miR-93-5p Transferred by Exosomes Promotes the Proliferation of Esophageal Cancer Cells *via* Intercellular Communication by Targeting PTEN. *BioMed Environ Sci* (2018) 31(3):171–85. doi: 10.3967/bes2018.023
127. Liu T, Du LT, Wang YS, Gao SY, Li J, Li PL, et al. Development of a Novel Serum Exosomal MicroRNA Nomogram for the Preoperative Prediction of Lymph Node Metastasis in Esophageal Squamous Cell Carcinoma. *Front Oncol* (2020) 10:573501. doi: 10.3389/fonc.2020.573501
128. Tian L, Yang L, Zheng W, Hu Y, Ding P, Wang Z, et al. RNA Sequencing of Exosomes Revealed Differentially Expressed Long Noncoding RNAs in Early-Stage Esophageal Squamous Cell Carcinoma and Benign Esophagitis. *Epigenomics* (2020) 12(6):525–41. doi: 10.2217/epi-2019-0371
129. Jiao Z, Yu A, Rong W, He X, Zen K, Shi M, et al. Five-lncRNA Signature in Plasma Exosomes Serves as Diagnostic Biomarker for Esophageal Squamous Cell Carcinoma. *Aging (Albany NY)* (2020) 12(14):15002–10. doi: 10.18632/aging.103559
130. Yan S, Du L, Jiang X, Duan W, Li J, Xie Y, et al. Evaluation of Serum Exosomal lncRNAs as Diagnostic and Prognostic Biomarkers for Esophageal Squamous Cell Carcinoma. *Cancer Manag Res* (2020) 12:9753–63. doi: 10.2147/CMAR.S250971
131. Fan L, Cao Q, Liu J, Zhang J, Li B. Circular RNA Profiling and Its Potential for Esophageal Squamous Cell Cancer Diagnosis and Prognosis. *Mol Cancer* (2019) 18(1):16. doi: 10.1186/s12943-018-0936-4
132. Lin Y, Dong H, Deng W, Lin W, Li K, Xiong X, et al. Evaluation of Salivary Exosomal Chimeric GOLM1-NAA35 RNA as a Potential Biomarker in Esophageal Carcinoma. *Clin Cancer Res* (2019) 25(10):3035–45. doi: 10.1158/1078-0432.CCR-18-3169
133. Ding Z, Yan Y, Guo YL, Wang C. Esophageal Carcinoma Cell-Excreted Exosomal Uc.189 Promotes Lymphatic Metastasis. *Aging (Albany NY)* (2021) 13(10):13846–58. doi: 10.18632/aging.202979
134. Tikhomirov R, Donnell BR, Catapano F, Faggian G, Gorelik J, Martelli F, et al. Exosomes: From Potential Culprits to New Therapeutic Promise in the Setting of Cardiac Fibrosis. *Cells* (2020) 9(3):592. doi: 10.3390/cells9030592
135. Dang XTT, Kavishka JM, Zhang DX, Pirisinu M, Le MTN. Extracellular Vesicles as an Efficient and Versatile System for Drug Delivery. *Cells* (2020) 9 (10):2191. doi: 10.3390/cells9102191
136. Shao J, Zaro J, Shen Y. Advances in Exosome-Based Drug Delivery and Tumor Targeting: From Tissue Distribution to Intracellular Fate. *Int J Nanomed* (2020) 15:9355–71. doi: 10.2147/IJN.S281890

137. Ha D, Yang N, Nadithe V. Exosomes as Therapeutic Drug Carriers and Delivery Vehicles Across Biological Membranes: Current Perspectives and Future Challenges. *Acta Pharm Sin B* (2016) 6(4):287–96. doi: 10.1016/j.apsb.2016.02.001
138. Liao W, Du Y, Zhang C, Pan F, Yao Y, Zhang T, et al. Exosomes: The Next Generation of Endogenous Nanomaterials for Advanced Drug Delivery and Therapy. *Acta Biomater* (2019) 86:1–14. doi: 10.1016/j.actbio.2018.12.045
139. Bari E, Ferrarotti I, Di Silvestre D, Grisoli P, Barzon V, Balderacchi A, et al. Adipose Mesenchymal Extracellular Vesicles as Alpha-1-Antitrypsin Physiological Delivery Systems for Lung Regeneration. *Cells* (2019) 8 (9):965. doi: 10.3390/cells8090965
140. Ullah M, Liu DD, Rai S, Razavi M, Choi J, Wang J, et al. A Novel Approach to Deliver Therapeutic Extracellular Vesicles Directly Into the Mouse Kidney *via* Its Arterial Blood Supply. *Cells* (2020) 9(4):937. doi: 10.3390/cells9040937
141. Matsumura S, Minamisawa T, Suga K, Kishita H, Akagi T, Ichiki T, et al. Subtypes of Tumour Cell-Derived Small Extracellular Vesicles Having Differently Externalized Phosphatidylserine. *J Extracell Vesicles* (2019) 8 (1):1579541. doi: 10.1080/20013078.2019.1579541
142. Ding L, Yang X, Gao Z, Effah CY, Zhang X, Wu Y, et al. A Holistic Review of the State-Of-the-Art Microfluidics for Exosome Separation: An Overview of the Current Status, Existing Obstacles, and Future Outlook. *Small* (2021) 17 (29):e2007174. doi: 10.1002/smll.202007174
143. Meng W, He C, Hao Y, Wang L, Li L, Zhu G. Prospects and Challenges of Extracellular Vesicle-Based Drug Delivery System: Considering Cell Source. *Drug Delivery* (2020) 27(1):585–98. doi: 10.1080/10717544.2020.1748758
144. Samal S, Dash P, Dash M. Drug Delivery to the Bone Microenvironment Mediated by Exosomes: An Axiom or Enigma. *Int J Nanomed* (2021) 16:3509–40. doi: 10.2147/IJN.S307843
145. Chen L, Hong W, Ren W, Xu T, Qian Z, He Z. Recent Progress in Targeted Delivery Vectors Based on Biomimetic Nanoparticles. *Signal Transduct Target Ther* (2021) 6(1):225. doi: 10.1038/s41392-021-00631-2
146. Go G, Lee J, Choi DS, Kim SS, Gho YS. Extracellular Vesicle-Mimetic Ghost Nanovesicles for Delivering Anti-Inflammatory Drugs to Mitigate Gram-Negative Bacterial Outer Membrane Vesicle-Induced Systemic Inflammatory Response Syndrome. *Adv Healthc Mater* (2019) 8(4): e1801082. doi: 10.1002/adhm.201801082
147. Gao X, Ran N, Dong X, Zuo B, Yang R, Zhou Q, et al. Anchor Peptide Captures, Targets, and Loads Exosomes of Diverse Origins for Diagnostics and Therapy. *Sci Transl Med* (2018) 10(444):eaat0195. doi: 10.1126/scitranslmed.aat0195
148. Kamerkar S, LeBleu VS, Sugimoto H, Yang S, Ruivo CF, Melo SA, et al. Exosomes Facilitate Therapeutic Targeting of Oncogenic KRAS in Pancreatic Cancer. *Nature* (2017) 546(7659):498–503. doi: 10.1038/nature22341

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 24 September 2021
doi: 10.3389/fonc.2021.734134

Targeted Next-Generation Sequencing of Cancer-Related Genes in a Norwegian Patient Cohort With Head and Neck Squamous Cell Carcinoma Reveals Novel Actionable Mutations and Correlations With Pathological Parameters

Harsh N. Dongre[1,2,3], Hilde Haave[4,5], Siren Fromreide[2,3], Fredrik A. Erland[4,5], Svein Erik Emblem Moe[4,5], Sophia Manueldas Dhayalan[1], Rasmus Kopperud Riis[1], Dipak Sapkota[6], Daniela Elena Costea[1,2,3], Hans Jorgen Aarstad[4,5] and Olav K. Vintermyr[1,2]*

[1] Department of Pathology, Haukeland University Hospital, Bergen, Norway, [2] Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway, [3] Centre for Cancer Biomarkers CCBIO, Department of Clinical Medicine, University of Bergen, Bergen, Norway, [4] Department of Otolaryngology/Head and Neck Surgery, Haukeland University Hospital, Bergen, Norway, [5] Otolaryngology, Department of Clinical Medicine, Faculty of Medicine, University of Bergen, Bergen, Norway, [6] Institute of Oral Biology, Faculty of Dentistry, University of Oslo, Oslo, Norway

OPEN ACCESS

Edited by:
Die Wang, Hudson Institute of Medical Research, Australia

Reviewed by:
Monica Charlotte Solomon, Manipal College of Dental Sciences, Manipal, India
Xi Yang, Shanghai Jiao Tong University, China

***Correspondence:**
Harsh N. Dongre
harsh.dongre@uib.no

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *30 June 2021*
Accepted: *27 August 2021*
Published: *24 September 2021*

Citation:
Dongre HN, Haave H, Fromreide S, Erland FA, Moe SEE, Dhayalan SM, Riis RK, Sapkota D, Costea DE, Aarstad HJ and Vintermyr OK (2021) Targeted Next-Generation Sequencing of Cancer-Related Genes in a Norwegian Patient Cohort With Head and Neck Squamous Cell Carcinoma Reveals Novel Actionable Mutations and Correlations With Pathological Parameters. Front. Oncol. 11:734134. doi: 10.3389/fonc.2021.734134

Background: Targeted next-generation sequencing (NGS) is increasingly applied in clinical oncology to advance personalized treatment. Despite success in many other tumour types, use of targeted NGS panels for assisting diagnosis and treatment of head and neck squamous cell carcinomas (HNSCC) is still limited.

Aim: The focus of this study was to establish a robust NGS panel targeting most frequent cancer mutations in long-term preserved formalin-fixed paraffin-embedded (FFPE) tissue samples of HNSCC from routine diagnostics.

Materials and Methods: Tumour DNA obtained from archival FFPE tissue blocks of HNSCC patients treated at Haukeland University Hospital between 2003-2016 (n=111) was subjected to mutational analysis using a custom made AmpliSeq Library PLUS panel targeting 31 genes (Illumina). Associations between mutational burden and clinical and pathological parameters were investigated. Mutation and corresponding clinicopathological data from HNSCC were extracted for selected genes from the Cancer Genome Atlas (TCGA) and used for Chi-square and Kaplan-Meier analysis.

Results: The threshold for sufficient number of reads was attained in 104 (93.7%) cases. Although the specific number of PCR amplified reads detected decreased, the number of NGS-annotated mutations did not significantly change with increased tissue preservation time. In HPV-negative carcinomas, mutations were detected mainly in *TP53* (73.3%), *FAT1* (26.7%) and *FLG* (16.7%) whereas in HPV-positive, the common mutations were in

FLG (24.3%) FAT1 (17%) and *FGFR3* (14.6%) genes. Other less common pathogenic mutations, including well reported SNPs were reproducibly identified. Presence of at least one cancer-specific mutations was found to be positively associated with an extensive desmoplastic stroma (p=0.019), and an aggressive type of invasive front (p=0.035), and negatively associated with the degree of differentiation (p=0.041). Analysis of TCGA data corroborated the association between cancer-specific mutations and tumour differentiation and survival analysis showed that tumours with at least one mutation had shorter disease-free and overall survival (p=0.005).

Conclusions: A custom made targeted NGS panel could reliably detect several specific mutations in archival samples of HNSCCs preserved up to 17 years. Using this method novel associations between mutational burden and clinical and pathological parameters were detected and actionable mutations in HPV-positive HNSCC were discovered.

Keywords: next-generation sequencing (NGS), actionable mutation, pathological parameters, stromal desmoplasia, inflammation, head and neck squamous cell carcinoma (HNSCC)

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC), arising in the oral cavity, oropharynx, hypopharynx and larynx, is a major public health concern worldwide (1). The incidence and mortality rates of HNSCC are increasing, with approximately 750,000 new cases and 360,000 deaths reported in 2020 (1). HNSCC are characterized by aetiological, phenotypical, biological, and clinical heterogeneity (2, 3). Smoking and smokeless tobacco chewing habits in developing countries and human papillomavirus (HPV) infection in developed countries are emerging as important risk factors for the rise in incidence rates of HNSCC (2, 3). Given the reported differences in survival outcomes based on HPV and smoking status, establishing HPV as a prognostic factor for better survival in oropharyngeal SCC (OPSCC) (4, 5), there is an additional interest in understanding the molecular differences between HPV-positive and HPV-negative lesions especially in OPSCC.

Current cancer treatment is moving towards more personalized and targeted treatment. Both targeted and immunotherapies have improved 5-year survival rates significantly in many types of cancers (6–8). However, for patients with HNSCC the only targeted therapy approved is cetuximab, a monoclonal antibody against epidermal growth factor receptor (EGFR) (6, 9). Recently, pembrolizumab and nivolumab, immunotherapeutic agents, have been approved for HNSCC (10, 11). Nonetheless, patients treated with targeted therapy such as cetuximab combined with radiotherapy develop resistance (6, 12) and less than 20% of HNSCC patients treated with immunotherapy achieve a lasting response (13).

Several studies have investigated genomic aberrations associated with HNSCC and this has greatly increased our understanding of the mutational landscape of HNSCC (2, 3, 14–16). Moreover, with the advent of next-generation sequencing (NGS), several studies have added to the ever-growing list of novel genetic alterations in HNSCC. Frequent mutations in several genes including *TP53, CDKN2A, PIK3CA, NOTCH1, CASP8* and *MLL2 (KMT2D)* among others as well as alterations in *EGFR, CCND1* and *FGFR* have been reported (14–16). However, these genes are yet to translate into clinically beneficial prognostic or predictive biomarkers. The incorporation of NGS into routine clinical setting has been hampered by lack of large number of biopsy specimens. Formalin-fixed paraffin-embedded (FFPE) archival tissues could overcome this as they are a rich resource that could allow correlations of specific mutational landscape with clinical and pathological parameters in long-term follow-up retrospective studies.

The aim of this study was to establish a robust NGS panel targeting most frequent cancer mutations in long-term preserved FFPE samples of HNSCC that would allow testing of clinical outcome and correlations with clinical and pathological parameters of cancer-specific mutations. This method could serve for identification of novel, actionable mutations that would open new avenues for personalized therapy in HNSCC. In addition, it will also bring new insights into the biology of this disease by allowing investigation of correlations between mutational landscape of tumours, clinical variables, and histopathological parameters, which is difficult to assess in fresh-frozen tissues. In this study we report that a custom made targeted NGS panel could reliably detect several specific mutations in archival samples of HNSCCs preserved up to 17 years. Using this method, we could identify targetable mutations in HNSCC and reveal novel associations between the mutational load and clinicopathological variables like tumour desmoplasia, pattern of invasion at the tumour front and the degree of differentiation.

MATERIALS AND METHODS

Patient Cohort

The cohort consisted of consecutive patients diagnosed with HNSCC at the Department of Otolaryngology/Head & Neck Surgery, Haukeland University Hospital (HUH), Bergen, Norway between 2003 to 2010 and 2013 to 2016, who consented

to participate in the study. In total, 111 HNSCC patients and 9 cancer-free control patients were enrolled in this retrospective study. Due to insufficient number of reads, seven HNSCC patients and two cancer-free controls were excluded from the final analysis. Out of 104 patients with NGS data that passed the threshold of reads, 43 (41.3%) patients were HPV-positive, 57 (54.8%) patients were HPV-negative, and 4 (3.8%) were not tested for HPV. Clinical data (age, gender, smoking, TNM staging at the point of diagnosis according to the American Joint Committee on Cancer (AJCC) manual 6th edition) was obtained from the Electronic Patient Journal at Haukeland University Hospital (DIPS). To ensure continuity in the study with respect to HNSCC classification during the whole period (2003-2016), staging of all the tumours has been set based on the AJCC 6th edition. The study was approved by the Regional (REK Vest) Committee for Medical and Health Research Ethics (2011/125).

Histopathological Evaluation of Haematoxylin and Eosin (HE) Sections

All morphological evaluations were done on representative HE sections from FFPE blocks by an expert pathologist (OKV) and a specialist in surgery (HH). The morphological scoring was done by consensus evaluation based on a pre-defined list of criteria and without knowledge of the HPV status following the scoring system described elsewhere (17, 18). The following tumour phenotypic criteria were evaluated: a) degree of keratinization, b) fraction maturating cells (%), and c) tumour stromal invasion pattern. In addition, the following host response patterns to tumour were evaluated: d) inflammatory response and e) stromal desmoplasia. Each parameter was scored 1-4 and further defined in **Supplementary Table 1** in detail. Degree of keratinization was defined as tumour area with apparent keratinized single cells, groups of cells or areas with remnants of keratinized cells. Fraction maturing cells was denoted as the fraction of immature (basal like cells) cells relative to the more differentiated cell fraction in the tumour sample. The invasion pattern was evaluated at the interphase between tumour and the intervening stroma at the invasive tumour front (19). The inflammatory host response was evaluated as the extent of chronic inflammation denoted by the presence of lymphocytes, plasma cells and histiocytic cells surrounding the invading tumour front. Areas with neutrophil granulocytic responses were not considered as these areas may represent secondary host responses due to tumour cell death, infection or ulcerations. Tumour stromal desmoplasia was evaluated only in the context of fibrillary and myxomatous tumour host responses *i.e.*, the angiogenetic tumour host responses was not included in this parameter.

DNA Isolation and HPV Detection

All tumour samples were reviewed, and representative tissue samples were selected. DNA was extracted from FFPE tissue blocks. These were tissues from primary HNSCC from diagnostic or surgical samples collected 5 to 17 years ago. For the NGS analysis at least 20% tumour fraction was required to be included in the analysis. Many of FFPE blocks were macro-dissected to enrich the tumour fraction of the patient samples. The mean tumour fraction was 47.5% (median 50%) in the samples that were analysed with NGS. Three to six 10 µm thick FFPE sections were de-paraffinized in deparaffinization solution and digested overnight in ATL-buffer and Proteinase K (Qiagen GmbH, Hilden, Germany) at 56°C. DNA was extracted using the E.Z.N.A tissue DNA kit (Omega BioTek, Norcross, GA, USA), and the DNA concentration was quantified using Qubit dsDNA BR assay kit (Thermo Fisher Scientific, Waltham, MA, USA) according to manufacturer's protocol.

HPV DNA detection was performed using standard Gp5+/Gp6+ primers as described elsewhere (20, 21). p16 status of all HPV-positive OPSCC from 2003-2010 (75% of cohort) was also determined by performing immunohistochemistry (IHC) using anti-p16^{INK4a} monoclonal antibody (clone E6H4, Roche diagnostics, Switzerland) as described elsewhere (5, 22). The cut-off percentage for p16 positivity was set as 70% as described in literature for OPSCC (5).

Design of Custom Made NGS Panel

The custom made NGS panel was based on previously reported mutational burden in HPV-negative and HPV-positive HNSCC (15, 23, 24). For some genes, the specific targeted regions were adapted from the commercially available Illumina TruSight Tumor 15 (*TP53, NRAS*) and TruSight Tumor 26 (*AKT 1, BRAF, CTNNB1, KIT, PTEN*) NGS panels (Illumina, San Diego, CA, USA). The PCR primer sets for the other targeted gene regions were *de novo* annotated and designed by Illumina DesignStudio. A custom-made NGS panel, targeting hotspot gene exon regions in 28 genes and complete exon coverage in 3 genes (*TP53, TRAF 3 and FAT1*) was designed. The list of genes is shown in **Supplementary Table 2**. Year- and side-matched normal samples (n=7), samples with known specific mutations, and Acrometrix oncology hotspot control (Thermo Fisher Scientific, Waltham, CA, USA) were used for quality control of the NGS panel.

NGS Library Preparation and Sequencing

DNA library preparation was conducted using the AmpliSeq Library PLUS kit (Illumina, San Diego, CA, USA) according to manufacturer's instructions. Briefly, 50 ng of DNA was used to prepare amplicon libraries with two custom-made primer pools. The amplicons were then partially digested, ligated to AmpliSeq CD indexes and purified with AMPure XP beads (Beckman Coulter, High Wycombe, UK). Further, the library was again amplified, purified with beads, and then quantified using Qubit dsDNA HS kit (Thermo Fisher Scientific, Waltham, MA, USA). Samples were pooled in equimolar concentrations, and finally 1.3 pM of library solution was loaded for paired-end sequencing on the Illumina MiniSeq platform using the BaseSpace Sequence Hub (Illumina, San Diego, CA, USA).

Bioinformatics

Bioinformatical analysis was performed using the DNA amplicon v.2.1.1 workflow in BaseSpace (Illumina, San Diego, CA, USA). The targeted regions (specified in a custom manifest file) were aligned to the reference genome hg19/GRCh37 using the Burrows Wheeler Aligner. Variants were called using the Somatic variant caller and annotated by RefSeq using

National Centre for Biotechnology Information (NCBI) database. An amplicon total depth (coverage) of more than 500 reads and a variant allele frequency (VAF) of more than 5% were set as strict thresholds to be included in the analysis. The recorded variants were evaluated in dbSNP (National Centre for Biotechnology Information) and the Cosmic (Catalog of Somatic Mutations in Cancer) databases and only coding sequences were considered (25). The European population frequencies for the variants were found in the ALFA Allele Frequency Aggregator (26). If a mutation had an allele frequency of more than 1% it was considered a single nucleotide polymorphism (SNP). Mutations with an allele frequency in the range 0.1% - 1% were also considered a SNP unless they had a pathogenic score of more than 0.5 in the Cosmic database. All mutations recorded were of the non-synonymous type, but for SNPs synonymous mutations were also recorded separately. Four normal FFPE tonsillar tissue samples from the period 2003-2010 and three normal FFPE tonsillar samples from the period 2012-2015 were analysed to validate mutations with respect to unspecific variants due to very aged FFPE tumour samples in this study.

Graphs related to mutational signatures were created using a web based application: Mutational Signatures in Cancer (MuSiCa) (27). Pictorial distribution of SNVs on protein domain structures was plotted using cBioPortal online tool (28, 29).

Statistical Analysis

The statistical analysis was performed using IBM SPSS Statistics Version 25.0 (Armonk, IBM Corp, USA). The data is presented as mean ± SEM (standard error over mean) and significance set at $p<0.05$. Tests of independence of clinicopathological parameters with mutational burden were assessed using Pearson's chi-squared test and Phi correlation coefficient was calculated to measure the strength of association between variables tested. For correlation analysis, cases were categorized into higher and lower mutational burden groups by mean value or a pre-set cut-off point; pathological parameters were dichotomized in lower (1&2) and higher scores (3&4).

Analysis of TCGA Data

Mutation status for *TP53, FAT1, FLG, CDKN2A, FGFR3, NSD-1*, and *KMT2C* genes and corresponding clinicopathological data for HNSCC (n=496) specimens were exported using cBioPortal online tool (28, 29). The data were imported to IBM SPSS for further statistical analysis. HNSCC were stratified with respect to the mutational status of the examined genes and associations between the mutational status and clinicopathological parameters and 5-year overall survival or 5-year disease (relapse) free survival were examined using Chi-square test, and Kaplan-Meier plot (Log Rank test), respectively.

RESULTS

Clinical Characteristics of the HNSCC Cohort

The NGS data from the HNSCC cohort after quality control consisted of 104 patients with SCC at different anatomical head and neck sites: oropharynx (61%), oral cavity (32%), hypopharynx (3%), larynx (2%), and metastatic lymph nodes (3%). The majority of cohort consisted of males (79%), the median age was 63 years (range 38-87), and the majority (68%) had a history of smoking. At the time of diagnosis, 61.5% presented with T1&T2 while 35.6% were at late T3&T4 stages, 54.8% had loco-regional lymph node metastasis, and 3% had distant metastasis.

For determining the HPV status, both HPV DNA and IHC for p16 were performed. The sensitivity and specificity of p16 IHC as compared to HPV DNA detection was 96% and 85.7%, respectively. From all HNSCC cases, 43 (41.3%) were HPV-positive as determined by presence of HPV DNA, with the majority being oropharyngeal SCC (95.4%) and only 2% of oral SCC being HPV-positive. Most of the HPV-positive cancers presented with lymph node metastasis (79.1%) at diagnosis, and only 38.6% of the HPV-negative. A positive moderate correlation between HPV status and locoregional lymph node metastasis was detected ($p<0.001$, $r=0.422$). In addition, among the oropharyngeal cancers, 41 out of 63 (66.7%) were HPV-positive while the rest were HPV-negative (33.3%). Out of these 41 HPV-positive cases, 24 cases (58.5) % were T1 and T2 early-stage tumours and 33 cases (80.5%) had lymph node metastasis ($p=0.018$, $r=0.30$). Clinical details of all the patients are presented in **Table 1**.

Clinico-Pathological Correlations

Including all subsites of HNSCC, HPV-positive cases presented with a non-aggressive tumour invasive front ($p<0.001$), little desmoplastic stromal reaction ($p=0.034$), but rich inflammatory host reaction ($p<0.001$) than the HPV-negative cases (**Figure 1**). Most of the HPV-positive cases presented with a less differentiated ($p<0.001$, $r=0.570$), basal like morphology ($p<0.001$, $r=0.363$), and had a more pushing (non-aggressive) border type of invasion front ($p<0.001$, $r=-0.538$). Consistent with these observations, HPV-positive OPSCC showed non-aggressive tumour invasive front ($p=0.001$), little desmoplastic stromal reaction ($p=0.059$, borderline significant), and rich inflammatory host reaction ($p=0.013$) than HPV-negative OPSCC. In addition, HPV-positive OPSCC displayed a less differentiated ($p=0.003$, $r=0.378$), and basal like morphology ($p=0.081$, $r=0.22$). Bigger tumours (T3&T4) were associated with a richer desmoplastic stroma ($p<0.001$, $r=0.398$), but with a little inflammatory infiltrate ($p=0.018$, $r=0.238$). Extensive stromal desmoplasia positively associated with an aggressive tumour invasion front ($p<0.001$, $r=0.426$) and presence of metastasis ($p=0.045$, $r=0.203$), but inversely associated with the rich inflammatory infiltrate ($p<0.001$, $r=0.-589$). The presence of a rich inflammatory infiltrate was negatively associated with an aggressive type of invasive front ($p=0.001$, $r=-0.330$) (**Figure 1**).

Performance of the NGS Method on FFPE Samples and the Effect of Preservation Time

Of special concern was the performance of the targeted NGS panel on aged FFPE material, especially in the FFPE tissue samples preserved in the diagnostics biobank for longer than

TABLE 1 | Clinical parameters of patients.

Characteristics	All patients n = 104 (%)	HPV-positive n = 43 (%)	HPV-negative n = 57 (%)	p-value
Age at diagnosis	63	59	64	0.381
Median	(38-87)	(38-87)	(39-84)	
Sex:				0.230
Male	82 (78.8)	36 (83.7)	42 (73.7)	
Female	22 (21.2)	07 (16.3)	15 (26.3)	
Tumour site:				**0.0001***
Oral cavity	33 (31.7)	01 (2.3)	32 (56.1)	
Oropharynx	63 (60.6)	41 (95.4)	22 (38.7)	
Hypopharynx	03 (2.9)	01 (2.3)	01 (1.7)[$]	
Larynx	02 (1.9)	00	02 (3.5)	
Metastasis	03 (2.9)	–	–	
Smoking:				0.137
Yes	71 (68.3)	25 (58.1)	42 (73.7)	
No	33 (31.7)	18 (41.9)	15 (26.3)	
T stage:				0.289
T1 and T2	64 (61.5)	25 (58.1)	39 (68.4)	
T3 and T4	37 (35.6)	18 (41.9)	19 (31.6)	
Unknown	03 (2.9)	–	–	
N stage:				**0.0001**
N0	44 (42.3)	9 (20.9)	35 (61.4)	
N+	57 (54.8)	34 (79.1)	23 (38.6)	
Unknown	03 (2.9)	–	–	
M stage:				0.127
M0	98 (94.2)	43 (100)	55 (94.7)	
M+	03 (2.9)	0	03 (5.3)	
Unknown	03 (2.9)	–	–	
Stage grouping[#]				–
Stage I	12 (11.5)	01 (2.3)	10 (17.5)	
Stage II	19 (18.3)	03 (6.9)	16 (28.1)	
Stage III	11 (10.6)	05 (11.7)	06 (10.6)	
Stage IV	59 (56.7)	34 (79.1)	25 (43.8)	
Unknown	03 (2.9)			

p-value is based on Chi-square test.
**Correlation between oral cavity and oropharynx based on HPV status. [#]Based on AJCC 6th edition. [$]HPV DNA test was inconclusive for one patient. Significant values are denoted in bold.*

10 years, from the 2003-2010 period. The average number of reads per amplicon in the longer time preserved group (2003-2010) was 1709 (n=46) *versus* 1862 (n=58) in the shorter time preserved group (2013-2016), compatible with some lower coverage in the more aged FFPE tissue samples. Of the seven cases that were excluded due to low coverage, five cases were from 2003/04, one case from 2010 and one case form 2015, indicating some concern with respect to performance due to ageing of the FFPE material after more than 5 years preservation time. The excluded cases had number of reads ranging from 20-435 with a median of 159. Targeted sequencing of 101 tumours and 3 tumours from metastatic sites resulted in a total of 2099 single nucleotide variations (SNVs) and 15 indels. No specific difference was, however, noted with respect to SNPs or mutational load between the older samples and the newer ones, stored for 17 and 5 years, respectively.

Mutational Landscape of HNSCC as Depicted by the Custom Made NGS Panel

The most frequently mutated genes in this Norwegian cohort of HNSCC were *TP53* (n=45, 43.3%), *FAT1* (n=23, 22.1%), *FLG* (n=20, 19.3%), *CDKN2A* (n=10, 9.6%), and *FGFR3* (n=7, 6.7%) (**Figure 2**). The tumour mutational burden (TMB), defined as total number of non-synonymous mutations per coding area of sequenced region, was 2.218 ± 0.2965 for the whole cohort. Smokers displayed higher TMB as compared to non-smokers (2.424 ± 0.4113 and 1.777 ± 0.2940, respectively), although the difference was not statistically significant. TMB was significantly higher in HPV-negative tumours (2.714 ± 0.4292) as compared to HPV-positive (1.515 ± 0.3576, p=0.04). Among the HPV-negative tumours, the highest TMB was detected in oral cavity SCCs (3.239 ± 0.7423), as compared to oropharyngeal SCCs (2.134 ± 0.3041, p=0.2471) (**Figure 2A**). Also, HPV-positive OPSCC had lower TMB (1.526 ± 0.375) as compared to

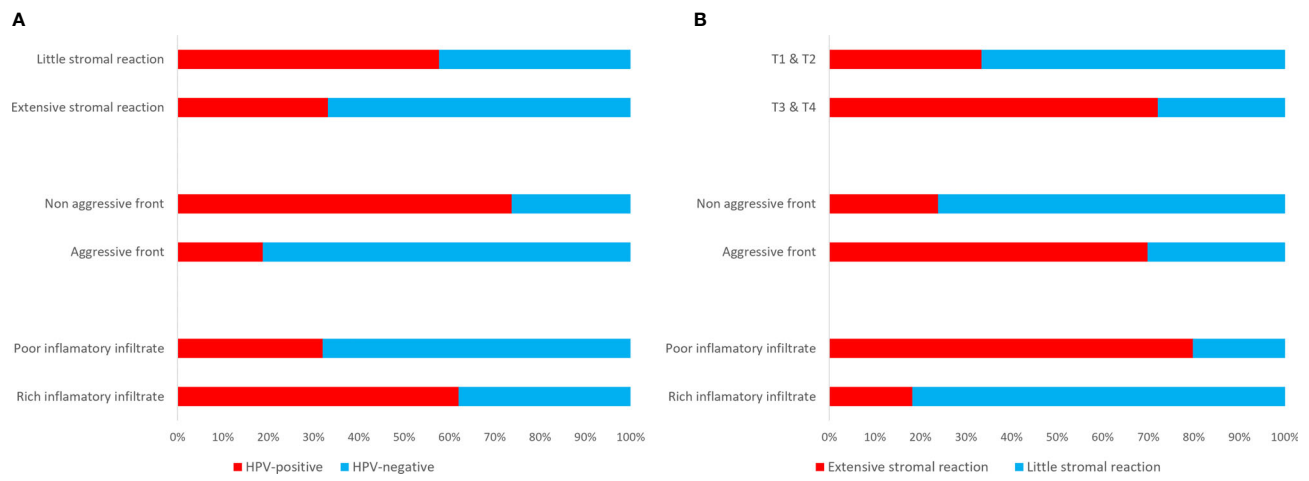


FIGURE 1 | Clinico-pathological correlations depicting an extensive (scores 3&4) stromal reaction, an aggressive (scores 3&4) invasive front and a poor (scores 3&4) inflammatory infiltrate in HPV-negative HNSCC as compared to the HPV-positive cases **(A)**, and associations between extensive stromal reaction and bigger size tumours, an aggressive invasive front, and a poor inflammatory infiltrate **(B)**.

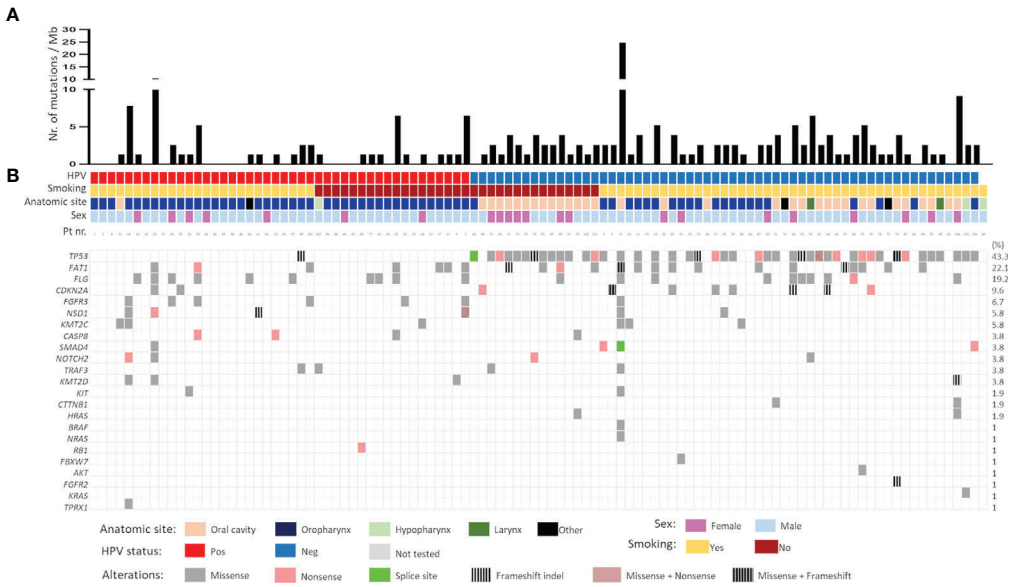


FIGURE 2 | **(A)** Mutational burden and **(B)** mutational spectrum of head and neck squamous cell carcinomas as detected by targeted sequencing. Samples were stratified according to HPV status followed by smoking status. Gene list is shown on the left whereas the corresponding frequency of mutation is shown on the right. Pos- positive; Neg- negative.

HPV-negative OPSCC (2.133 ± 0.303, p=0.005). There was no difference in TMB in the two different types of sexes, female: (2.1927 ± 0.448) and male: (2.2260 ± 0.357).

The mutational landscape of HPV-positive and HPV-negative tumours was distinct. In HPV-negative carcinomas, mutations were detected mainly in *TP53* (73.3%), *FAT1* (26.7%) and *FLG* (16.7%) whereas in HPV-positive, the common mutations were in *FLG (24.3%) FAT1* (17%) and *FGFR3* (14.6%) genes. A majority of *TP53* mutations were missense (66.7%) followed by nonsense (17.7%) mutations. Frameshift indels (9%) were also frequently observed in the *TP53* gene. Interestingly, 20% (9/44) of the HPV-negative *TP53* mutated tumours had more than one pathogenic *TP53* mutation. Nearly 82% of *FAT1* mutations were missense and only 8% were nonsense, rest being frameshift indels. Inactivating mutations in *FAT1* (n=3) and *CDKN2A* (n=3) were exclusively found in HPV-negative tumours. Contrary to this, genetic alterations in fibroblasts growth factor receptor 3 (*FGFR3*), a potent receptor-tyrosine kinases (RTK) involved in growth and proliferation of cancer cells, were significantly higher in HPV-positive *vs.* HPV-negative tumours (13.95% *vs* 1.73%, p=0.017). All *FGFR3* mutations were missense mutations. Notably the mutations in genes encoding histone methyltransferases, *NSD1* (5.8%) and *KMT2C* (5.8%) were reported, with majority of them being missense. Commonly known mutations in *RAS* (3.9%, combined for *HRAS, NRAS* and *KRAS*), *CASP8* (3.8%), *SMAD4* (3.8%), *NOTCH2* (3.8%), and *TRAF3* (3.8%) were also registered (**Figure 2B**). Focussing only on OPSCC, most mutations in HPV-positive OPSCC were *FLG* (22%) followed by *FAT1* (17.1%), *FGFR3* (14.6%) and *NSD1* (9.8%). Notable mutations in *CASP8* and *KMT2D* (both 7.3%) were also found (**Figure 2B**).

Mutational Signature of HNSCC

SNVs identified in the HNSCC cohort were further evaluated for mutational signatures with respect to presence of HPV and smoking status. The most common class of variant were missense mutations in both HPV-negative and HPV-positive tumours (**Figure 3A**). Also, the frequency of single nucleotide polymorphisms (SNPs) was greater than insertions and deletions both in HPV-positive and HPV-negative cases (**Figure 3B**). Interestingly, more than 21% (13/60) of HPV-negative patients had two mutations in the same gene as compared to only 7.3% (3/41) patients in HPV-positive subtype (**Figure 3C**). As reported before (30), C>T transitions were the most common SNV type in both HPV-negative and HPV-positive tumours, with the frequency being greater in the HPV-positive tumours (**Figures 3D, E**).

Similar observations were made in smokers and non-smokers subtypes with missense mutations being the most common class of variant (**Figure 4A**). Also, the frequency of SNVs was greater than other types of variants in smokers (**Figure 4B**). Surprisingly, the percentage of patients having two mutations in the same gene was more in non-smokers as compared to smokers (**Figure 4C**). Based on the mutational signature, smokers had higher frequency of C>T transitions among different SNVs, as compared to non-smokers (**Figures 4D, E**).

Potentially targetable mutations in *FGFR3* were identified. Well established canonical missense mutations in R248C and S249C were identified in five patients (4.8%) (**Figure 5**). Surprisingly, all five patients were HPV-positive oropharyngeal SCC with lymph node metastasis. These mutations are used as predictive biomarkers for the use of erdafitinib in patients with locally advanced or metastatic urothelial and bladder carcinomas

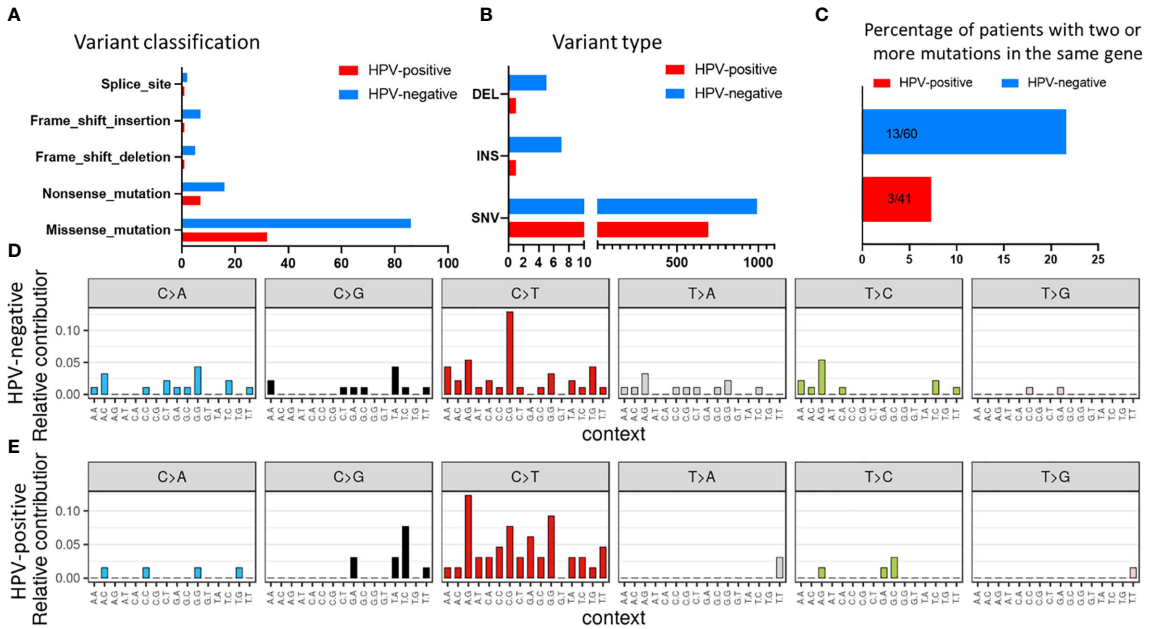


FIGURE 3 | Mutational signature in the HNSCC samples stratified with respect to HPV status for **(A)** variant classification and **(B)** variant type. **(C)** Comparison of number of patients with two or more mutations in the same gene. Mutational profile based on C:G>T:A transitions in **(D)** HPV-negative and **(E)** HPV-positive samples.

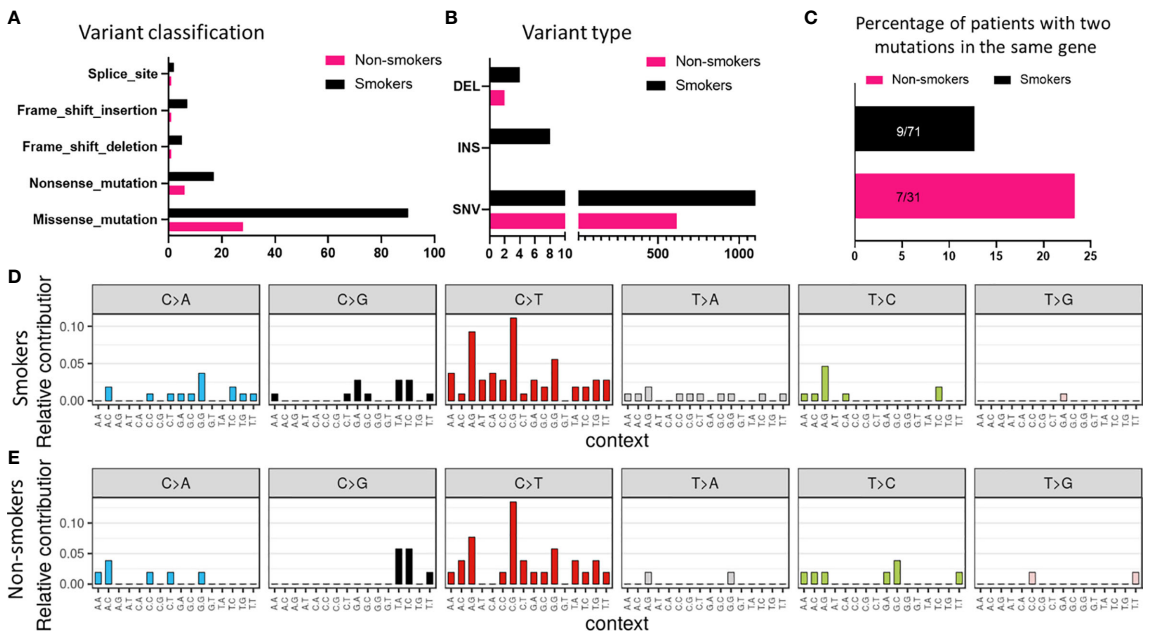


FIGURE 4 | Mutational signature associated with smokers and non-smokers based on **(A)** variant classification, and **(B)** variant type. **(C)** Comparison of number of patients with two or more mutations in the same gene based on smoking status. Mutational profile based on C:G>T:A transitions in **(D)** smokers and **(E)** non-smokers.

(NCT02365597). The mutations have been proposed to activate the FGF-receptor *via* ligand-independent dimerization by generating novel cysteine residues (31, 32). Further, different SNVs in proteins coding for *FAT1* and *NSD1* genes are also shown in **Figure 5**. However, the distribution is of *FAT1* and *NSD1* is heterogenous and does not point to any specific pattern.

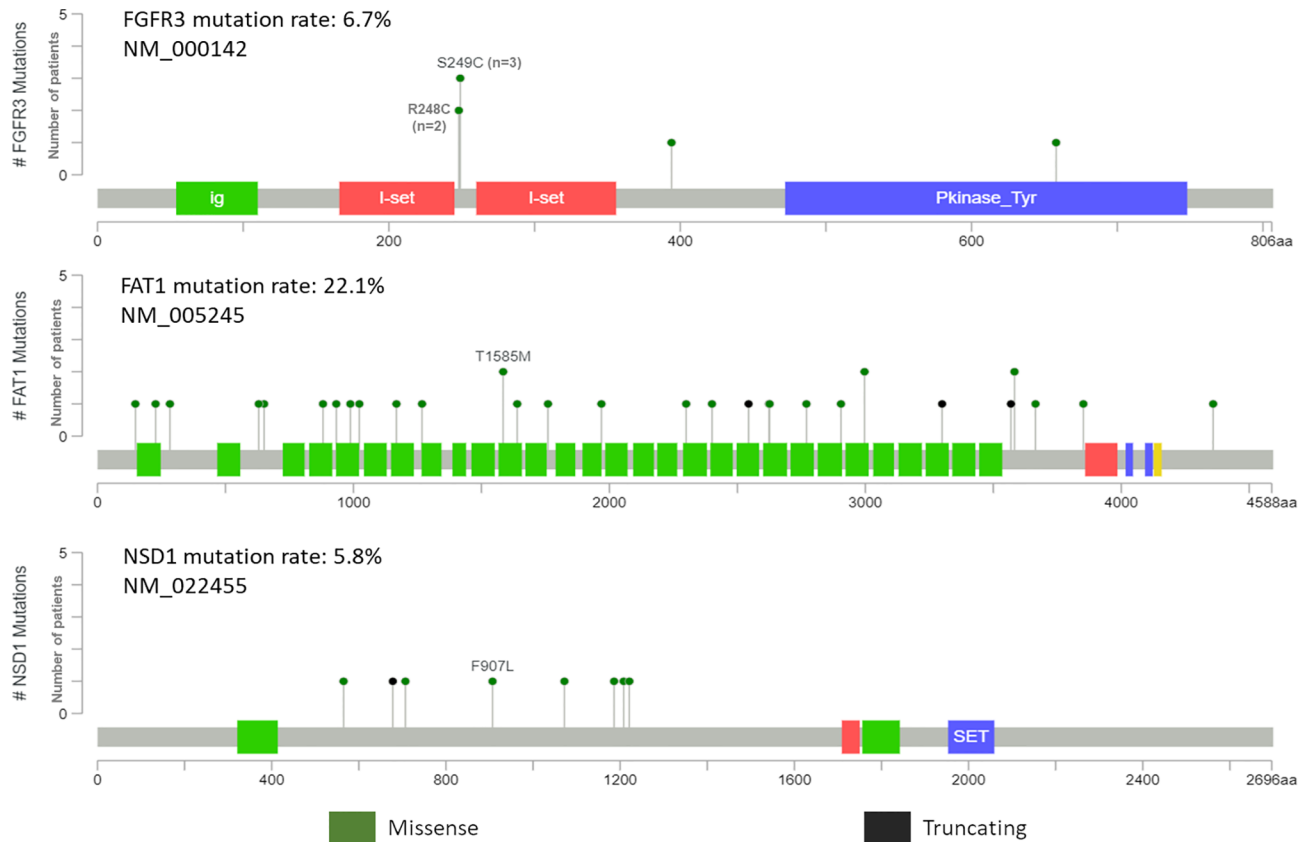


FIGURE 5 | Pictorial representation of single nucleotide variant distribution on domain structure of *FGFR3, FAT1* and *NSD1*.

Correlations Between Presence of Cancer-Specific Mutations and Clinico-Pathological Parameters

A higher number of cancer-specific mutations were associated with a negative HPV status (p<0.001, r=0.448). No association was observed between the number of SNPs per Mb detected and HPV status or any other clinic-pathological parameters tested. For the whole cohort, presence of at least one cancer-specific mutation was found to be positively associated with an aggressive type of invasive front (p=0.035, r=0.210), and extensive desmoplastic stroma (p=0.019, r=0.234), and negatively associated with the degree of differentiation (p=0.041, r=-203) (**Figure 6**).

Validation of Correlations Between Mutational Burden and Clinico-Pathological Parameters in TCGA Dataset

The Chi Square analysis between mutational score reflecting the presence of at least one mutation in any of the seven most frequently cancer-specific mutated genes identified in our panel (*TP53, FAT1, FGF3, FLG, CDKN2A, KMT2C*, and *NSD1*) and clinicopathological variables available in the TCGA data set for HNSCC showed that the mutation score was significantly associated negatively with differentiation (p<0.001), and positively with perineural invasion (p=0.010), and clinical T-stage (p<0.001). The survival analysis showed that tumours with at least one mutation in one of these genes had shorter 5-year disease-free and overall survival (p=0.005 for both, Log Rank) (**Figure 7**). Further, based on HPV-status stratification, it can be observed that the seven gene panel predicts better disease-free and overall survival in HPV-negative HNSCC (**Figure 7**).

DISCUSSION

Formalin-fixed paraffin-embedded (FFPE) tumour tissue samples are less optimal for NGS based applications than fresh tissue samples. A major problem is cytosine deamination to uracil, smaller insert sizes, higher duplication rates and an increased frequency of falsely mapped SNVs. Still, Sweiger et al. could show successful low coverage sequencing and copy number detection in FFPE tissue samples stored for more than 18 years (33). Significant improvements were made by Hedegaard and co-workers to improve the output of NGS whole exome sequencing in aged FFPE tissue material, however the successful sequencing was obtained in only 29,5%

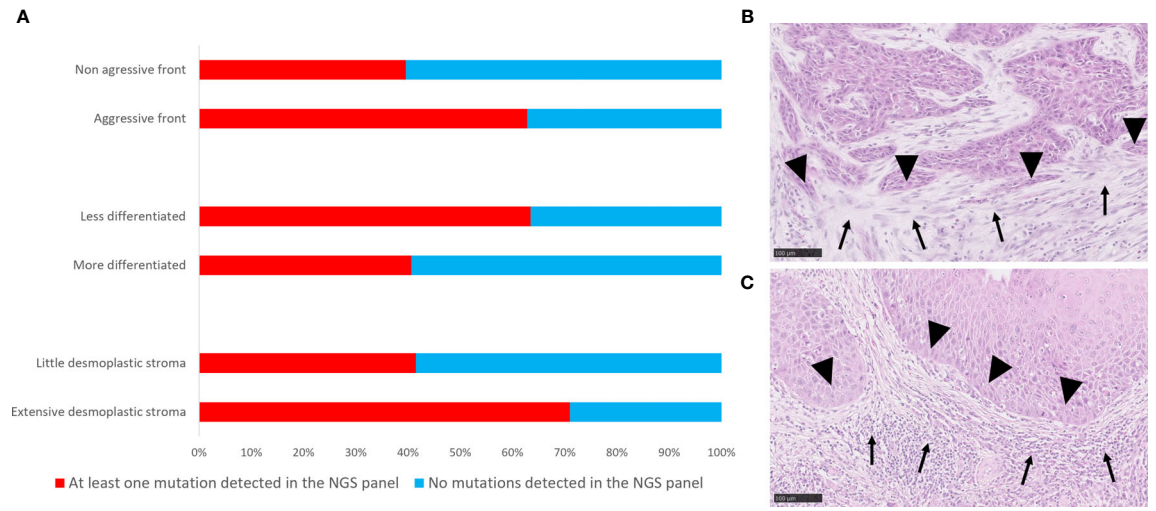


FIGURE 6 | Correlations of mutational landscape with histopathological parameters. **(A)** Bar graph showing distribution of cases with at least one mutation and without mutations in the genes targeted by the NGS panel according to histopathological parameters. **(B)** Histological picture of a tumour with at least one mutation detected by targeted NGS panel, and an extensive stromal desmoplastic response (solid arrows), an aggressive invasive tumour front (solid arrowheads), a less differentiated phenotype and a poor inflammatory response. **(C)** Histological representation of a tumour with no mutations detected by targeted NGS panel and a little stromal desmoplastic response, non-aggressive invasive tumour front (solid arrowheads), a more differentiated phenotype and an intense inflammatory response (solid arrows). Scale bar = 100 μm.

of the samples (34). They noted in particular an increase in false positive SNVs in FFPE tissue after prolonged storage for more than 3-4 years (34). A similar observation has also been made by others (35). On the other hand, Kerick and co-workers have shown that smaller targeted NGS panels using PCR amplification and an increased coverage will reduce the impact of FFPE induced false positive SNVs in the work up of NGS analysis (36). Our study has used long-term preserved (range 4-17 years) FFPE tumour tissue samples for the NGS analysis. To reduce the impact of fixation, the NGS library was re-amplified before paired-end sequencing on the Illumina MiniSeq platform based on manufacturer's suggestion to improve number of reads per sample. Also, a robust cut-off of 500 reads for total allelic depth was used to minimize the risk of false positive SNVs due to aged FFPE tumour tissue in our study. We have tested all mutations/annotations involving C-T and A-G conversions in the period 2003-2010 and 2013-2016 without finding significant differences between these time periods. With these limitations kept in mind, the study presents data on the application of a custom-made, HNSCC targeted NGS panel using FFPE material preserved for up to 17 years in a routine diagnostic pathology service. We used the targeted NGS panel to characterize the mutational landscape of a cohort of HNSCC patients with tissues preserved at the Diagnostics Biobank at Department of Pathology, Haukeland University Hospital. Our results are in line with previous studies that found higher TMB in HPV-negative cancers compared to HPV-positive HNSCC (14, 16).

In addition, the panel of genes found most frequently mutated in our study is comparable to the findings from previous studies performed using whole exome sequencing (WES) (14–16). Due to strict cut-off of 500 reads, similar to what is used in diagnostics at HUS, several well reported mutations in the PIK3CA/PTEN pathways were excluded from the final analysis. However, from the remaining detected genes, of interest is our finding of frequent mutations in the gene encoding Nuclear Set Domain Containing Protein 1 (*NSD1*), a histone methyltransferase, in both HPV-positive and negative HNSCC. *NSD1* has been identified as a biomarker for global epigenetic changes in cancer (37, 38) and based on TCGA data, it was recently suggested as a prognostic biomarker in patients with HPV-negative HNSCC (39). Using TCGA data, Bui et al. identified a survival advantage for patients with mutations in the *NSD1* gene. This gene was found altered in approximately 10% of patients with HNSCC, and they proposed that patients with loss-of-function *NSD1* mutations should be considered a distinct clinical subclass of HPV-negative HNSCC, with increased cisplatin sensitivity (39). In addition, we found mutations in *FGFR3* gene, particularly at R248C and S249C that could be used as a targeted therapy especially in HPV-positive HNSCC. Combined together, such findings are usually the result of WES studies and are not prone to be further implemented in the clinical oncology. However, WES is not available at most hospitals owing to its high cost, operational complexity, and long turnover times. On the other hand, lower-cost targeted NGS for cancer specific genes are increasingly affordable and finding its application in clinical oncology. We show here that using a targeted NGS panel is an effective and more affordable way for detecting actionable mutations in archival material preserved for more than 17 years. Moreover, novel mutations unearthed in *FGFR3* in HPV-positive OPSCC justifies the use of NGS based gene panel and suggest valid targets in personalized treatment.

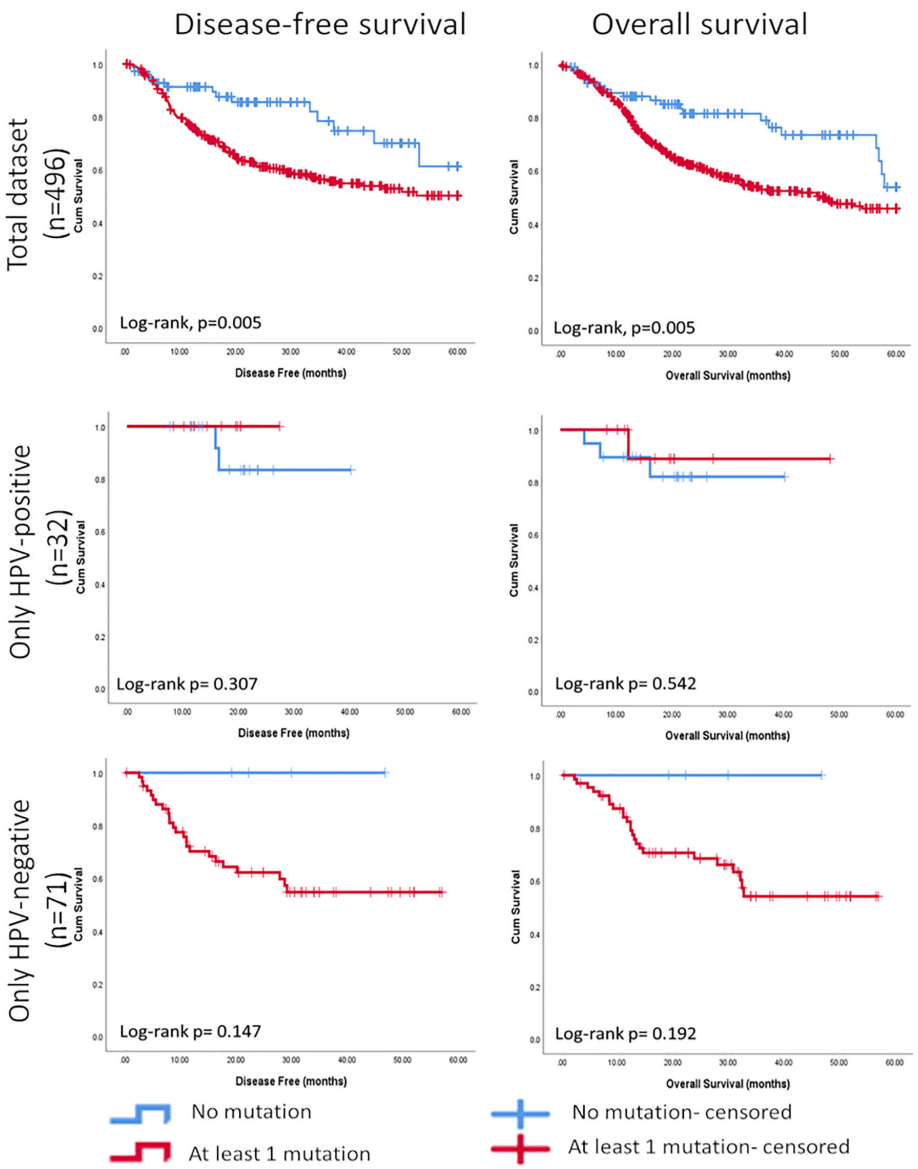


FIGURE 7 | Kaplan-Meier curves showing significant associations between the mutational score based on seven cancer-specific genes found most frequently mutated in HNSCC and further stratified based on HPV-status using the targeted NGS panel and disease-free survival (5 years) and overall survival (5 years).

We further wanted to investigate possible association between the presence of mutations in any of the genes investigated and clinical and histopathological parameters. Histopathological assessment of FFPE tissue and surgical resections still remains the cornerstone of diagnosis and are readily available to a pathologist with simple chemical stains. Not only this, but it is well known fact that DNA mutations and their impact on cellular mRNA and protein expression are the main drivers of changes that are evident in the histopathological assessment. As reported in this study, the presence of at least one cancer-specific mutation among the seven most altered genes was associated with less differentiation, an aggressive type of invasive front and a extensive desmoplastic stroma. To our knowledge, this is the first attempt to study and identify associations between mutational load and its stromal host response. In the light of the recent development of immunotherapy, mutational landscape has been previously investigated as a predictor of immune response, but the stromal response has not been investigated in this respect. TMB has been reported as a biomarker for predicting response to immunotherapy in cancer patients (40, 41). This seems to be explained by the thought that tumours with high TMB express a greater diversity of neoantigens, resulting in increased immune recognition when immune checkpoint inhibitors release natural brakes on the immune system. Of note, we did not identify any

correlations between TMB or presence of mutations in the investigated genes and presence of a heavier immune infiltrate, but correlations with a more aggressive type of invasion and fibroblastic stromal response were found.

Looking more into the histopathological variables, we have revealed a correlation between presence of at least one cancer specific-mutation and a more desmoplastic response and an aggressive type of invasion at the tumour front but less inflammatory infiltrate. With the new knowledge on the role of cancer associated fibroblasts (CAF) on inducing an immune suppressive phenotype in HNSCC (42) our findings are not surprising. Takahashi et al. found CAFs from HNSCC to express higher B7H1 (PDL1) and B7DC (PDL2) than normal fibroblasts, both putative negative regulators of immune function. Functionally, they were able to suppress T cell proliferation, induce T cell apoptosis and recruit Treg (CD4+Foxp3+) cells. Subpopulations of CAFs with immunosuppressive functions have been recently described in other cancer tissues as well (43). In this respect it is also worth mentioning that our analysis revealed that the HPV-positive cases presented with a non-aggressive tumour invasive front, little desmoplastic stromal reaction, but rich inflammatory host reaction than the HPV-negative cases. This indicates a differential host response to HPV-positive cancers compared to HPV-negative.

The findings of this study were validated on an independent data set. Using TCGA data available for HNSCC, we found that the presence of mutations in at least one of the seven most frequently mutated genes identified by the NGS panel was significantly associated with less differentiation, perineural invasion, and clinical T stage. In addition, we could also perform survival analysis on TCGA data since these data were not available for our HNSCC cohort. The survival analysis showed that tumours with at least one mutation in one of these genes had a shorter 5-year disease-free and overall survival, indicating a clinical relevance for the panel of genes identified by the targeted NGS panel.

CONCLUSIONS

A custom made targeted NGS panel could reliably detect several specific mutations in archival samples of HNSCCs within 17 years of preservation. Using this method, novel associations between mutational burden and clinical and pathological parameters were detected and actionable mutation in HPV-positive HNSCC were discovered. A specific focus was on NGS data obtained on preserved FFPE material, especially with respect to number of PCR amplified reads with a strict cut-off, to uncover more novel and translational mutations.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Regional Committee for Medical and Health Research Ethics (REK Vest, 2011/125). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HD, HH, DC, HA, and OV have conceived the study. FE, SM, and HA recruited the patients and collected clinical data. HH and OV performed tissue analysis. SF, SD and RR performed the NGS experiments. HD, HH, DS, OV, and DC analysed data. HD and DC prepared the figures and the manuscript. HA, OV, and DC were supervising the work. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by The Western Norway Regional Health Authority (Helse Vest Grant No. 911902/2013 and 912260/2019), The Research Council of Norway through its Centres of Excellence funding scheme, (Grant No. 22325), and The Norwegian Centre for International Cooperation in Education (project number CPEA-LT-2016/10106).

ACKNOWLEDGMENTS

The authors would like to acknowledge all the patients involved in this study and their families.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.734134/full#supplementary-material

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
2. Lawrence MS, Sougnez C, Lichtenstein L, Cibulskis K, Lander E, et al. Comprehensive Genomic Characterization of Head and Neck Squamous Cell Carcinomas. *Nature* (2015) 517(7536):576–82. doi: 10.1038/nature14129
3. Leemans CR, Braakhuis BJM, Brakenhoff RH. The Molecular Biology of Head and Neck Cancer. *Nat Rev Cancer* (2011) 11(1):9–22. doi: 10.1038/nrc2982
4. Ljøkjel B, Haave H, Lybak S, Aarstad HH, Karlsdottir À, Vintermyr OK, et al. The Impact of HPV Infection, Smoking History, Age and Operability of the Patient on Disease-Specific Survival in a Geographically Defined Cohort of

Patients With Oropharyngeal Squamous Cell Carcinoma. *Acta Oto-Laryngol* (2014) 134(9):964–73. doi: 10.3109/00016489.2014.927590
5. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, et al. Human Papillomavirus and Survival of Patients With Oropharyngeal Cancer. *N Engl J Med* (2010) 363(1):24–35. doi: 10.1056/NEJMoa0912217
6. Alsahafi E, Begg K, Amelio I, Raulf N, Lucarelli P, Sauter T, et al. Clinical Update on Head and Neck Cancer: Molecular Biology and Ongoing Challenges. *Cell Death Dis* (2019) 10(8):540. doi: 10.1038/s41419-019-1769-9
7. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N Engl J Med* (2017) 377(25):2500–1. doi: 10.1056/NEJMc1713444
8. Ferris RL. Immunology and Immunotherapy of Head and Neck Cancer. *J Clin Oncol* (2015) 33(29):3293–304. doi: 10.1200/jco.2015.61.1509
9. Vermorken JB, Trigo J, Hitt R, Koralewski P, Diaz-Rubio E, Rolland F, et al. Open-Label, Uncontrolled, Multicenter Phase II Study to Evaluate the Efficacy and Toxicity of Cetuximab as a Single Agent in Patients With Recurrent and/or Metastatic Squamous Cell Carcinoma of the Head and Neck Who Failed to Respond to Platinum-Based Therapy. *J Clin Oncol* (2007) 25(16):2171–7. doi: 10.1200/jco.2006.06.7447
10. De Felice F, Musio D, Tombolini V. Immune Check-Point Inhibitors and Standard Chemoradiotherapy in Definitive Head and Neck Cancer Treatment. *J Personalized Med* (2021) 11(5):393. doi: 10.3390/jpm11050393
11. Chow LQM, Haddad R, Gupta S, Mahipal A, Mehra R, Tahara M, et al. Antitumor Activity of Pembrolizumab in Biomarker-Unselected Patients With Recurrent and/or Metastatic Head and Neck Squamous Cell Carcinoma: Results From the Phase Ib KEYNOTE-012 Expansion Cohort. *J Clin Oncol* (2016) 34(32):3838–45. doi: 10.1200/jco.2016.68.1478
12. Sacco AG, Worden FP. Molecularly Targeted Therapy for the Treatment of Head and Neck Cancer: A Review of the ErbB Family Inhibitors. *Oncol Targets Ther* (2016) 9:1927–43. doi: 10.2147/ott.S93720
13. Wang Z, Goto Y, Allevato MM, Wu VH, Saddawi-Konefka R, Gilardi M, et al. Disruption of the HER3-PI3K-mTOR Oncogenic Signaling Axis and PD-1 Blockade as a Multimodal Precision Immunotherapy in Head and Neck Cancer. *Nat Commun* (2021) 12(1):2383. doi: 10.1038/s41467-021-22619-w
14. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. *Science* (2011) 333(6046):1157–60. doi: 10.1126/science.1208130
15. Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, et al. Exome Sequencing of Head and Neck Squamous Cell Carcinoma Reveals Inactivating Mutations in NOTCH1. *Science* (2011) 333(6046):1154–7. doi: 10.1126/science.1206923
16. Maitra A, Biswas NK, Amin K, Kowtal P, Kumar S, Das S, et al. Mutational Landscape of Gingivo-Buccal Oral Squamous Cell Carcinoma Reveals New Recurrently-Mutated Genes and Molecular Subgroups. *Nat Commun* (2013) 4 (1):2873. doi: 10.1038/ncomms3873
17. Haave H, Gulati S, Brekke J, Lybak S, Vintermyr OK, Aarstad HJ, et al. Tumor Stromal Desmoplasia and Inflammatory Response Uniquely Predict Survival With and Without Stratification for HPV Tumor Infection in OPSCC Patients. *Acta Otolaryngol* (2018) 138(11):1035–42. doi: 10.1080/00016489.2018.1497809
18. Bryne M, Jenssen N, Boysen M. Histological Grading in the Deep Invasive Front of T1 and T2 Glottic Squamous Cell Carcinomas has High Prognostic Value. *Virchows Archiv* (1995) 427(3):277–81. doi: 10.1007/BF00203395
19. Brandwein-Gensler M, Teixeira MS, Lewis CM, Lee B, Rolnitzky L, Hille JJ, et al. Oral Squamous Cell Carcinoma: Histologic Risk Assessment, But Not Margin Status, Is Strongly Predictive of Local Disease-Free and Overall Survival. *Am J Surg Pathol* (2005) 29(2):167–78. doi: 10.1097/01.pas.0000149687.90710.21
20. Ojesina AI, Lichtenstein L, Freeman SS, Pedamallu CS, Imaz-Rosshandler I, Pugh TJ, et al. Landscape of Genomic Alterations in Cervical Carcinomas. *Nature* (2014) 506(7488):371–5. doi: 10.1038/nature12881
21. Lybak S, Ljøkjel B, Haave H, Karlsdottir À., Vintermyr OK, Aarstad HJ, et al. Primary Surgery Results in No Survival Benefit Compared to Primary Radiation for Oropharyngeal Cancer Patients Stratified by High-Risk Human Papilloma Virus Status. *Eur Arch Otorhinolaryngol* (2017) 274 (1):477–87. doi: 10.1007/s00405-016-4203-2
22. Begum S, Gillison ML, Ansari-Lari MA, Shah K, Westra WH. Detection of Human Papillomavirus in Cervical Lymph Nodes: A Highly Effective Strategy for Localizing Site of Tumor Origin. *Clin Cancer Res* (2003) 9 (17):6469–75.
23. Lechner M, Frampton GM, Fenton T, Feber A, Palmer G, Jay A, et al. Targeted Next-Generation Sequencing of Head and Neck Squamous Cell Carcinoma Identifies Novel Genetic Alterations in HPV+ and HPV- Tumors. *Genome Med* (2013) 5(5):49. doi: 10.1186/gm453
24. Seiwert TY, Zuo Z, Keck MK, Khattri A, Pedamallu CS, Stricker T, et al. Integrative and Comparative Genomic Analysis of HPV-Positive and HPV-Negative Head and Neck Squamous Cell Carcinomas. *Clin Cancer Res* (2015) 21(3):632–41. doi: 10.1158/1078-0432.Ccr-13-3310
25. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet* (2017) 100(2):267–80. doi: 10.1016/j.ajhg.2017.01.004
26. Phan L, Jin Y, Zhang H, Qiang W, Shekhtman E, Shao D, et al. ALFA: Allele Frequency Aggregator (National Center for Biotechnology Information, US National Library of Medicine). *In* (2020).
27. Díaz-Gay M, Vila-Casadesús M, Franch-Expósito S, Hernández-Illán E, Lozano JJ, Castellví-Bel S, et al. Mutational Signatures in Cancer (MuSiCa): A Web Application to Implement Mutational Signatures Analysis in Cancer Samples. *BMC Bioinf* (2018) 19(1):224. doi: 10.1186/s12859-018-2234-y
28. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The Cbio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov* (2012) 2(5):401–4. doi: 10.1158/2159-8290.Cd-12-0095
29. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the Cbioportal. *Sci Signal* (2013) 6(269):pl1. doi: 10.1126/scisignal.2004088
30. Patel K, Bhat FA, Patil S, Routray S, Mohanty N, Nair B, et al. Whole-Exome Sequencing Analysis of Oral Squamous Cell Carcinoma Delineated by Tobacco Usage Habits. *Front Oncol* (2021) 11:660696(1962). doi: 10.3389/fonc.2021.660696
31. di Martino E, L'Hôte CG, Kennedy W, Tomlinson DC, Knowles MA. Mutant Fibroblast Growth Factor Receptor 3 Induces Intracellular Signaling and Cellular Transformation in a Cell Type- and Mutation-Specific Manner. *Oncogene* (2009) 28(48):4306–16. doi: 10.1038/onc.2009.280
32. d'Avis PY, Robertson SC, Meyer AN, Bardwell WM, Webster MK, Donoghue DJ, et al. Constitutive Activation of Fibroblast Growth Factor Receptor 3 by Mutations Responsible for the Lethal Skeletal Dysplasia Thanatophoric Dysplasia Type I. *Cell Growth Differ* (1998) 9(1):71–8.
33. Schweiger MR, Kerick M, Timmermann B, Albrecht MW, Borodina T, Parkhomchuk D, et al. Genome-Wide Massively Parallel Sequencing of Formaldehyde Fixed-Paraffin Embedded (FFPE) Tumor Tissues for Copy-Number- and Mutation-Analysis. *PloS One* (2009) 4(5):e5548. doi: 10.1371/journal.pone.0005548
34. Hedegaard J, Thorsen K, Lund MK, Hein AM, Hamilton-Dutoit SJ, Vang S, et al. Next-Generation Sequencing of RNA and DNA Isolated From Paired Fresh-Frozen and Formalin-Fixed Paraffin-Embedded Samples of Human Cancer and Normal Tissue. *PloS One* (2014) 9(5):e98187. doi: 10.1371/journal.pone.0098187
35. Fujii T, Uchiyama T, Matsuoka M, Myojin T, Sugimoto S, Nitta Y, et al. Evaluation of DNA and RNA Quality From Archival Formalin-Fixed Paraffin-Embedded Tissue for Next-Generation Sequencing – Retrospective Study in Japanese Single Institution. *Pathol Int* (2020) 70(9):602–11. doi: 10.1111/pin.12969
36. Kerick M, Isau M, Timmermann B, Sültmann H, Herwig R, Krobitsch S, et al. Targeted High Throughput Sequencing in Clinical Cancer Settings: Formaldehyde Fixed-Paraffin Embedded (FFPE) Tumor Tissues, Input Amount and Tumor Heterogeneity. *BMC Med Genomics* (2011) 4, 68. doi: 10.1186/1755-8794-4-68
37. Lee ST, Wiemels JL. Genome-Wide CpG Island Methylation and Intergenic Demethylation Propensities Vary Among Different Tumor Sites. *Nucleic Acids Res* (2016) 44(3):1105–17. doi: 10.1093/nar/gkv1038
38. Papillon-Cavanagh S, Lu C, Gayden T, Mikael LG, Bechet D, Karamboulas C, et al. Impaired H3K36 Methylation Defines a Subset of Head and Neck Squamous Cell Carcinomas. *Nat Genet* (2017) 49(2):180–5. doi: 10.1038/ng.3757
39. Bui N, Huang JK, Bojorquez-Gomez A, Licon K, Sanchez KS, Tang SN, et al. Disruption of NSD1 in Head and Neck Cancer Promotes Favorable

Chemotherapeutic Responses Linked to Hypomethylation. *Mol Cancer Ther* (2018) 17(7):1585–94. doi: 10.1158/1535-7163.Mct-17-0937
40. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of Tumor Mutation Burden as an Immunotherapy Biomarker: Utility for the Oncology Clinic. *Ann Oncol* (2019) 30(1):44–56. doi: 10.1093/annonc/mdy495
41. Gandara DR, Paul SM, Kowanetz M, Schleifman E, Zou W, Li Y, et al. Blood-Based Tumor Mutational Burden as a Predictor of Clinical Benefit in Non-Small-Cell Lung Cancer Patients Treated With Atezolizumab. *Nat Med* (2018) 24(9):1441–8. doi: 10.1038/s41591-018-0134-3
42. Takahashi H, Sakakura K, Kawabata-Iwakawa R, Rokudai S, Toyoda M, Nishiyama M, et al. Immunosuppressive Activity of Cancer-Associated Fibroblasts in Head and Neck Squamous Cell Carcinoma. *Cancer Immunol Immunother* (2015) 64(11):1407–17. doi: 10.1007/s00262-015-1742-0
43. Kieffer Y, Hocine HR, Gentric G, Pelon F, Bernard C, Bourachot B, et al. Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discov* (2020) 10(9):1330–51. doi: 10.1158/2159-8290.Cd-19-1384

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 07 October 2021
doi: 10.3389/fonc.2021.750794

Integrative Analysis of Metabolomics and Transcriptomics Data Identifies Prognostic Biomarkers Associated With Oral Squamous Cell Carcinoma

Lihua Zuo[1†], Zhuo Chen[2†], Lihuang Chen[3], Jian Kang[1], Yingying Shi[1], Liwei Liu[1], Shuhua Zhang[4], Qingquan Jia[1], Yi Huang[5] and Zhi Sun[1]*

[1] Department of Pharmacy, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [2] Department of Oral and Maxillofacial Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [3] School and Hospital of Stomatology, Weifang Medical University, Weifang, China, [4] Clinical Laboratory, Chongqing Southeast Hospital, Chongqing, China, [5] Research and Development Department, Chongqing Huangjia Biotechnology Limited Company, Chongqing, China

OPEN ACCESS

Edited by:
Die Wang,
Hudson Institute of Medical Research, Australia

Reviewed by:
Le Ying,
Hudson Institute of Medical Research, Australia
Qiqi Xie,
Affiliated Hospital of Qinghai University, China
Wei Cao,
Shanghai Jiao Tong University, China

***Correspondence:**
Zhi Sun
sunzhi2013@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *31 July 2021*
Accepted: *08 September 2021*
Published: *07 October 2021*

Citation:
Zuo L, Chen Z, Chen L, Kang J, Shi Y, Liu L, Zhang S, Jia Q, Huang Y and Sun Z (2021) Integrative Analysis of Metabolomics and Transcriptomics Data Identifies Prognostic Biomarkers Associated With Oral Squamous Cell Carcinoma. Front. Oncol. 11:750794. doi: 10.3389/fonc.2021.750794

Background: Oral squamous cell carcinoma (OSCC) is the most malignant neoplasm in oral cancer. There is growing evidence that its progression involves altered metabolism. The current method of evaluating prognosis is very limited, and metabolomics may provide a new approach for quantitative evaluation. The aim of the study is to evaluate the use of metabolomics as prognostic markers for patients with OSCC.

Methods: An analytical platform, Ultra-Performance Liquid Chromatography-Quadrupole/Orbitrap High Resolution Mass Spectrometry (UHPLC-Q-Orbitrap HRMS), was used to acquire the serum fingerprinting profiles from a total of 103 patients of OSCC before and after the operation. In total, 103 OSCC patients were assigned to either a training set (n = 73) or a test set (n = 30). The potential biomarkers and the changes of serum metabolites were profiled and correlated with the clinicopathological parameters and survival of the patients by statistical analysis. To further verify our results, we linked them to gene expression using data from the Kyoto Encyclopedia of Genes and Genomes (KEGG).

Results: In total, 14 differential metabolites and five disturbed pathways were identified between the preoperative group and postoperative group. Succinic acid change-low, hypoxanthine change-high tumor grade, and tumor stage indicated a trend towards improved recurrence-free survival (RFS), whether in a training set or a test set. In addition, succinic acid change-low, hypoxanthine change-high, and tumor grade provided the highest predictive accuracy of the patients with OSCC. KEGG enrichment analysis showed that the imbalance in the amino acid and purine metabolic pathway may affect the prognosis of OSCC.

Conclusions: The changes of metabolites before and after operation may be related to the prognosis of OSCC patients. UHPLC-Q-Orbitrap HRMS serum metabolomics analysis could be used to further stratify the prognosis of patients with OSCC. These

results can better understand the mechanisms related to early recurrence and help develop more effective therapeutic targets.

Keywords: OSCC, metabolomics, prognostic biomarkers, succinic acid, hypoxanthine

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is the most common malignant neoplasm in oral cancer, and patients with the carcinoma had a low 5-year survival rate and poor prognosis (1–3). Although the treatment of OSCC has improved, including surgery, radiotherapy, chemotherapy, and immunotherapy (1), the current worldwide average 5-year overall survival (OS) rate is only 65% (4). Over the past decade, increasing evidence has implicated altered metabolic homeostasis to be dysregulated with OSCC malignant progression (5). The most striking feature of cancer cells is that they rewire their metabolism and nutrient acquisition patterns to meet cancer cell energy needs, and oncogene signaling pathways and OSCC metabolic activity established a strong link. Cell metabolic phenotypes can be used to predict patients' outcomes (6). Therefore, given the feature of cancer cells, identification of more sensitive prognostic biomarkers and novel therapeutic targets are important targets of research in OSCC.

The molecular pathogenesis of OSCC is complex, which is the result of the interaction of several molecular networks (5). It involves not only the changes of specific gene and protein expressions but also a change inmetabolic processes (7). Metabolomics is an important branch of omics science, which is used to evaluate the changes of metabolites in biological samples (8, 9). Recently, a large number of metabolomics studies have focused on the exploration of disease mechanisms, the identification of potential biomarkers, the prediction of cancer prognosis, and the evaluation of treatment effect (10, 11). Most metabolomics studies of OSCC are mainly based on the metabolic profiles of saliva, serum, and tumor tissues to identify potential biomarkers for screening and early diagnosis (12). Fu et al. detected 25 amino acids in OSCC tissue by targeted metabolomics technology and proved that L-asparagine metabolism disorder mediated by asparagine synthase promoted the perineural invasion of OSCC (13). In addition, Yang et al. showed that L-glutamate, L-aspartic acid, and L-proline were identified as a group of potential diagnostic biomarkers of OSCC (14). This research group also proved that amino acid signatures are also different at different distances from the surgical margins of OSCC, which provides a new idea for determining the intraoperative safety boundary (15). However, there is no prognostic study of OSCC based on metabolomics.

In this study, ultra-high-performance liquid chromatography–quadrupole/Orbitrap high-resolution mass spectrometry (UHPLC-Q-Orbitrap HRMS) was used to acquire the serum fingerprinting profiles from a total of 103 patients of OSCC before and after the operation. The serum fingerprint information of 73 OSCC patients was used as the training set to find metabolites related to prognosis, and the serum fingerprint information of the remaining 30 OSCC patients was used as the test set to confirm the results of the training set. Transcriptome data from the Kyoto Encyclopedia of Genes and Genomes (KEGG) are also used to detect gene expression levels and find key genes and pathways related to diseases. The study was designed to uncover transcription programs, driving the observed metabolic phenotype, and a framework for future studies designed to determine how specific metabolic programs may influence the prognosis of OSCC.

MATERIAL AND METHODS

Reagents and Chemicals

The HPLC-grade methanol and acetonitrile were acquired from Fisher Scientific (Fair Lawn, NJ, USA). HPLC-grade formic acid was purchased from Aladdin Industrial Co., Ltd. (Shanghai, China). HPLC-grade water was obtained by the Millipore system (Shanghai, China). The internal standards and all the endogenous metabolite standards were acquired from J&K Scientific Ltd. (Beijing, China) and Sigma-Aldrich (St Louis, MO, USA).

Study Design and Participant

This cross-sectional study recruited 103 patients with OSCC from The First Affiliated Hospital of Zhengzhou University who were diagnosed for the first time by oral clinicians based on the clinical criteria and postoperative pathology report (16). We excluded the patients with substance abuse, viral hepatitis, severe nephropathy, malignancies, metabolic diseases, and long-term use of estrogens, tamoxifen, or corticosteroids. In all these patients, preoperative radiotherapy or chemotherapy has not been administrated. The ethical approval for this study was obtained from the Ethical Committees of The First Affiliated Hospital of Zhengzhou University (name of IRB: Ethics Committee of Scientific Research Project of The First Affiliated Hospital of Zhengzhou University; ethical code: SB201902006). This research was conducted in accordance with the ethical guidelines of the 1975 Declaration of Helsinki.

Treatment and Follow-Up

In the first half of the year after surgery, patients were followed up every 1 month and then every 3 months until May 2021, disease progression, death, or follow-up failure. The study was conducted at each scheduled time by patient follow-up or telephone follow-up. Progression-free survival (PFS) was selected as the endpoint and defined as the time interval from surgery to local or distant recurrence and/or metastasis,

whichever occurred first. If the patient died, was lost to follow-up, or did not progress at the end of the study, the survival time was considered censored.

Sample Collection and Preparation

At patients' initial visit and 7 days after the operation, the venous blood of each volunteer patient was collected in the morning after overnight fasting. The blood was put into polypropylene tubes containing coagulant and cooled down in an insulated box with ice. The fresh blood samples were centrifuged at 3,000 rpm for 10 min at 4°C (Centrifuge CF16RN HITACHI, Tokyo, Japan). Then we separated and transferred supernatants (serum) into new Eppendorf tubes and immediately froze them at −80°C until analysis.

After melting on ice, the serum (100 µl of sample into 300 µl of methanol solution containing 0.05 µg/ml of L-2-chlorophenylalanine and 0.5 µg/ml of ketoprofen as internal standard) was added to the samples. After being vortexed for 1 min, the mixture was centrifuged at 13,000 pm at 4°C for 10 min, and then 200 µl of supernatant was transferred to the autosampler vial for UHPLC-MS/MS analysis.

The reproducibility and reliability of UHPLC-MS/MS system were evaluated by quality control (QC) samples. After the equipment was stabilized, six QC samples were analyzed primarily. A blank (acetonitrile) was inserted after each QC sample to wash the column. One QC sample was injected at the beginning analysis, and QC samples were evenly inserted every 10 samples in the sequence of analytical workflow.

Ultra-High-Performance Liquid Chromatography–Quadrupole-Orbitrap Analysis

We used an UHPLC system to achieve chromatographic separation (Dionex, Themo Fisher Scientific, Waltham, MA, USA). Gradient elution was performed. Aliquots measuring five microliters from each sample were injected into an ACQUITY UHPLC® BEH C_{18} (2.1 mm × 100 mm, 1.7 µm, Waters, USA). Mobile phase A was acetonitrile, and mobile phase B was water containing 0.1% formic acid. The gradient elution was as follows: 0–0.5 min, 5% A; 0.5–1.0 min, 5%–60% A; 1.0–7.0 min, 60%–80% A; 7.0–9.0 min, 80%–100% A; 9.0–11.0 min, 100% A; 11.0–11.2 min, 100%–5% A; and 11.2–13.0 min, 5% A. The flow rate was 0.2 ml/min.

The Q-Exactive Orbitrap MS was combined with the UHPLC system, which used a heated electrospray ionization (HESI) source. The mass spectra were respectively acquired in the positive and negative modes through full-mass/dd MS^2 (data-dependent MS) scanning patterns. The instrument scanned a mass range from 80 to 1,200 *m*/*z* with a mass resolution power of 17,500 in MS/MS. The temperature of the auxiliary gas was 300°C with a flow rate of 10 arb. The ion source temperature was 350°C and the capillary temperature 320°C. The collision energy was set at 20, 40, and 60 eV with the spray voltage at 3.50 kV in the positive mode or 2.8 kV in the negative mode. The analytical sequence of every experimental sample was random.

Identification of Differential Metabolites and Kyoto Encyclopedia of Genes and Genomes Enrichment Analysis

The comprehensive peak table (molecular weight, retention time (RT), and peak area) generated by metabolites was extracted from the raw data file using compound discoverer 3.1 software (Version 3.0, Thermo Scientific). Import the comprehensive peak table into Xcalibur™ software to realize the visualization (Version 3.0, Thermo Fisher Scientific). Then, the mass spectra and spectral data entered into Xcalibur™ software were compared with the human metabolomics database (Human Metabolome Database (HMDB), http://hmdb.ca/) and PubChem compound database to determine the different metabolites.

Obtain the genes corresponding to the differential metabolites in the HMDB, import the above genes into KEGG database for enrichment analysis, and visualize the pathways with $p < 0.05$ and false discovery rate (FDR) <0.05.

Determine the Change Multiple of Metabolites

The fold change (FC) value of each metabolite in each patient was calculated by dividing the peak area of each patient's corresponding preoperative group by the peak area of each patient's corresponding postoperative group, to observe the relationship between the changes of metabolites and the prognosis of patients with OSCC. Instead of the average FC value, the best cutoff value was calculated using X-Tile software (https://medicine.yale.edu/lab/rimm/research/software/) to divide the samples into metabolites change-high group and metabolites change-low group.

Statistical Analysis

The data result set, which corresponds to the concentration of certain metabolites, contained all the *m*/*z* value, RT, and ion peak area of each sample. Principal component analysis (PCA), orthogonal partial least square discrimination analysis (OPLS-DA), variable importance in projection (VIP), and the 200 times permutation test were obtained from multivariate statistical SIMCA software (Version 14.0 Umetrics, Umea, Sweden). A Student's t-test and FC of all the peaks were used by the SPSS 21.0 software (IBM, Chicago, IL, USA). MetaboAnalyst (https://www.metaboanalyst.ca/) was used to generate the heatmap to show the trend of change, which was created by these screened metabolites.

The 3-year recurrence-free survival (RFS) rate was evaluated using the Kaplan–Meier method and log-rank test. The Cox proportional hazards model was used to estimate the independent prognostic factors for RFS. *p*-Values <0.05 were considered statistically significant. The area under the curve (AUC) and receiver operating characteristic (ROC) curve were used to evaluate and compare the prognostic value of prognostic biomarkers.

RESULTS

Clinical Characteristics of the Study Subjects

We collected a total of 73 cases with primary OSCC, including 35 males and 38 females, with a mean age of 58.5 (median 56, range 28–75) years. The mean follow-up period was 14.9 months (median = 14.2, range 2.4 to 31.8). Of these 73 patients, 39 patients were alive without recurrence, and three patients died of the disease.

The remaining 31 patients were alive but relapsed or metastasized. Due to the low number of deaths, no OS analysis was performed. The clinical parameters of all 73 patients are summarized in **Supplementary Table S1**, with all the details in the Supplementary Material.

Metabolomics Analysis and Biomarker Identification

Multiple methods, including the use of internal standards and QC samples, were used to ensure stable and reliable metabolic profiling results. In PCA score plots, QC samples were clustered closely in the positive mode (**Figure 1A**), which showed that the analytical process was stable and credible. To gain insights into the metabolic features of OSCC before and after the resection, UHPLC/Q-Orbitrap HRMS was performed on these serum samples. All data of preoperative and postoperative groups were processed to normalize the ion peak areas and exported to the SIMCA 14.0 software to perform multivariate statistical analysis. A total of 2,451 ion peaks in positive ion modes were extracted. In both PCA and OPLS-DA score plots, the preoperative and postoperative groups showed a clear separation (**Figures 1A, B**), R^2Y at 0.702 and Q^2 at 0.553. The 200 times permutation test (**Figure 1C**) showed that the model was not over-fitting (R^2 at 0.372 and Q^2 at 0.997 in the posit ion mode). The results of the negative ion pattern also distinguished the preoperative group from the postoperative group in terms of metabolic changes (**Supplementary Figure S1**, Supporting Information). A combination of p-values <0.05 and VIP values >1 was used to identify metabolic biomarkers. In addition, a total of 14 significant metabolites (6 increased and 8 decreased in the postoperative group of patients) were annotated using online databases and reference standards, including succinic acid, hypoxanthine, glutamine, and arginine (**Table 1**). The heatmap (**Figure 2**) shows the differences in metabolite distribution between the two groups. The correlation among these 14 differentially expressed metabolites was explored using Spearman's correlation analysis. As **Figure 3B** shows, the metabolites with smaller p-values were enriched in either the preoperative subjects or the postoperative groups that had stronger correlations. To further explore the underlying molecular mechanism of OSCC, the metabolic pathways of the metabolites were analyzed by MetaboAnalyst (**Figure 3A**). The

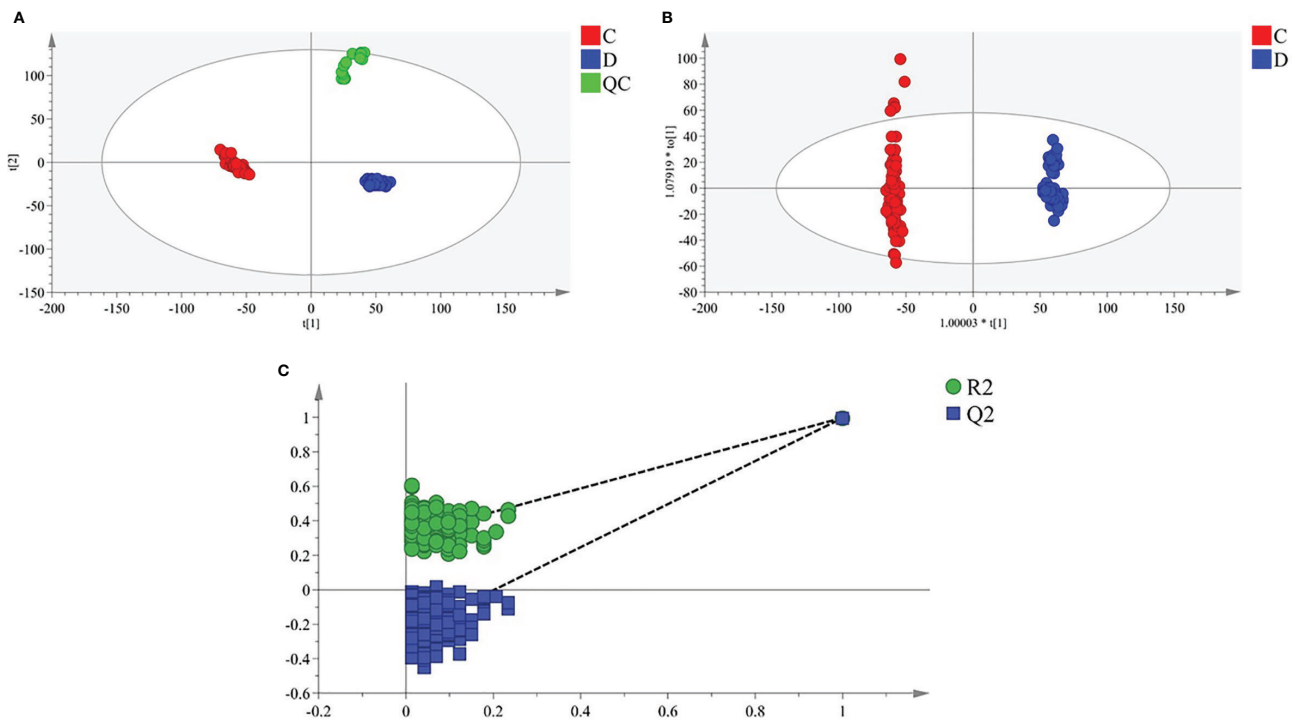


FIGURE 1 | Multivariate statistical analysis of two groups. The principal component analysis (PCA) plot of QC and samples in **(A)** positive ion mode. The orthogonal partial least square discrimination analysis (OPLS-DA) score plots of preoperative group *vs.* postoperative group in **(B)** positive ion mode. Cross-validation plot with a permutation test repeated 200 times of preoperative group vs. postoperative group in **(C)** positive ion mode. C, preoperative group; D, postoperative group; QC, quality control.

TABLE 1 | Statistical analysis of potential metabolic biomarkers.

No.	Metabolites	Ion mode	RT (min)	Molecular	VIP	FC (preoperative/postoperative)	*p*-Value	Class
1	Succinic acid	P	1.54	130.12	1.63	3.34	2.88E−159	OAC
2	Arginine	P	1.558	174.11	1.69	2.86	4.21E−110	AA
3	9-Decanoylcarnitine	P	1.316	313.22	1.68	2.77	7.44E−110	CC
4	Asparaginyl-Valine	P	1.143	231.122	1.63	2.59	1.51E−86	AA
5	Glutamine	P	0.968	146.07	1.62	1.83	1.48E−82	AA
6	Hypoxanthine	N	1.852	136.04	1.62	1.47	6.49E−82	PPN
7	Sphingosine	P	1.472	299.28	1.6	1.43	1.30E−80	SC
8	Palmitoyl ethanolamide	P	1.302	299.28	1.7	1.35	4.98E−78	OTH
9	Hexanoylcarnitine	P	1.725	259.18	1.61	0.64	1.10E−76	CC
10	Orotic acid	P	1.421	156.01	1.66	0.6	3.77E−74	VIT
11	Uric acid	P	9.753	168.03	1.5	0.44	4.55E−74	PPN
12	Vanillyl mandelic acid	N	1.374	198.05	1.59	0.43	1.18E−67	OAC
13	Ethyl acetate	P	1.156	88.05262	1.67	0.39	1.13E−66	OTH
14	Thromboxane B2	P	1.082	408.19712	1.5	0.39	1.29E−66	OAC

p-value, the analysis was adjusted by gender, age, BMI, smoking status, steroid use, seizure medication use, and diabetes medication use.
RT, retention time; FC, fold change; VIP, variable importance in projection, obtained from preoperative group vs. postoperative group in discovery cohort; OAC, organic acid and conjugates; AA, amino acids and derivatives; CC, carnitines and conjugates; PPN, purine derivatives and purine nucleosides; VIT, vitamins; SC, sphingolipids and conjugates; OTH, others.

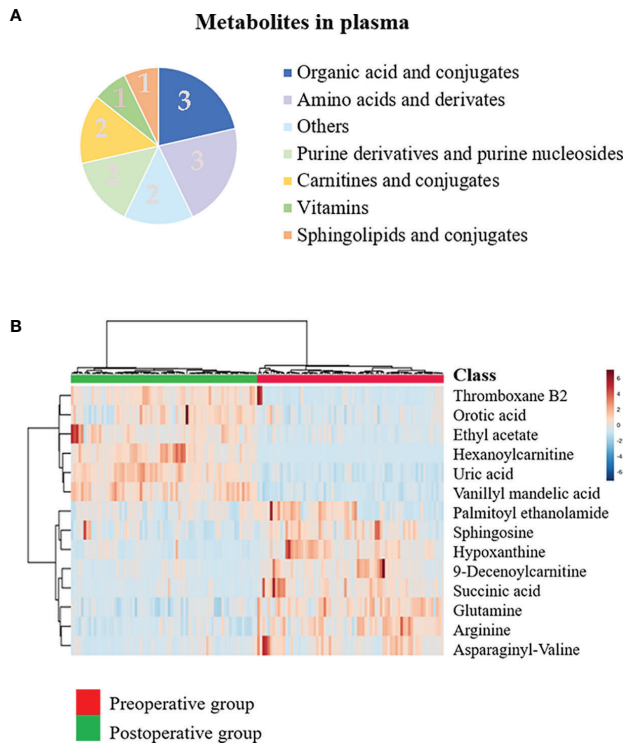


FIGURE 2 | Discovery of the prognostic biomarkers for oral squamous cell carcinoma (OSCC) in plasma metabolomics. **(A)** Pie chart summarizing the taxonomic annotation of differentially expressed metabolites in the preoperative group vs. postoperative group (numbers in the pie chart represent the number of metabolites). **(B)** Heatmap representing the relative levels of 14 significantly different metabolites in the plasma samples of the preoperative group and postoperative group.

results showed that citrate cycle metabolism, purine metabolism, alanine, aspartate and glutamate metabolism, pyrimidine metabolism, and sphingolipid metabolism were associated with OSCC.

Succinic Acid Change-Low and Hypoxanthine Change-High Were Independent Prognostic Factors for 3-Year Recurrence-Free Survival

Of the 14 metabolites identified, ROC was performed to calculate the AUC. Four metabolites had AUC >0.9 (**Figure 4**), and ROC curves for the other four metabolites in plasma with AUC >0.8 are shown in **Supplementary Figure S2** in the Supporting Information. In the postoperative group compared with the preoperative group, the levels of succinic acid, arginine, 9-decanoylcarnitine, asparagine-valine, glutamine, hypoxanthine, sphingosine, and palmitoyl ethanolamide decreased with multiplicative changes of 3.34, 2.86, 2.77,2.59, 1.83, 1.47, 1.43, and 1.35, respectively. The levels of hexanoylcarnitine, orotic acid, uric acid, vanillyl mandelic acid, ethyl acetate, and thromboxane B2 were elevated with multiplicative changes of 0.64, 0.44, 0.6, 0.43, 0.39, and 0.39, respectively. These metabolites may provide new clues for future prognosis of OSCC.

Next, we investigated the prognostic value of various clinicopathological parameters and the change of metabolites in our cohort. We divided the patients into two groups according to the cutoff value of metabolites' FC obtained from X-Tile, with metabolites with AUC >0.8 including thromboxane B2, succinic acid, glutamine, hypoxanthine, arginine, 9-decanoylcarnitine, orotic acid, and asparaginyl-valine (**Table 2**). The univariate analysis using the log-rank test showed that tumor grade (differentiation), tumor T stage, succinic acid change-low, and hypoxanthine change-high indicated a trend towards improved RFS (**Figure 5**). Succinic acid change-low and hypoxanthine change-high were significantly associated with a better 3-year RFS rate (**Figure 6**). A total of 30 additional plasma samples in the test set were used to evaluate the potential prognostic evaluation ability of identified metabolites. Succinic acid change-low and hypoxanthine change-high are also related to the better 3-year RFS rate (**Figure 5**). Variables that showed statistically significant associations with 3-year RFS rates in the

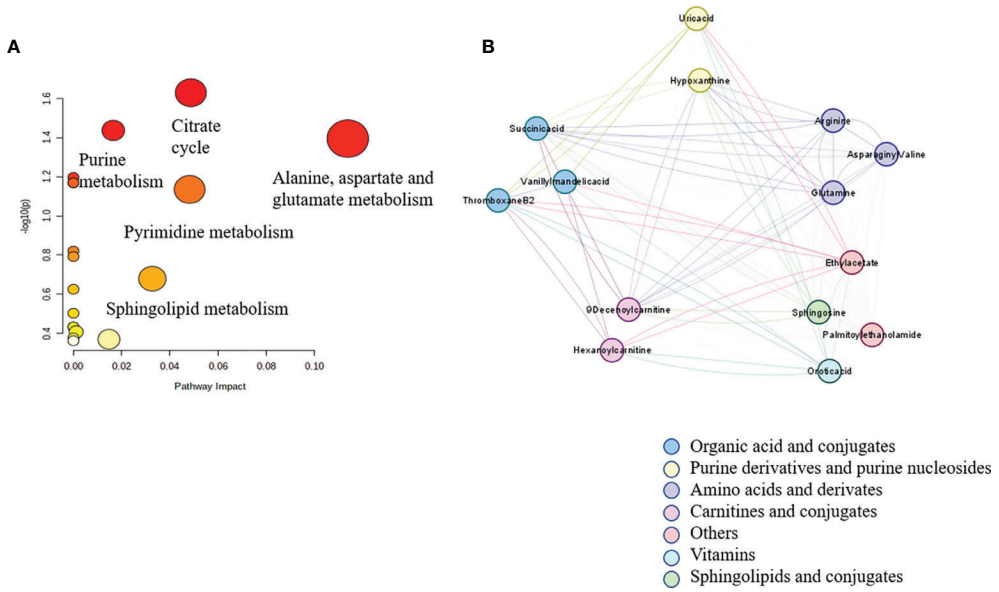


FIGURE 3 | Correlation network analysis of metabolites identified in untargeted metabolomics. Spearman's correlation analysis of 14 significantly different metabolites in the plasma samples of the preoperative group and the postoperative group in graph **(A)**. The disturbed metabolic pathways showed various metabolism changes between the preoperative group and postoperative group in graph **(B)**.

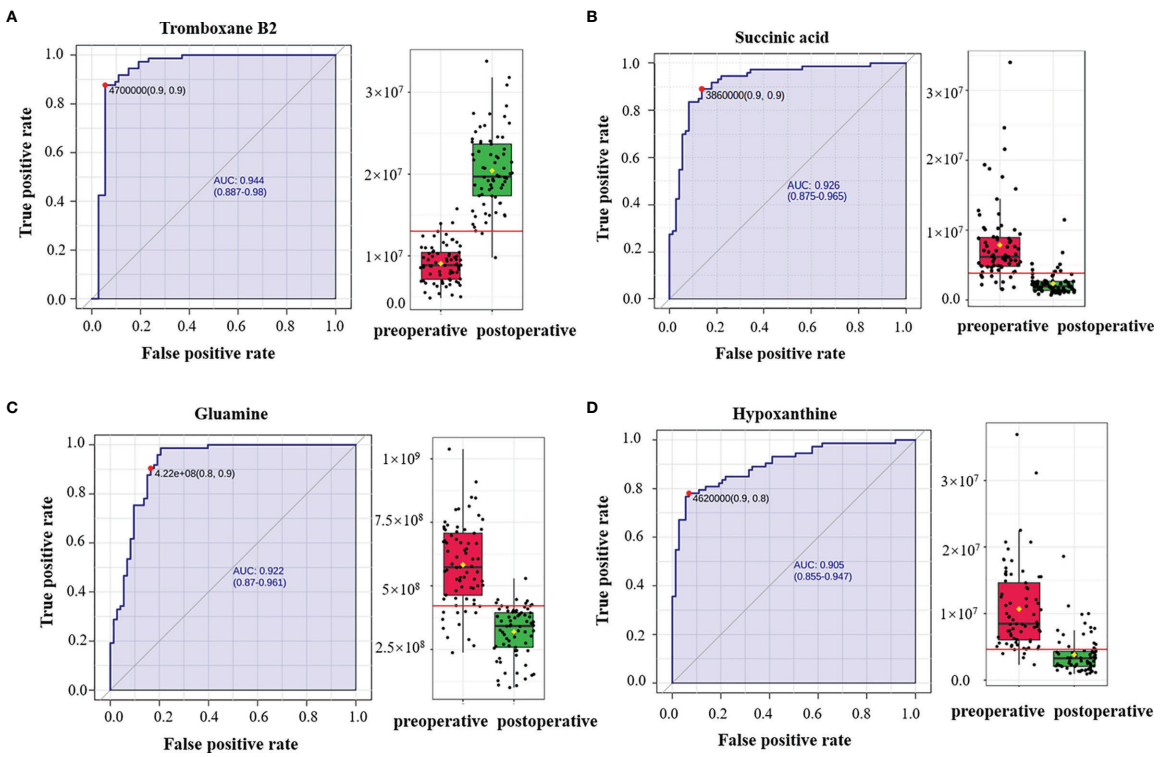


FIGURE 4 | Receiver operating characteristic (ROC) curves of plasma metabolites with area under the curve (AUC) values exceeding 0.9. ROC curves of **(A)** thromboxane B2, **(B)** succinic acid, **(C)** glutamine, and **(D)** hypoxanthine. The AUCs of thromboxane B2, succinic acid, glutamine, and hypoxanthine were 0.944 (95% CI = 0.887–0.98), 0.926 (95% CI = 0.875–0.965), 0.922 (95% CI = 0.87–0.961), and 0.905 (95% CI = 0.855–0.947), respectively. The box plots show the median, the quartiles, and the whole range of peak area of these metabolites.

TABLE 2 | Univariate and multivariate survival analysis of clinicopathological parameters.

Variables		3-year RFS (%)	Univariate analysis (*p*)	Multivariate analysis (*p*)
Gender	Male	47.6	0.552	
	Female	61.3		
Age	<60	54.3	0.913	
	>60	52.6		
Smoking	Current	67.9	0.651	
	Never or former	72.6		
Alcohol	Current	68.1	0.727	
	Never or former	72.7		
Tumor site	Tongue	75.3	0.3	
	Buccal	61.5		
	Gingiva	75.7		
	Others	57.9		
Differentiation	Grade I	75.7	0.025*	0.589
	Grade II/III	61.5		
T stage	T1/T2	82.4	0.003*	0.384
	T3/T4	51.5		
Succinic acid	Low	21.4	0.015*	0.001*
	High	79.5		
Hypoxanthine	Low	63.3	<0.001*	0.023*
	High	80		
Thromboxane B2	Low	48.9	0.332	
	High	40		
Asparaginyl-valine	Low	44	0.582	
	High	52.2		
Glutamine	Low	47.6	0.552	
	High	61.3		
Arginine	Low	54.5	0.198	
	High	34.5		
9-Decanoylcarnitine	Low	43.1	0.723	
	High	54.5		
Orotic acid	Low	35.4	0.24	
	High	52		

Variables that showed statistically significant associations with 3-year RFS rates in the univariate analyses were entered into multivariate Cox regression analyses.
RFS, recurrence-free survival.
**p < 0.05.*

univariate analyses were entered into multivariate Cox regression analyses (**Table 2**). In multivariate analyses, succinic acid change-low expression and hypoxanthine change-high expression were independent prognostic factors for the 3-year RFS rate (hazard ratio [HR] = 5.730, 95% CI, 1.667–19. 694; [HR] = 3.221, 95% CI, 1.233–8.414).

Metabolic Prognostic Panel Had High Predictive Accuracy on 3-Year Recurrence-Free Survival

To determine the predictive accuracy of succinic acid change-low and hypoxanthine change-high on 3-year RFS, we performed ROC curve analyses. As shown in **Figure 6** and **Table 3**, T-stage, tumor differentiation, succinic acid change-low, and hypoxanthine change-high had similar predictive accuracies ($p < 0.05$), but the panel of succinic acid change-low, hypoxanthine change-high, and tumor grade (differentiation) provided the highest predictive accuracy (AUC = 0.730; 95% CI, 0.654–0.805).

Biological Networks Associated With the Differentially Expressed Metabolites and Genes

The HMDB and KEGG database revealed some genes that are closely correlated with the 14 differentially expressed metabolites. The "metabolites-genes" network (**Figure 7A**) was then constructed by using Cytoscape software 3.8.2. However, no related genes were found for palmitoyl ethanolamide, hexanoylcarnitine, asparaginyl-valine, and 9-decanoylcarnitine.

KEGG enrichment analysis showed that a total of 13 pathways ($p < 0.05$, FDR < 0.05) were significantly disturbed in the postoperative group compared with the preoperative group (**Figure 7B**). The most important genes are involved in amino acid metabolism and purine metabolism. In order to further explore the relationship between genes and metabolism, the changed metabolites and genes were mapped to the relevant networks by searching the online KEGG databases and HMDB. The metabolic profiles of succinic acid and its regulatory genes and the metabolites of hypoxanthine and its regulatory genes are shown in **Figure 8**. These results suggest that a considerable number of genes in amino acid metabolism and purine metabolism pathway are closely related to the prognosis of OSCC patients.

DISCUSSION

As one of the major components of systems biology, metabolomics is a well-established method to assess global metabolic profiles through biomarker discovery in accessible biofluids (15, 17, 18). In this study, global non-targeted metabolomics was established to investigate changes in metabolic phenotypes associated with OSCC, and transcriptome analysis was performed to reveal genes associated with metabolites found to be differentially expressed in OSCC patients. This suggests that several metabolites and genes are commonly involved in metabolic pathways and regulatory signaling in OSCC. As a series of works, our study not only dissects the regulatory features of metabolic networks in OSCC but also explores their ability to predict prognosis in OSCC.

Low succinate in OSCC is associated with better 3-year RFS of the patients, suggesting that succinate accumulation is associated with a worse prognosis. Succinate is an inhibitor of prolyl hydroxylase (PHD) (19), which is responsible for hydroxylation of hypoxia-inducible factor 1-alpha (HIF1α), causing its degradation. Then, succinate accumulation results in a pseudo-hypoxic response that is caused by HIF1α stabilization and activation of genes containing HIF response elements (HREs) (20–22). Overall, succinate accumulation plays an important role in the epigenetic alteration of cancer cells, cancer cell metabolism, epithelial-to-mesenchymal transition (EMT), and angiogenesis.

Succinate accumulation induces epigenetic alterations in cancer cells, which causes competitive inhibition of several alpha-ketoglutarate (αKG)-dependent dioxygenase. In these

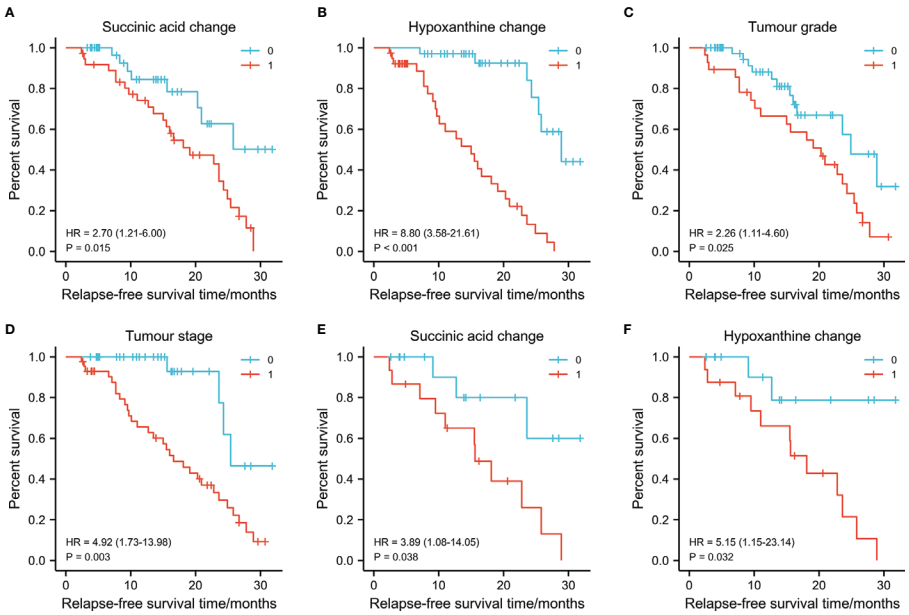


FIGURE 5 | Graphs **(A–D)** showing the 3-year survival rates of succinic acid change, hypoxanthine change tumor stage, and tumor grade in 73 oral squamous cell carcinoma (OSCC) patients (training set) using the Kaplan–Meier method and log-rank test. **(A)** Succinic acid change-low had a higher 3-year recurrence-free survival (RFS) rate than succinic acid change-low (1) (79.4% vs. 30.8%, $p = 0.015$). **(B)** Hypoxanthine change-high had a higher 3-year RFS rate than hypoxanthine-low (81.6% vs. 22.9%, $p < 0.001$). **(C)** Grade I (well differentiated) had a higher 3-year RFS rate than grade II/III (moderately or poor differentiation) (53.3% vs. 17.9%, $p = 0.025$). **(D)** Early T stage (T1/T2) had a higher 3-year RFS rate than late T stage (T3/T4) (86.7% vs. 55.9%, $p = 0.003$). Graphs **(E)** and **(F)** show the 3-year survival rates of succinic acid change and hypoxanthine change in 30 OSCC patients (test set) using the Kaplan–Meier method and log-rank test. **(E)** Succinic acid change-low had a higher 3-year RFS rate than succinic acid change-low (62.4% vs. 42.8%, $p < 0.038$). **(F)** Succinic acid change-low had a higher 3-year RFS rate than hypoxanthine-low (65.6% vs. 41.6%, $p = 0.032$). 0 in graphs **(A–F)** represents succinic acid change-low, hypoxanthine change-high, grade I, early T stage (T1/T2), succinic acid change-low, and succinic acid change-low, respectively. 1 in graphs **(A–F)** represents succinic acid change-high, hypoxanthine change-low, grade II/III, late T stage (T3/T4), succinic acid change-low, and succinic acid change-high, respectively.

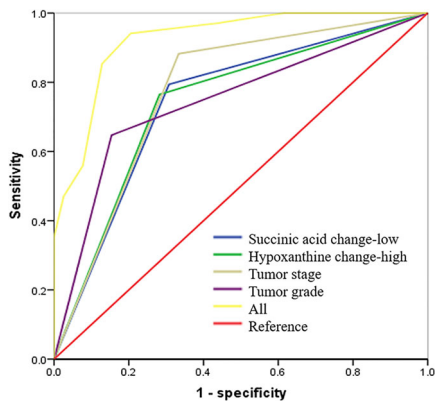


FIGURE 6 | Succinic acid change-low and hypoxanthine change-high had high predictive accuracy for 3-year recurrence-free survival (RFS). receiver operating characteristic (ROC) curves indicating the predictive accuracy, sensitivity, and specificity of each potential parameter. The area under the curve (AUC) values of succinic acid change-low, hypoxanthine change-high, tumor stage, and tumor grade (differentiation) were 0.743 (95% CI = 0.627–0.859), 0.794 (95% CI = 0.686–0.903), 0.775 (95% CI = 0.664–0.885), and 0.747 (95% CI = 0.629–0.864), respectively; but succinic acid change-low combined with hypoxanthine change-high and tumor grade provided the highest predictive accuracy (AUC = 0.900; 95% CI = 0.822–0.967).

αKGs, Ten-eleven-translocation (TET) and Jumonji domain-containing histone demethylases (JMHD) are responsible for histone hypermethylation and decrease of hydroxylation of 5mC (20, 23). JMHD causes the oxidation of methyl groups on lysine residues of histones H3 and H4. Its inhibition induces global histone hypermethylation that alters epigenetic control of gene expression, with potential tumorigenic consequences (3, 24). The effect of succinic acid accumulation on cell transcriptome mediates the pseudo-hypoxic phenotype and induces the change of metabolic phenotype, which leads to the bioenergy conversion from mitochondrial respiration to cytosolic glycolysis (22, 25, 26). The accumulation of succinic acid can lead to the loss of succinate dehydrogenase (SDH) activity and lead to changes in the metabolism of non-essential amino acids, especially aspartic acid, which is the main precursor of protein and nucleotide biosynthesis, as well as other non-essential amino acids such as arginine and asparagine (27). Hypermethylation induced by succinic acid accumulation promotes EMT, migration, and invasion (27, 28).

EMT allows epithelial cancer cells to present mesenchymal features, providing them with enhanced motility and invasiveness, thus allowing cancer to spread and metastasize. Hypermethylation induced by succinic acid accumulation promotes EMT, migration, and invasion (28). Succinate

TABLE 3 | Summary of the RFS predictive accuracy of succinic acid and hypoxanthine.

Predictive factors	AUC	SE	*p*-Value	95% CI	
T stage	0.775	0.056	<0.000*	0.664	0.885
Succinic acid change-low	0.641	0.059	<0.000*	0.627	0.859
Hypoxanthine change-high	0.682	0.055	<0.000*	0.686	0.903
Differentiation	0.656	0.06	<0.000*	0.629	0.864
All[a]	0.9	0.034	<0.000*	0.822	0.967

Note. SE, standard error; 95% CI, 95% confidence interval; RFS, recurrence-free survival; AUC, area under the curve.

[a]*Succinic acid change-low, hypoxanthine change-high, and differentiation.*

*$p < 0.05$.

accumulation also promotes angiogenesis. In SDH deficient prostaglandins and prostate cancer tissues, learner found that succinic acid accumulation was associated with expression of inducible factor-1 α, angiogenic genes and high density of microvessels. (19, 29, 30).

In the past, succinic acid was considered as an intermediate of citric acid cycle. However, it also plays a role in gene expression and intercellular communication (10). Recently, the importance of succinic acid accumulation in carcinogenesis progression has been fully demonstrated, which fully proves that succinic acid is a tumor-related metabolite. Serum organic acid analysis can be used as an effective and cheap broad-spectrum screening method to narrow the scope of more expensive gene sequencing (31).

In our study, hypoxanthine change-high is associated with better 3-year RFS, suggesting that consumption of hypoxanthine is associated with a worse prognosis. Hypoxanthine-guanine phosphoribosyl transferase (HPRT) is an enzyme in the DNA salvage pathway responsible for recycling GTP and is involved in the production and regulation of the purinosome, with a significant regulatory role in the synthesis rate of purines during the cell cycle. It is significantly elevated in cancer cells (32, 33). Hypoxanthine is one of the substrates of HPRT. With the increase of HPRT level, the consumption of hypoxanthine is more. Wang et al. reported that HPRT promotes proliferation and metastasis in head and neck squamous cell carcinoma, which concurs with the results in the present study (34).

The discovery and detection of metabolites in serum of patients with cancer have created a new paradigm of cancer biology. It is possible to detect metabolites related to early prognosis and take corresponding treatment. This includes the discovery of new therapeutic targets that exploit vulnerabilities of cancer cells, such as their dependence on oncometabolites. Measurement of succinic acid, hypoxanthine, and other metabolites will be an ideal tool for screening and tracking

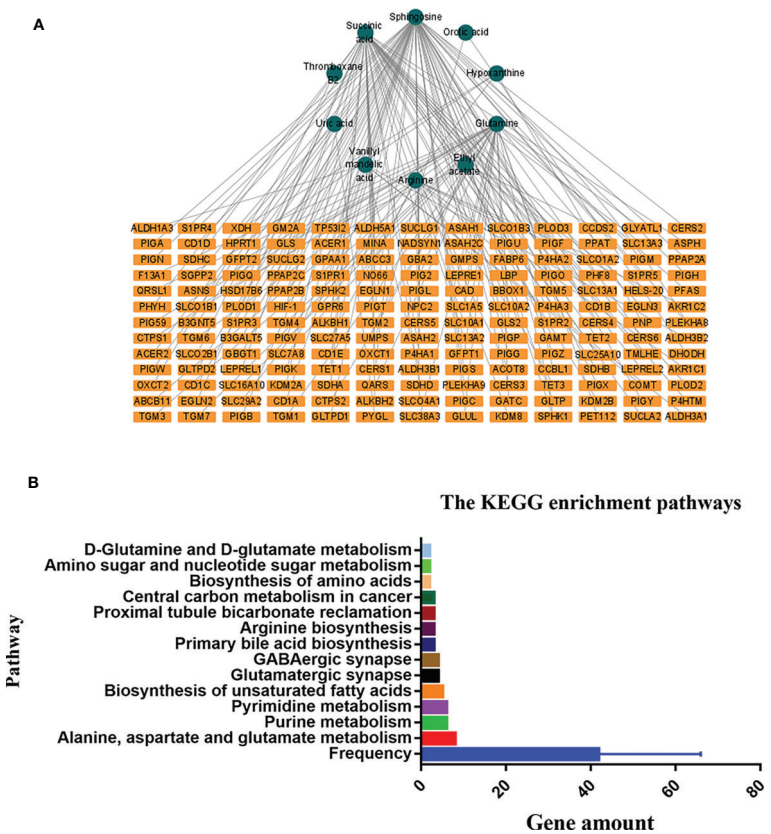


FIGURE 7 | The connected network of metabolites and genes in graph **(A)**. The orange rectangles circles node represents the differentially expressed metabolites in the preoperative group vs. postoperative group. The blue rectangles represent the genes closely correlated with those metabolites. The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathways in the preoperative group and postoperative group in graph **(B)**.

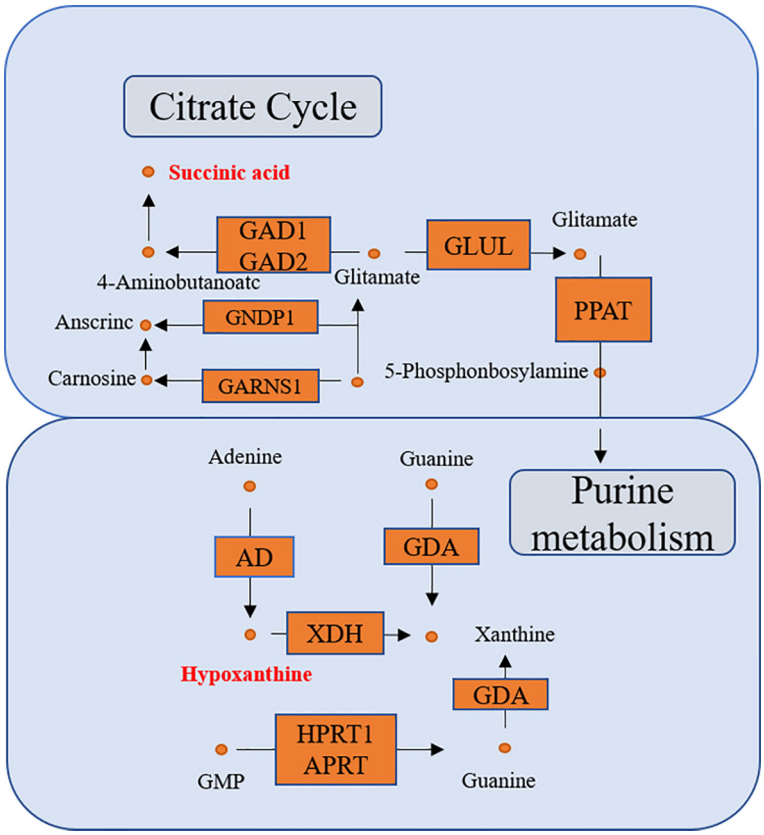


FIGURE 8 | The connected network of genes, enzymes, and metabolites. The orange rectangles and circles represent the significant genes and metabolites, respectively.

OSCC with corresponding metabolic disorders (35). Advances in MS and nuclear magnetic resonance technology have promoted high-resolution metabolite mapping of cells and tumors and have identified the accumulation of metabolites associated with specific gene defects (36).

In conclusion, this study established a new method to evaluate the prognosis of patients with OSCC, using UHPLC-Q-Orbitrap HRMS serum metabolomics analysis, which showed higher predictive accuracy in patients with OSCC. However, it is not so accurate to only use FC mean values as a standard. In the future, more samples need to be collected to standardize the measurement.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committees of The First Affiliated Hospital of Zhengzhou University (name of IRB: Ethics Committee of Scientific Research Project of The First Affiliated Hospital of Zhengzhou University; ethical code: SB201902006). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ZS designed the study. LZ assisted in the conceptualization of the study. ZC and QJ conducted the experiments. ZC and LC undertook the data analysis and wrote the manuscript. SZ and YH collected the assay samples. JK, YS, and LL aided in editing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 82003921), the Foundation of Beijing medical and health (No. YWJKJJHKYJJ-B17818), the key scientific research project of Henan institution of higher education (No. 19A360008 and No. 21A320057), the Joint Construction Project of Henan Medical Science Project (No. SBGJ22002078), and the Health Science and Technology

Innovation Excellent Talents Training Project for Young and Middle-aged in Henan Province (No. YXKC2020058).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.750794/full#supplementary-material

Supplementary Tables S1 | The baseline characteristics of 73 participants

Supplementary Figures S1 | Multivariate statistical analysis of two groups. The principal component analysis (PCA) plot of QC and samples in **(A)** the negative positive ion mode. The orthogonal partial least square discrimination analysis (OPLS-DA) score plots of preoperative group vs. postoperative group in **(B)** negative ion mode. Cross-validation plot with a permutation test repeated 200 times of preoperative group vs. postoperative group in **(C)** negative ion mode. C, preoperative group; D, postoperative group; QC, quality control.

Supplementary Figures S2 | ROC curves of plasma metabolites with AUC values exceeding 0.8 in our study.

Supplementary Figures S3 | The complete analyses steps of a flowchart in our study.

REFERENCES

1. Chi AC, Day TA, Neville BW. Oral Cavity and Oropharyngeal Squamous Cell Carcinoma-an Update. *CA Cancer J Clin* (2015) 65(5):401–21. doi: 10.3322/caac.21293
2. Hedberg ML, Goh G, Chiosea SI, Bauman JE, Freilino ML, Zeng Y, et al. Genetic Landscape of Metastatic and Recurrent Head and Neck Squamous Cell Carcinoma. *J Clin Invest.* (2016) 126(4):1606. doi: 10.1172/JCI86862
3. Lee YS, Johnson DE, Grandis JR. An Update: Emerging Drugs to Treat Squamous Cell Carcinomas of the Head and Neck. *Expert Opin Emerg Drugs* (2018) 23:283–99. doi: 10.1080/14728214.2018.1543400
4. Wang Y, Ow TJ, Myers JN. Pathways for Cervical Metastasis in Malignant Neoplasms of the Head and Neck Region. *Clin Anat.* (2012) 25(1):54–71. doi: 10.1002/ca.,21249
5. Wei J, Xie G, Zhou Z, Shi P, Qiu Y, Zheng X, et al. Salivary Metabolite Signatures of Oral Cancer and Leukoplakia. *Int J Cancer.* (2011) 129(9):2207–17. doi: 10.1002/ijc.25881
6. Vander Heiden MG, DeBerardinis RJ. Understanding the Intersections Between Metabolism and Cancer Biology. *Cell* (2017) 168(4):657–69. doi: 10.1016/j.cell.2016.12.039
7. Xie GX, Chen TL, Qiu YP. Urine Metabolite Profiling Offers Potential Early Diagnosis of Oral Cancer. *Metabolomics* (2012) 8:220–31. doi: 10.1007/s11306-011-0302-7
8. Armitage EG, Barbas C. Metabolomics in Cancer Biomarker Discovery: Current Trends and Future Perspectives. *J Pharm BioMed Anal* (2014) 87 (1):11. doi: 10.1016/j.jpba.2013.08.041
9. Naz S, Moreira dos Santos DC, García A, Barbas C. Analytical Protocols Based on LC-MS, GC-MS and CE-MS for Nontargeted Metabolomics of Biological Tissues. *Bioanalysis* (2014) 6(12):1657–77. doi: 10.4155/bio.14.119
10. Peruzzotti-Jametti L, Bernstock JD, Vicario N. Macrophage-Derived Extracellular Succinate Licenses Neural Stem Cells to Suppress Chronic Neuroinflammation. *Cell Stem Cell* (2018) 22(3):355–368.e13. doi: 10.1016/j.stem.2018.01.020
11. Wishart DS. Emerging Applications of Metabolomics in Drug Discovery and Precision Medicine. *Nat Rev Drug Discovery* (2016) 15(7):473–84. doi: 10.1038/nrd.2016.32
12. Chen X, Yu D. Metabolomics Study of Oral Cancers. *Metabolomics* (2019) 15 (2):22. doi: 10.1007/s11306-019-1483-8
13. Fu Y, Ding L, Yang X, Ding Z, Huang X, Zhang L, et al. Asparagine Synthetase-Mediated L-Asparagine Metabolism Disorder Promotes the Perineural Invasion of Oral Squamous Cell Carcinoma. *Front Oncol* (2021) 11:637226. doi: 10.3389/fonc.2021.637226
14. Yang XH, Jing Y, Wang S, Ding F, Zhang X-X, Chen S, et al. Integrated non-Targeted and Targeted Metabolomics Uncovers Amino Acid Markers of Oral Squamous Cell Carcinoma. *Front Oncol* (2020) 10:426. doi: 10.3389/fonc.2020.00426
15. Yang XH, Zhang XX, Jing Y, Ding L, Fu Y, Wang S, et al. Amino Acids Signatures of Distance-Related Surgical Margins of Oral Squamous Cell Carcinoma. *EBioMedicine* (2019) 48:81–91. doi: 10.1016/j.ebiom.2019.10.005
16. Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head and Neck Squamous Cell Carcinoma. *Nat Rev Dis Primers.* (2020) 6 (1):92. doi: 10.1038/s41572-020-00224-3
17. Wang X, Zhang A, Han Y, Wang P, Sun H, Song G, et al. Urine Metabolomics Analysis for Biomarker Discovery and Detection of Jaundice Syndrome in Patients With Liver Disease. *Mol Cell Proteomics.* (2021) 11(8):370–80. doi: 10.1074/mcp.M111.016006
18. Zhang SN, Li XZ, Wang Y. Neuroprotection or Neurotoxicity? New Insights Into the Effects of Acanthopanax senticosus Harms on Nervous System Through Cerebral Metabolomics Analysis. *J Ethnopharmacol* (2014) 156:290–300. doi: 10.1016/j.cca.2018.09.021
19. Pollard PJ, Brière JJ, Alam NA, Barwell J, Barclay E, Wortham NC, et al. Accumulation of Krebs Cycle Intermediates and Over-Expression of HIF1alpha in Tumours Which Result From Germline FH and SDH Mutation. *Hum Mol Genet* (2015) 14(15):2231–9. doi: 10.1093/hmg/ddi227
20. Laukka T, Mariani CJ, Ihantola T. Fumarate and Succinate Regulate Expression of Hypoxia-Inducible Genes *via* TET Enzymes. *J Biol Chem* (2016) 291(8):4256–65. doi: 10.1074/jbc.M115.688762
21. Tretter L, Patocs A, Chinopoulos C. Succinate, an Intermediate in Metabolism, Signal Transduction, ROS, Hypoxia, and Tumorigenesis. *Biochim Biophys Acta* (2016) 1857(8):1086–101. doi: 10.1016/j.bbabio.2016.03.012
22. Weber A, Klocker H, Oberacher H, Gnaiger E, Neuwirt H, Sampson N, et al. Succinate Accumulation is Associated With a Shift of Mitochondrial Respiratory Control and HIF-1α Upregulation in PTEN Negative Prostate Cancer Cells. *Int J Mol Sci* (2018) 19(7):2129. doi: 10.3390/ijms19072129
23. Xiao M, Yang H, Xu W. Inhibition of α-KG-Dependent Histone and DNA Demethylases by Fumarate and Succinate That are Accumulated in Mutations of FH and SDH Tumor Suppressors. *Genes Dev* (2012) 26(12):1326–38. doi: 10.1101/gad.191056.112
24. Hoekstra AS, de Graaff MA, Briaire-de Brujin IH. Inactivation of SDH and FH Cause Loss of 5hmc and Increased H3k9me3 in Paraganglioma/Pheochromocytoma and Smooth Muscle Tumors. *Oncotarget* (2015) 6 (36):38777–88. doi: 10.18632/oncotarget.6091
25. Lee MR, Mantel C, Lee SA, Moon SH, Broxmeyer HE. Mir-31/SDHA Axis Regulates Reprogramming Efficiency Through Mitochondrial Metabolism. *Stem Cell Rep* (2016) 7(1):1–10. doi: 10.1016/j.stemcr.2016.05.012
26. Tseng PL, Wu WH, Hu TH, Chen C-W, Cheng H-C, Li C-F, et al. Decreased Succinate Dehydrogenase B in Human Hepatocellular Carcinoma Accelerates Tumor Malignancy by Inducing the Warburg Effect. *Sci Rep* (2018) 8(1):3081. doi: 10.1038/s41598-018-21361-6
27. Lussey-Lepoutre C, Hollinshead KE, Ludwig C. Loss of Succinate Dehydrogenase Activity Results in Dependency on Pyruvate Carboxylation for Cellular Anabolism. *Nat Commun* (2015) 6:8784. doi: 10.1038/ncomms9784
28. Loriot C, Domingues M, Berger A. Deciphering the Molecular Basis of Invasiveness in Sdhb-Deficient Cells. *Oncotarget* (2015) 6(32):32955–65. doi: 10.18632/oncotarget.5106
29. Gimenez-Roqueplo AP, Favier J, Rustin P, Mourad JJ, Plouin PF, Corvol P, et al. The R22X Mutation of the SDHD Gene in Hereditary Paraganglioma Abolishes the Enzymatic Activity of Complex II in the Mitochondrial Respiratory Chain and Activates the Hypoxia Pathway. *Am J Hum Genet* (2001) 69(6):1186–97. doi: 10.1086/324413

30. Gimenez-Roqueplo AP, Favier J, Rustin P, Rieubland C, Kerlan V, Plouin P, et al. Functional Consequences of a SDHB Gene Mutation in an Apparently Sporadic Pheochromocytoma. *J Clin Endocrinol Metab* (2012) 87(10):4771–4. doi: 10.1210/jc.2002-020525
31. Dalla Pozza E, Dando I, Pacchiana R, Liboi E, Scupoli MT, Donadelli M, et al. Regulation of Succinate Dehydrogenase and Role of Succinate in Cancer. *Semin Cell Dev Biol* (2020) 98:4–14. doi: 10.1016/j.semcdb.2019.04.013
32. Townsend MH, Anderson MD, Weagel EG, Velazquez EJ, Weber KS, Robison RA, et al. Non-Small-Cell Lung Cancer Cell Lines A549 and NCI-H460 Express Hypoxanthine Guanine Phosphoribosyltransferase on the Plasma Membrane. *Onco Targets Ther* (2017) 10:1921–32. doi: 10.2147/OTT.S128416
33. Townsend MH, Robison RA, O'Neill KL. A Review of HPRT and its Emerging Role in Cancer. *Med Oncol* (2018) 35(6):89. doi: 10.1007/s12032-018-1144-1
34. Wang L, Wang Y, Han N, Wang X, Ruan M. HPRT Promotes Proliferation and Metastasis in Head and Neck Squamous Cell Carcinoma Through Direct Interaction With STAT3. *Exp Cell Res* (2021) 399(1):112424. doi: 10.1016/j.yexcr.2020.112424
35. Collins RRJ, Patel K, Putnam WC, Kapur P, Rakheja D. Oncometabolites: A New Paradigm for Oncology, Metabolism, and the Clinical Laboratory. *Clin Chem* (2017) 63(12):1812–20. doi: 10.1373/clinchem.2016.267666
36. Jones PM, Bennett MJ. Urine Organic Acid Analysis for Inherited Metabolic Disease Using Gas Chromatography-Mass Spectrometry. *Methods Mol Biol* (2010) 603:423–31. doi: 10.1007/978-1-60761-459-341

Conflict of Interest: Author YH was employed by Chongqing Huangjia Biotechnology Limited Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 18 October 2021
doi: 10.3389/fonc.2021.731752

Systematic Analysis and Identification of Dysregulated Panel lncRNAs Contributing to Poor Prognosis in Head-Neck Cancer

Shang-Ju Tang[1,2], *Guo-Rong You*[1], *Joseph T. Chang*[3,4*] *and Ann-Joy Cheng*[1,2,3*]

[1] Department of Medical Biotechnology and Laboratory Science, College of Medicine, Chang Gung University, Taoyuan, Taiwan, [2] Graduate Institute of Biomedical Sciences, College of Medicine, Change Gung University, Taoyuan, Taiwan, [3] Department of Radiation Oncology, Chang Gung Memorial Hospital, Taoyuan, Taiwan, [4] Department of Medical School, College of Medicine, Chang Gung University, Taoyuan, Taiwan

OPEN ACCESS

Edited by:
Die Wang,
Hudson Institute of Medical Research, Australia

Reviewed by:
Beifang Niu,
Chinese Academy of Sciences (CAS), China
Manuel Pires Bicho,
University of Lisbon, Portugal
Lei Lyu,
Wuhan No.1 Hospital, China

***Correspondence:**
Ann-Joy Cheng
annjoycheng@gap.cgu.edu.tw
Joseph T. Chang
jtchang@cgmh.org.tw

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *28 June 2021*
Accepted: *04 October 2021*
Published: *18 October 2021*

Citation:
Tang S-J, You G-R, Chang JT and Cheng A-J (2021) Systematic Analysis and Identification of Dysregulated Panel lncRNAs Contributing to Poor Prognosis in Head-Neck Cancer. Front. Oncol. 11:731752. doi: 10.3389/fonc.2021.731752

Head and neck cancer (HNC) is one of the most prevalent cancers worldwide, accounting for approximately 5% of all cancers. While the underlying molecules and their pathogenetic mechanisms in HNC have yet to be well elucidated, recent studies have shown that dysregulation of lncRNAs may disrupt the homeostasis of various biological pathways. However, the understanding of lncRNAs in HNC is still limited by the lack of expression profiling. In the present study, we employed a systematic strategy to identify a panel of lncRNA associated with HNC. A cancer-related lncRNA profile PCR array was screened to explore potential molecules specific for HNC. A total of 55 lncRNAs were found to be dysregulated in HNC cells when compared to normal keratinocytes. Further analysis of the prognostic significance using The Cancer Genome Atlas (TCGA) database revealed 15 lncRNAs highly correlated with overall survival in HNC patients. Additionally, clinical sample expression analysis of the TCGA-HNSC cohort revealed 16 highly dysregulated lncRNAs in HNC, resulting in a combined 31-lncRNA signature panel that could predict prognosis. Validation of these molecules confirmed the considerable level of altered expressions in HNC cells, with XIST, HOXA11-AS, TSIX, MALAT1, WT1-AS, and IPW being the most prominently dysregulated. We further selected a molecule from our panel (XIST) to confirm the validity of these lncRNAs in the regulation of cancer aggressiveness. Gene ontology (GO) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analyses demonstrated that XIST participated in various cancer-related functions, including cell proliferation and metastasis. XIST silencing with the RNAi technique substantially reduced invasion and migration in several HNC cell lines. Thus, our study defined a 31-lncRNA panel as prognostic signatures in HNC. These perspective results provide a knowledge foundation for further application of these molecules in precision medicine.

Keywords: lncRNA - long noncoding RNA, head and neck cancer, prognostic panel, XIST (X-inactive specific transcript), altered gene expression

INTRODUCTION

Head and neck cancer (HNC) is a complex and difficult to treat disease. While this type of cancer encompasses dysregulations at areas including the mouth, nasal cavity, larynx, and pharynx, over 90% of all HNCs are squamous cell carcinomas, and of the oral region (1). Together, they account for approximately 5% of all cancers worldwide, according to GLOBOCAN 2020 estimates (2). Like other cancers, standard treatment methods include surgery, radiotherapy, chemotherapy, or combination therapy (3). Nevertheless, even after months of treatment, relapse is always a potential problem. Although the 5-year survival rate of HNC patients is roughly 80% when detected at the earliest stages, mid-to late-stage detection causes that number to decrease by over two-fold (4, 5). Therefore, it is critical to identify and understand how these carcinogenic mechanisms and molecules affect cancer, as we may be able to uncover the mysteries behind HNC and how to treat and/or prevent it.

It is well established that while more than 75% of the human genome is transcribed, only 2% consists of coding genes (6). The remaining non-coding RNAs, previously merely labeled as transcriptional noise or garbage sequences and disregarded, have recently become much more recognized. The largest class of the non-coding RNA family, with transcripts longer than 200 nucleotides, are known as long non-coding RNAs (LncRNAs). They have gained significant attention over the past decade, as many studies have confirmed their roles in various biological processes involving transcriptional and epigenetic regulation, metabolism, and multiple cellular functions (7–9).

So far, there has been no distinctive markers for HNC. Nevertheless, many studies have shown that aberrantly expressed lncRNAs may potentially play important roles in this particular cancer type. Currently, only a handful of lncRNAs have been implicated in different cancerous functions such as migration, invasion, and metastasis of HNC, including HOTAIR, UCA1, and MALAT1 (10). A variety of lncRNAs have been discovered to play roles in various cancers. For example, lncRNA HOTAIR and UCA1 have both been found to play carcinogenic roles in multiple cancer types (11, 12). A recent review by Zhou et al. also depicted various lncRNAs that were implicated in HNC metastasis (13). Moreover, since not many studies have profiled lncRNAs in cancers of the head and neck region in combination with prognosis analysis, the results of this research will allow us to better understand the mechanisms behind HNC, and provide new insights on the development of diagnostic, prognostic, or treatment markers.

However, the screening and selection of these molecules are mostly ambiguous, and their prognosis abilities have yet to be thoroughly investigated.

The Cancer Genome Atlas (TCGA) is a comprehensive database for identifying and annotating different genes across multiple cancers. With the recent development in genomic sequencing, many cancer-associated lncRNA studies have been accomplished by solely analyzing and constructing data based on these clinical datasets (14–16). While various studies have profiled lncRNAs across different cancers, including breast cancer (17), lung cancer (18), and esophageal cancer (19), very few focus on the intricacies of HNC. Additionally, although the high-throughput TCGA datasets offer a large library of potential candidate molecules, exclusively relying on database information without experimental validation may limit insight for substantiation of prognostic cancer markers, as many have noted (20–24). Therefore, examination of validated lncRNAs in combination with big-data analysis would be ideal.

PCR array is a relatively new method of gene expression analysis. While it may lack discovery power or high-throughput abilities, it makes up for in its sensitivity, specificity, and great dynamic range. Additionally, data analysis is quick and efficient, as opposed to the cumbersome bioinformatics analysis required for genome-wide methods. Previous studies have reported the use of PCR arrays to analyze genes in specific pathways or diseases. For example, Boone et al. performed two pathway-specific arrays for apoptosis and neurotrophins & receptor genes to elucidate the changes in post-traumatic brain injury (25). Zhang et al. also used a panel of 54 genes specific to Alzheimer's disease to observe changes in gene expression in mice (26).

To our knowledge, there are very few reports that show a systematic profiling investigation of HNC lncRNAs, nor their correlation with prognosis. Herein, we systematically examined differentially expressed lncRNAs in HNC cells using a PCR array-based method. We further assessed our results with prognostic information obtained from a high-throughput database, the TCGA-HNSC cohort. Additionally, expression levels of the top dysregulated lncRNAs from the same TCGA dataset were parallelly assessed to provide a base foundation for our research. In combination with *in silico* and *in vitro* analysis, we defined a panel of 31-lncRNA signatures with valuable prognostic information.

MATERIALS AND METHODS

Cells and Cell Cultures

A total of 10 HNC cell lines, SAS, OECM1, FaDu, Detroit, SCC4, SCC25, OC3, BM1, BM2, and NPC076, and six normal keratinocyte cell lines, CGHNK2, CGHNK4, CGHNK6, CGHNK16, CGHNK47, and NOK were used. Cells are cultured and maintained as previously described (27). Briefly, SAS and NPC076 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM, Gibco®), SCC4 and SCC25 cells were maintained in DMEM/F12 medium (D-MEM/F-12, Gibco®), OECM1, BM1, BM2, and KYSE cells were maintained in Roswell Park Memorial Institute 1640 medium (RPMI 1640, Gibco®), FaDu and Detroit cells were maintained in Minimum Essential Media (MEM, Gibco®), OC3 was maintained in 1:2 DMEM/Keratinocyte Serum-Free Medium (KSFM, Gibco®), and normal keratinocyte cell lines were cultured in KSFM (KSFM, Gibco®). Cancer cell line mediums were supplemented with 7% FBS and 1% Antibiotic-Antimycotic, and all cells were incubated at 37°C in a humidified atmosphere of 5% CO_2.

LncRNA Screening *via* RT^2 PCR Array

Systematic gene profiling was accomplished using Qiagen's PCR array kit, according to the manufacturer's protocol. Briefly, RNA

from cell pellets were extracted and quantified. cDNA synthesis was performed using the RT^2 First Strand Kit (Cat. No. 330401; Qiagen, GmbH), and subsequently combined with the RT^2 SYBR® Green PCR master mix (Cat. No. 330504; Qiagen, GmbH). The master mix was then used in combination with the Human Cancer PathwayFinder RT^2 lncRNA PCR Array (Cat. No. LAHS-002Z; Qiagen, GmbH), and the output data was analyzed using the GeneGlobe Data Analysis Center at http://www.qiagen.com/geneglobe.

LncRNA Analysis *via* RT-qPCR

Cell pellets were washed with PBS and collected for RNA isolation. Total RNA extraction was performed using TRIzol reagent (Gibco BRL), and quantification was achieved with a Nanovue™ spectrophotometer (GE Healthcare). cDNA synthesis was achieved by combining total RNA (2 μg) with 5x first-strand buffer (GeneDireX, Inc.), 0.1M DTT (Invitrogen; Thermo Fisher Scientific, Inc.), 1 unit of RNase inhibitor (Invitrogen; Thermo Fisher Scientific, Inc.), 25 mM dNTPs (Thermo Scientific, Thermo Fisher Scientific, Inc.), and random hexamer primers to a total reaction volume of 30μl. TaqMan qPCR assay kit (Applied Biosystems, Thermo Fisher Scientific, Inc.) was combined with the cDNA to create a 20 μL reaction volume to measure lncRNA expression after 50 cycles. For SYBR green reactions, iQ™ SYBR® Green Supermix (Bio-Rad, Inc.) was used instead. Specific lncRNA PCR primers were designed through primer blast. The PCR primers used in this study are listed in **Supplementary Table S1**. Results were normalized against GAPDH internal control.

Western Blot Analysis of lncRNA Targets

Cell lysates were isolated by homogenization in CHAPS lysis buffer (10 mM Tris, pH 7.4, 1 mM MgCl2, 1 mM EGTA, 150 mM NaCl, 0.5% CHAPS and 10% glycerol; Sigma-Aldrich; Merck KGaA) containing a protease and phosphatase inhibitor. Briefly, cell pellets were resuspended in ice-cold buffer and incubated on ice for 30 minutes. Protein collection *via* centrifugation at 13000 g for 30 minutes at 4°C was performed, followed by protein concentration quantification with Bradford assay (Bio-Rad, Inc.), according to the manufacturer's instructions. Protein separation was performed using 10% sodium dodecyl sulfate-polyacrylamide (SDS-PAGE) gel with 30 μg of protein and transferred onto a nitrocellulose membrane. 5% milk was used for blocking, and specific primary antibodies were hybridized overnight at 4°C. Subsequently, the membranes were incubated with secondary antibodies and visualized through chemiluminescent detection. GAPDH was used as an internal control.

Clinical Evaluation of LncRNAs Related to Prognosis in HNC Patients

The RNA-seq data was obtained through the UALCAN web portal (http://ualcan.path.uab.edu/) (28). The Kaplan-Meier survival curves were plotted to evaluate the prognosis of lncRNAs in high- and low-risk patient groups. The head and neck RNA-seq dataset on the Kaplan-Meier Plotter pan-cancer database (https://kmplot.com/analysis/) was selected to assess the overall survival of lncRNAs. Head-neck squamous cell carcinoma data was collected from sources including TCGA, European Genome-phenome Archive (EGA), and Gene Expression Omnibus (GEO) (n = 500), and cohorts were split by automatic sel7best cut-off for median expression values. Additionally, univariate proportional cox hazard ratios (HRs) with 95% confidence intervals, along with survival p-values calculated by log-rank test were obtained for each lncRNA. TCGA clinical expression level analysis was performed using SurvExpress (http://bioinformatica.mty.itesm.mx:8080/Biomatec/SurvivaX.jsp), and pan-cancer analysis was performed with Gepia2 (http://gepia2.cancer-pku.cn/#index).

Molecular Targets and Pathway Analyses *via* Bioinformatic Methods

Potential lncRNA-binding axes were investigated through various online databases and prediction algorithms. Potential lncRNA-mRNA or protein bindings were screened using ENCORI (version 3.0, http://starbase.sysu.edu.cn/index.php) (29). Gene Ontology (GO) and pathway analysis were conducted with the GO and KEGG database from The Database for Annotation, Visualization and Integrated Discovery (DAVID, version 6.8, https://david.ncifcrf.gov/).

Knockdown LncRNA XIST Expression *via* Specific siRNA Transfection

Knockdown of lncRNAs was accomplished with specific siRNAs (Thermo Scientific, Thermo Fisher Scientific, Inc.). siRNA sequences are listed in **Supplementary Table S1**. Transfection was performed with Lipofectamine 2000™ reagent (Invitrogen, Thermo Fisher Scientific, Inc.) in OPTI-MEM medium (Invitrogen, Thermo Fisher Scientific, Inc.), according to the manufacturer's instructions. Briefly, cells were seeded in a 10-cm dish and incubated for 24 hours prior to transfection. 20 to 40 μg of siRNA was transfected and incubated for another 24 hours before the cells were counted and/or pelleted for functional assay examination.

Determination of Cellular Functions: Growth, Migration, and Invasion

Colony formation assay was performed by seeding 5 x 10^2 to 5 x 10^3 transfected cells into 6-well plates and incubated without disturbance for 10 to 14 days. The cells were then fixed and stained with crystal violet for 2 hours, and the number of colonies formed was counted.

Cell invasion was performed using Millicell® (Millipore) cell culture inserts. Transwell chambers were coated with matrigel, and 1 x 10^6 transfected cells were seeded into the upper chamber. The lower chamber was filled with 20% FBS culture medium to promote cell invasion. After 16 to 24 hours of incubation at 37°C, invaded cells were fixed with formaldehyde for 30 minutes, stained with crystal violet for 2 hours, and the number of invaded cells which passed through the Matrigel-coated membranes were quantified and compared to their control counterparts.

Cell migration assay was performed using the wound-healing method. 1 x 10^5 transfected cells were seeded into Ibidi® culture

inserts for 16 hours. After the cells adhered, the inserts were removed, leaving a cell-free gap in the monolayer of cells. Migration towards the gaps was then photographed and measured at 4-hour intervals, up to 12 hours.

Statistical Analysis

RT-qPCR data was performed with at least three independent experiments for each experimental cohort. Unpaired t-test was used to compare the normal and cancer groups (Graphpad Prism 8.0), and a p-value of ≤ 0.05 was considered statistically significant.

RESULTS

LncRNA Expression Profiling in HNC Cell Lines

To profile lncRNAs associated with HNC, a PCR array with 84 cancer-related lncRNAs was used to examine the differential expressions between three HNC cell lines (SAS, OECM1, and FaDu) and two lines of normal keratinocytes (CGHNK2 and CGHNK6). The three HNC cell line's geometric mean fold regulation (FR) and fold change (FC) of each lncRNA was compared to the mean of the normal cell lines, as summarized in **Supplementary Table S2**, and the relationship between cancerous and normal groups were shown in **Figure 1A**. A screening criterion of a mean |FR| ≥ 1.5 was established, resulting in 55 significantly dysregulated lncRNAs (**Figure 1B**). Among these, 27 lncRNAs were upregulated, and 28 were downregulated (**Figure 1B** and **Tables 1**, **2**). Hierarchical clustering analysis was used to visualize these differentially expressed lncRNAs (**Figure 1C**). These results suggest a panel candidate of lncRNA that may participate in the carcinogenesis of HNC.

Prognostic Significance of the Panel lncRNAs in HNC Patients

The clinical significance of the lncRNA in HNC patients was determined by examining the association between each lncRNA expression and clinical patients' prognosis. The Kaplan-Meier Plotter (KM Plotter) suite was used to analyze overall survival in HNC patients with the TCGA-HNSC dataset (n=500). A total of 41 lncRNAs was examined, which included the 55 candidates post-exclusion of nine molecules without information in KM Plotter. **Figure 2A** shows a few examples of highly significant results. As depicted, the upregulated lncRNAs, XIST, HOXA11-AS, and TERC, were significantly associated with poor prognosis, while IPW, a downregulated lncRNA, was correlated with good

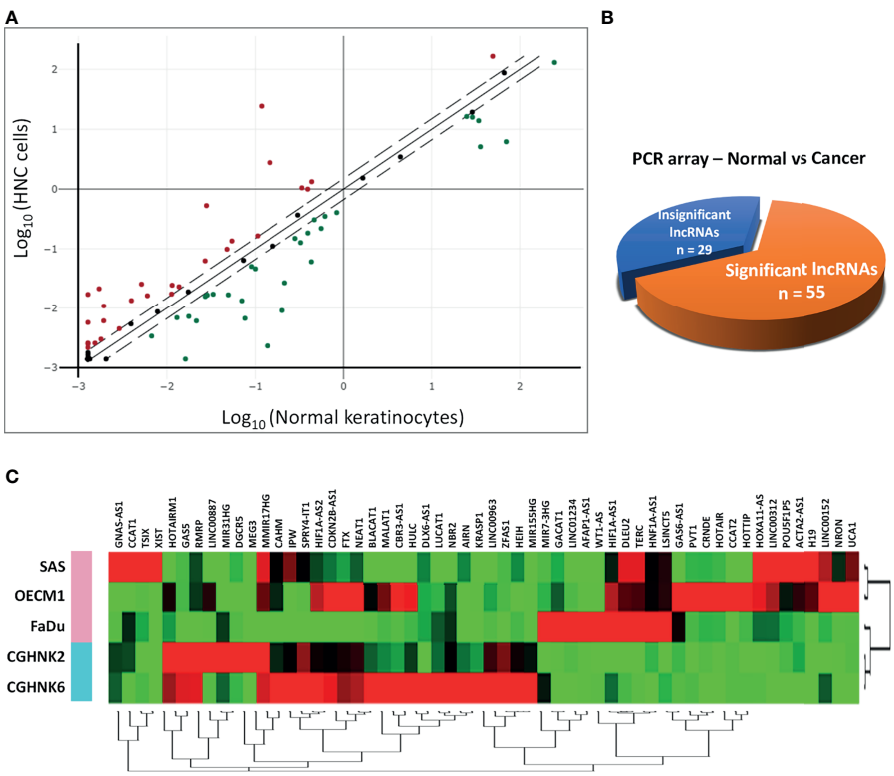


FIGURE 1 | Profiling of 84 lncRNAs in three HNC cell lines (SAS, OECM1, FaDu) versus two normal keratinocyte cell lines (CGHNK2, CGHNK6) *via* lncRNA PCR Array. **(A)** Scatterplot of 84 lncRNAs. Dotted lines represent the selection criteria threshold of |FR| ≥ 1.5. Red dots represent upregulated lncRNAs (n = 27), and green dots represent downregulated lncRNAs (n = 28). Black dots represent insignificant lncRNAs according to the selection criterion (n = 29). **(B)** Pie chart representation of the screening composition between significantly dysregulated lncRNAs (n = 55) and insignificant lncRNAs (n = 29) in HNC, according to the PCR array results. **(C)** Clustergram of the 55 significantly dysregulated lncRNA expression profiles between the HNC and normal keratinocyte cell groups.

TABLE 1 | List of upregulated genes across the HNC cell line group compared with the normal keratinocyte cell line group.

Up-Regulation (n = 27, comparing to control group)				
Symbol	**SAS**	**OECM1**	**FaDu**	**Geometric mean**
	Fold Regulation	**Fold Regulation**	**Fold Regulation**	**Fold Regulation**
H19	895.97	612.4	16.05	206.51
UCA1	43.33	58.42	2.68	18.94
CRNDE	10.36	65.04	9.98	18.88
XIST	1785.75	-1.03	1.38	13.41
HOTAIR	6.33	43.81	7.01	12.48
NRON	7.4	16.89	1.43	5.63
LINC00312	7.17	6.12	2.77	4.95
HOXA11-AS	6.16	5.83	2.46	4.45
LINC01234	-3.34	3.06	43.4	3.41
GACAT1	1.91	2.66	6.15	3.15
PVT1	2.05	9.6	1.53	3.11
LINC00152	4.48	5.22	1.21	3.05
LSINCT5	2.4	2.42	3.48	2.72
TERC	2.68	2.18	2.81	2.54
DLEU2	2.62	2.16	2.74	2.49
GAS6-AS1	-1.75	6.03	3.63	2.32
DLX6-AS1	2.59	1.91	1.99	2.14
AFAP1-AS1	-1.17	-1.05	10.78	2.06
WT1-AS	-1.08	-1.03	8.87	2
HOTTIP	-1.08	5.63	1.38	1.93
CCAT1	8.77	-6.76	3.71	1.69
TSIX	3.51	-1.03	1.38	1.68
HIF1A-AS1	1.32	1.78	1.94	1.66
HNF1A-AS1	1.5	1.51	2.02	1.66
ACTA2-AS1	2.14	1.66	1.05	1.55
CCAT2	-1.63	10.65	-1.77	1.55
POU5F1P5	2.14	1.47	1.18	1.55

TABLE 2 | List of downregulated genes across the HNC cell line group compared with the normal keratinocyte cell line group.

Down-Regulation (n = 28, comparing to control group)				
Symbol	**SAS**	**OECM1**	**FaDu**	**Geometric mean**
	Fold Regulation	**Fold Regulation**	**Fold Regulation**	**Fold Regulation**
KRASP1	-87.59	-31.38	-77.46	-59.71
IPW	1.07	-101.59	-113.01	-22.08
MEG3	-13.62	-12.97	-9.13	-11.73
MIR155HG	-13.86	-10.28	-10.38	-11.39
ZFAS1	-7.61	-8.15	-23.43	-11.33
SPRY4-IT1	-1.65	-13.56	-22.55	-7.96
LINC00963	-14.45	-3.37	-7.39	-7.11
GAS5	-5.79	-5.47	-10.82	-7
BLACAT1	-12.98	-1.22	-9.07	-5.24
DGCR5	-2.11	-6.01	-3.53	-3.55
LUCAT1	-6.18	-2.63	-1.53	-2.92
MALAT1	-6.72	1.35	-3.41	-2.57
HEIH	-1.92	-3.32	-2.58	-2.54
LINC00887	-6.2	2.11	-4.98	-2.45
HIF1A-AS2	-1.78	1.12	-6.39	-2.16
MIR17HG	-1.02	-1.28	-7.6	-2.15
MIR31HG	-2.63	-2.61	-1.28	-2.07
GNAS-AS1	2.55	-5.38	-3.79	-2
CDKN2B-AS1	-1.93	1.33	-4.89	-1.92
HOTAIRM1	-2.71	-1.29	-1.93	-1.89
HULC	-3.67	1.36	-2.5	-1.89
RMRP	-1.63	-1.58	-2.58	-1.88
CAHM	-1.2	-1.36	-3.51	-1.79
NBR2	-2.63	-1.49	-1.4	-1.76
CBR3-AS1	-3.17	1.68	-2.63	-1.71
AIRN	-1.38	-1.89	-1.89	-1.7
FTX	-1.76	1.29	-2.59	-1.52
NEAT1	-1.56	1.26	-2.84	-1.52

survival. The hazard ratio (HR) of the prognostic association for each lncRNA was summarized in **Figure 2B**. In total, there were 24 lncRNAs with HR ≥ 1.0, implying the higher risk of these lncRNA expressions to be correlated with worse survival, whereas 17 lncRNAs were found with HR < 1.0, alluding to a lower lncRNA level, favoring good prognosis in HNC patients.

Dysregulated lncRNA Signatures in Cells Correlated With Prognosis in Patients

To parallelly assess lncRNA expression level and the prognostic significance in HNC patients, **Figure 3A** was plotted to show the association between these two parameters of each molecule. As shown, a total of 27 lncRNAs exhibited correlative levels of dysregulation and prognostic risk (HRs), with 16 being positive-risk and 11 negative-risk to the prognosis of HNC. To further assess the prognostic prediction power, the statistical significance on the overall survival of these 27 lncRNAs was examined. **Figure 3B** depicts the 15 molecules that exhibited altered expression and statistical correlation (p-value ≤ 0.05) in HNC patients. Of these lncRNAs, 9 molecules were upregulated and associated with poor prognosis, including TERC, LINC01234, CCAT1, XIST, GACAT1, WT1-AS, CCAT2, HOXA11-AS, and TSIX. In contrast, a total of 6 molecules, NEAT1, MALAT1, CDKN2B-AS1, CBR3-AS1, IPW, and AIRN, were downregulated and related to good prognosis.

Prognostic Significance of the Top 30 Up- and Down-Regulated lncRNAs From TCGA-HNSC Database

Apart from the data collected from our PCR array panel, we also analyzed the clinical RNA-seq data from TCGA. Here, we selected the 30 most upregulated and the 30 most downregulated lncRNAs in the TCGA-HNSC cohort

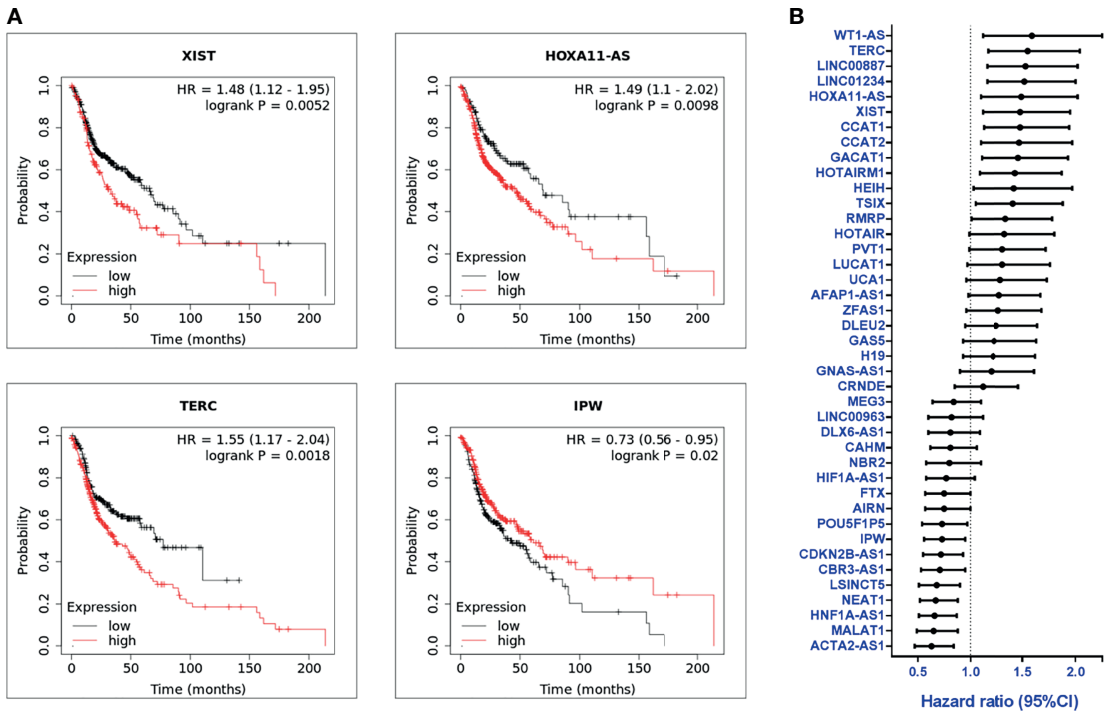


FIGURE 2 | Overall survival and prognosis analysis of profiled lncRNAs **(A)** Kaplan-Meier survival curve examples of dysregulated lncRNAs in HNC. Hazard ratios (HRs) were calculated with 95% confidence intervals (CI), and p-values ≤ 0.05 were considered significant. **(B)** HRs of 41 significantly dysregulated lncRNAs. A total of 41 lncRNAs were examined, post-exclusion of nine lncRNAs from the 55 highly dysregulated candidates with no information provided in KM plotter.

(**Figure 4A**). Out of the top 60 dysregulated lncRNAs, we found 24 lncRNAs with HRs significantly correlating with their expression levels, including 14 upregulated lncRNAs with corresponding HRs ≥ 1.0, and 10 downregulated lncRNAs displaying HRs < 1.0 (**Figure 4B**). Further investigation of these molecules revealed 16 genes with p-values ≤ 0.05, signifying a high correlation with prognosis (**Figure 4C**). The collective correlated HRs and significant p-values of the lncRNAs screened are summarized in **Supplementary Table S3**. In combination with our panel derived from the PCR array results, we established a comprehensive panel of 31-lncRNA signatures associated with HNC prognosis.

A Panel of 31-lncRNA Signature Can Potentially Predict HNC Prognosis

Since the expression levels of the lncRNAs from the TCGA dataset are already established from clinical HNC patients, we wanted to specifically confirm and authenticate the extensive capabilities of the 15 lncRNA signatures screened from our own PCR array sample set through RT-qPCR expression analysis. In addition to the five HNC and normal cell lines used in the PCR array, we also included four additional normal cell lines and seven additional cancer cell line samples. **Figure 5A** shows the expression folds of six panel lncRNAs: HOXA11-AS, IPW, MALAT1, TSIX, WT1-AS, and XIST, with statistically significant dysregulation in HNC cells compared to normal cells. As seen in the example lncRNAs shown, the panel lncRNAs are verified across multiple HNC cell lines to be significantly correlated with our PCR array results. Thus, our defined 31-lncRNA signature panel provides insight on dysregulation in cells, as well as correlation with prognosis in HNC patients.

LncRNA XIST Is Significantly Correlated With HNC

Herein, we selected lncRNA XIST for further functional analysis, due to its significant FR and correlation with TCGA datasets (**Figure 3B**), as well as its high endogenous expression level in HNC cell lines. **Figure 5B** is a schematic representation of our screening and selection. To further verify the potential significance of XIST in cancer, we confirmed its expression in clinical sample data from TCGA (**Supplementary Figure S1A**). Furthermore, we examined its pan-cancer expression. TCGA clinical samples and tissues from various other carcinomas, such as lung, liver/bile duct, and thyroid cancers showed that XIST was upregulated in multiple cancers (**Supplementary Figure S1B**). Additionally, the pan-cancer overall survival for XIST was seen to correlate with poor prognosis, with an HR of 1.2 and a p-value of 0.023, signifying high expression risk, resulting in poor prognosis (**Supplementary Figure S1C**). The data collected from clinical resources coincided

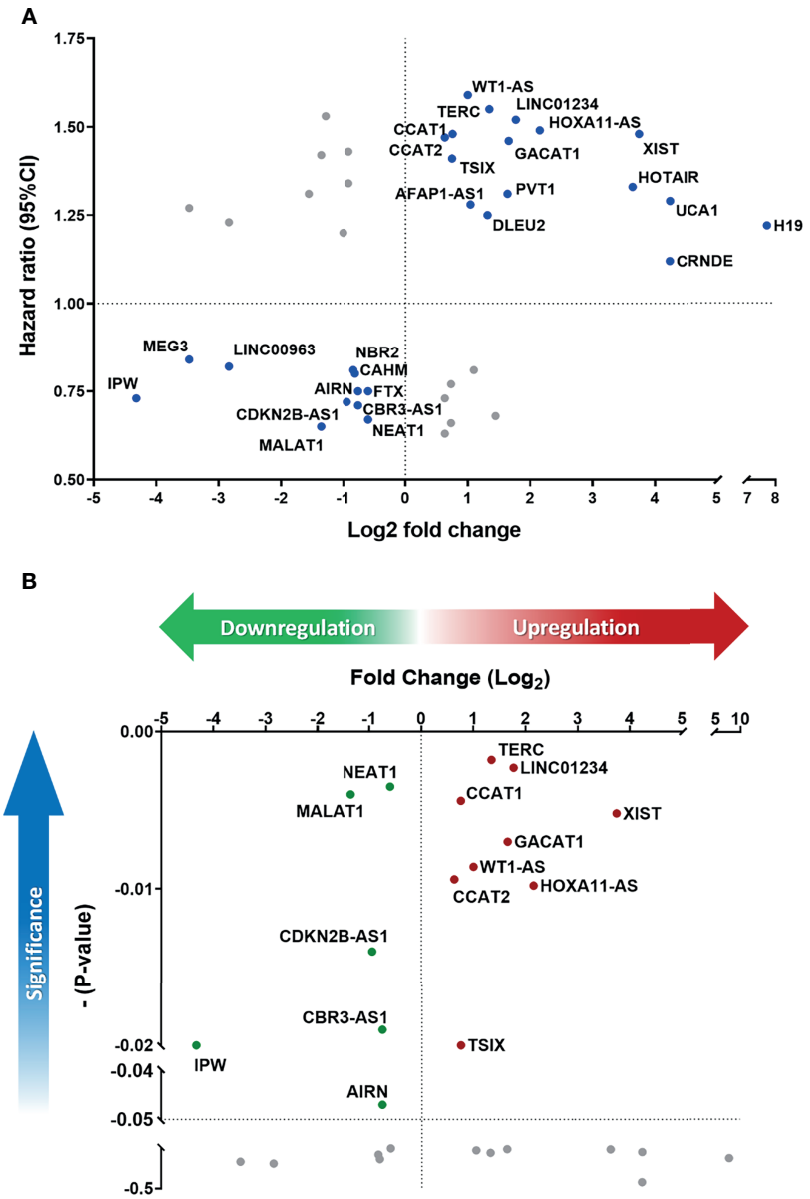


FIGURE 3 | Defining the lncRNA panel based on correlation between gene expression and prognostic information. **(A)** HR versus log_2-fold change (FC) dot-plot of 41 significantly dysregulated lncRNAs from the PCR array. In total, 27 lncRNAs were correlated with high risk score. **(B)** Overall survival p-value versus log_2-FC of 27 positively correlated lncRNAs. A total of 15 lncRNAs were significantly correlated with cancer prognosis (p-value ≤ 0.05).

with our results with XIST in HNC, verifying that XIST, as well as our other panel lncRNAs could be significant for prognostic analysis of HNC.

To acquire comprehensive information related to XIST modulated functional pathway, we performed Gene Ontology (GO) and KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis and found that many of the genes associated with XIST participate in pathways and functions related to cancer metastasis, including various adhesion-specific functions (**Figures 6A, B**).

Silencing of XIST was performed with siRNA as the cancer function model (**Figure 7A**). While analysis of long-term cell growth *via* colony formation assay did not show any significant increase or decrease of colony formation ability across different cell lines (**Figure 7B**), both migration and invasion abilities were prominently inhibited when XIST was silenced (**Figures 7C, D**). Migration was partially inhibited in SAS cell lines, while CGHNC9 and FaDu cells had at least a 40% inhibition rate. Similarly, all three cell lines exhibited highly repressed invasion rates in the siRNA group.

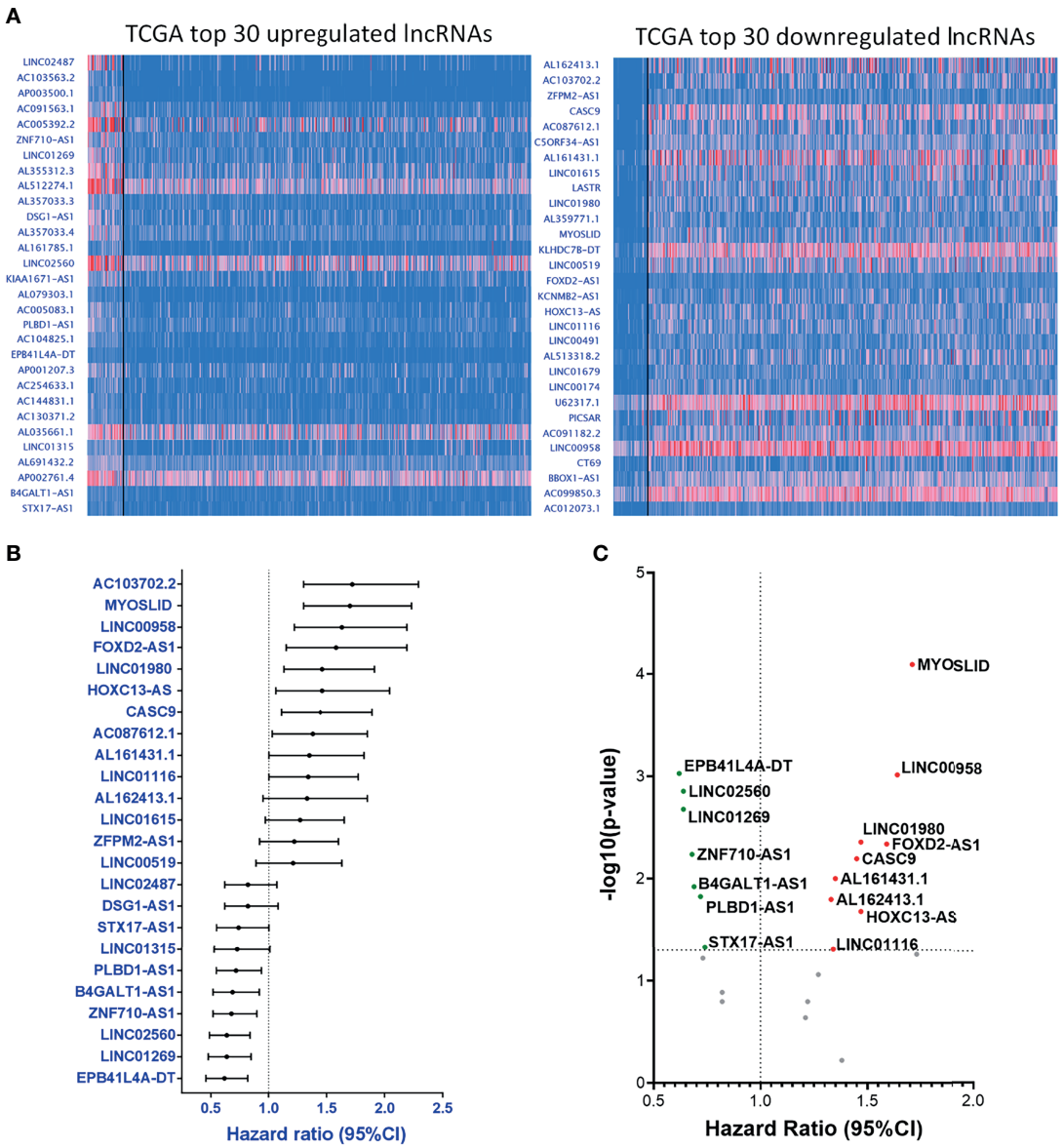


FIGURE 4 | Parallel prognostic significance analysis of clinical samples based on the top 60 dysregulated lncRNA expressions from TCGA-HNSC dataset by UALCAN. **(A)** RNA-seq data of the top 30 upregulated (left) and downregulated (right) lncRNAs were analyzed and depicted through the UALCAN resource. **(B)** HRs of 24 significantly dysregulated lncRNAs. A total of 34 lncRNAs were examined, post-exclusion of lncRNAs with incomplete prognostic information. **(C)** -(Log_{10}-p-value) versus HR volcano plot of the 24 significantly dysregulated lncRNAs. A total of 16 lncRNAs were significantly correlated with cancer prognosis (p-value ≤ 0.05).

To analyze the downstream mechanisms of XIST-mediated migration and invasion in HNC, we performed RT-qPCR and Western blotting of molecules associated with EMT, including MMP2, MMP7, MMP9, E-cadherin, and N-cadherin. RT-qPCR analysis showed that silencing of XIST also decreased the levels of MMPs and mesenchymal markers, while the epithelial marker E-cadherin was significantly increased across all three cell lines (**Figure 8A**). The protein levels of these genes were also similarly affected. While N-cadherin were inhibited by the knockdown of XIST, E-cadherin was significantly upregulated (**Figure 8B**). Thus, as demonstrated by the analysis of XIST, our panel of lncRNAs can potentially be effectively used as biomarkers that can predict prognosis for HNC.

DISCUSSION

Cancer has recently become the most common cause of death in higher-income countries. Therefore, it is of utmost importance to find precise molecules that can detect the cancers before it

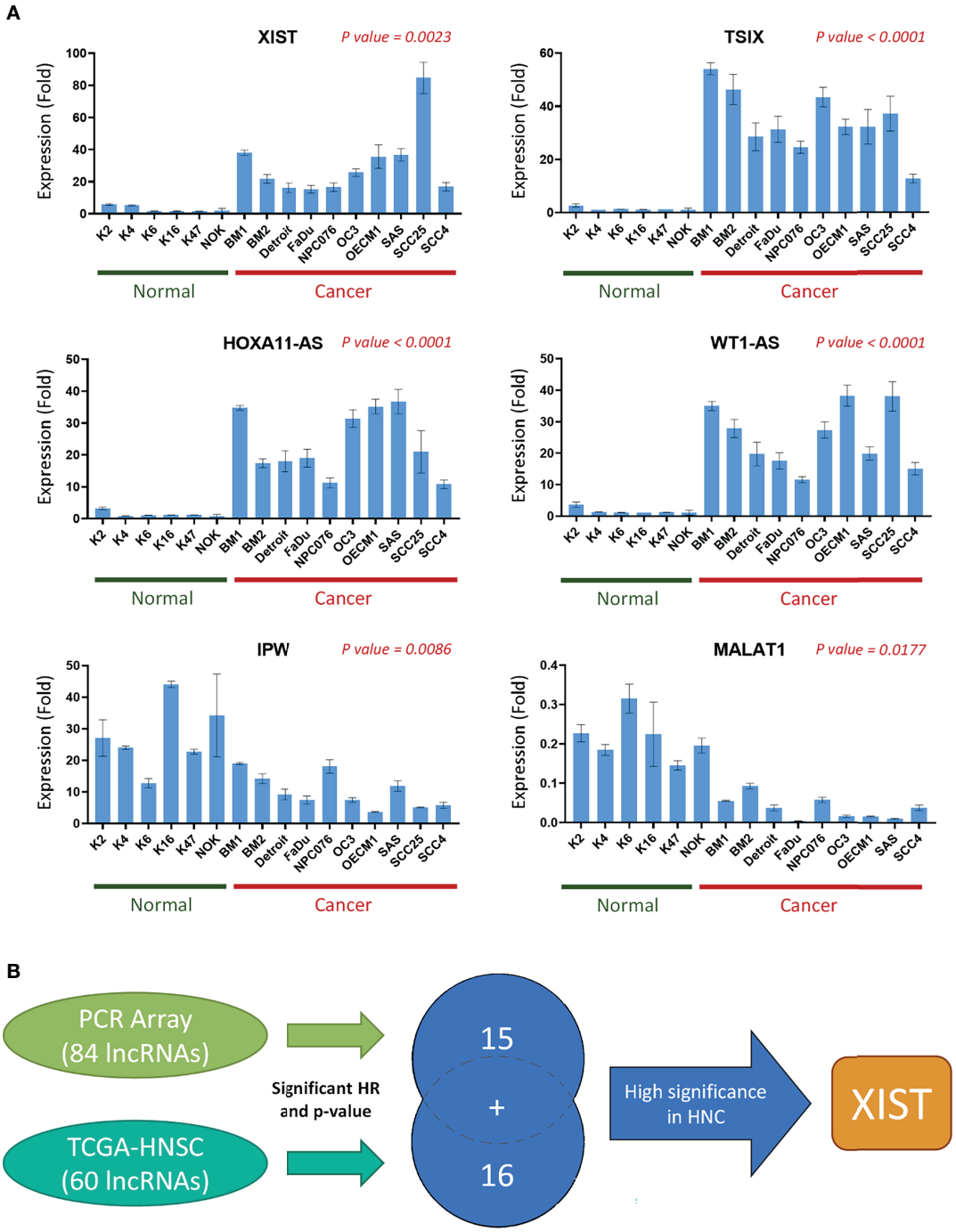


FIGURE 5 | A panel of 31-lncRNA signature predicts the prognosis of HNC. **(A)** The 15 PCR array Panel lncRNAs were verified with various HNC cell lines (n = 10) and normal keratinocytes (n = 6) using RT-qPCR. The six most significantly dysregulated lncRNAs are shown. P-values were calculated using t-test, where p-value ≤ 0.05 was considered significant. **(B)** Schematic flowchart of the systematic screening process. A total of 31 lncRNAs were found to be significantly correlated with HNC prognosis, including 15 lncRNAs found through the PCR array, and 16 found through TCGA database analysis.

reaches the late stages. LncRNAs are a class of non-coding RNAs that have the ability to modify and/or regulate biological activities, which contributes greatly to all types of diseases, including cancers. A variety of lncRNAs have been implicated from previous studies to play roles in various cancers, including HNC (30). However, to the best of our knowledge, no specific prognostic lncRNA(s) have been derived from a systematically verified study. In this study, we designed a comprehensive

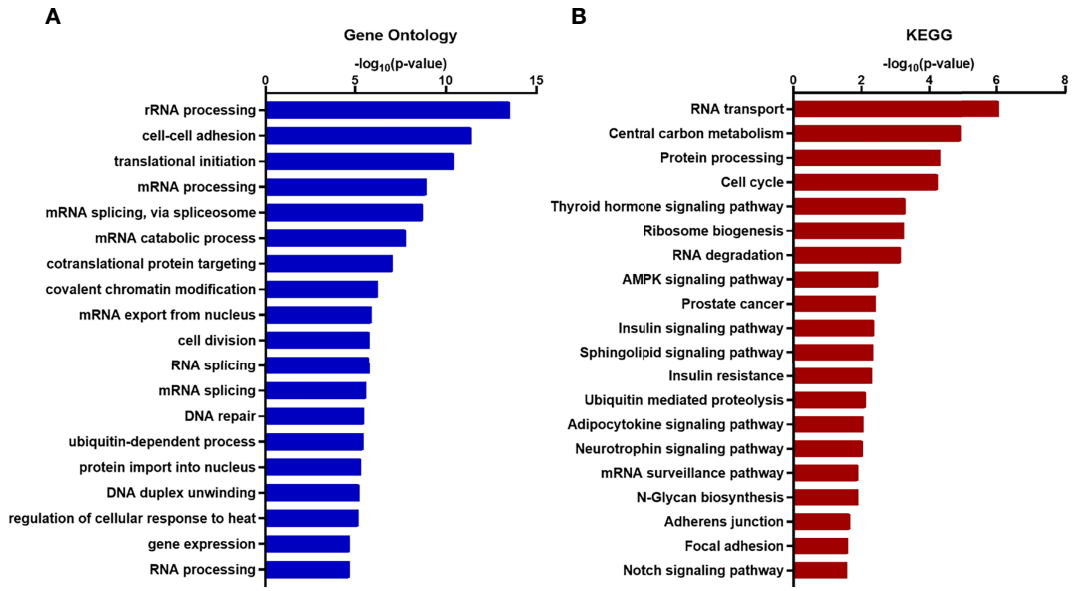


FIGURE 6 | Functional enrichment and annotation of genes associated with XIST *via* bioinformation software. **(A)** GO annotation and **(B)** KEGG pathway enrichment analysis of XIST genes shown in blue and red, respectively. The original significance values obtained from DAVID were transformed to "-log (p-value)" for plotting. Functions and/or pathways with the highest significance are shown.

strategy to systematically profile prognosis-associated lncRNAs in HNC. A few highlights are noted from our work. [1] Profiling of 84 lncRNAs was performed with a PCR array panel. [2] A total of 55 lncRNAs were found to be highly dysregulated in HNC, with 27 upregulated and 28 downregulated genes. [3] A panel of 31-signature prognosis-associated lncRNAs in HNC was defined. [4] XIST was demonstrated as a critical lncRNA molecule in carcinogenic functions, such as cell migration and invasion. Thus, our defined panel of lncRNAs can be used as potential HNC prognostic markers.

After validating the 84 cancer-associated lncRNAs with PCR array using the criteria $|FR| \geq 1.5$, we identified 55 dysregulated lncRNAs in HNC. Many of these lncRNAs have appeared across multiple previous studies. A cancer lncRNA consensus by Carlevaro-Fita et al. listed H19, HOTAIR, MALAT1, and MEG3 as the most prolific lncRNAs, all of which were consistent with our results (31). A review by Cossu et al. also pointed out various lncRNAs, such as AFAP1-AS1, PVT1, MALAT1, H19, DLEU2, CCAT1, and more, that are potentially associated with HNC, many of which also agreed with our findings (32).

Following our initial screening, we investigated the prognostic abilities of these lncRNAs. We chose to evaluate prognosis through HR and overall survival, as the risk of disease in conjunction with time represents imperative determining factors of cancer progression. Univariate cox proportional HRs have been used by multiple studies to represent prognosis potential, as it estimates the relative risk of each lncRNA (33). The results of our HR analysis showed 27 candidates with significant prognosis implications. Then, utilizing overall survival analysis (p-value ≤ 0.05), we evaluated the significance between lncRNA expression and cancer survival. Here, we discovered 15 prognosis-associated lncRNAs. To broaden the scope of our study to include data based on clinical samples, we analyzed highly dysregulated lncRNAs from the TCGA-HNSC dataset in conjunction with the PCR array screening results. A total of 16 lncRNAs were found to be significant in the cancer survival and progression of HNC patients. Altogether, we established a 31-lncRNA signature panel that predicts HNC prognosis.

Upon further detailed investigation, we found some notable lncRNAs in our panel, including XIST, TSIX, HOXA11-AS, WT1-AS, IPW, and MALAT1, with significant dysregulation in multiple HNC cell lines. A study by Yao et al. also identified HOXA11-AS and MALAT1 as potential biomarkers for HNC (34). The prognostic risk of MALAT1 has also been well established in various previous studies (32, 35). Interestingly, although studies have deemed MALAT1 as an oncogene across many cancer types (35), our results indicated that it was downregulated in HNC. TCGA data analysis also showed that high expression of MALAT1 resulted in higher overall survival, which correlates with our study. Thus, these common lncRNAs may have high potential for future HNC-specific studies. On the other hand, some lncRNAs from our panel are relatively novel lncRNAs, such as WT1-AS, where only a handful of studies have proposed its carcinogenic function in lung, cervical, and breast cancer (36–38). Knowledge regarding HOXA11-AS is also sparse, although recent studies have elaborated on its role in liver (39), lung (40), head-neck (34), and other cancers (41). Not much is known about lncRNA IPW and TSIX either, but some preliminary studies have pointed out potential interactions between TSIX and the more well-known

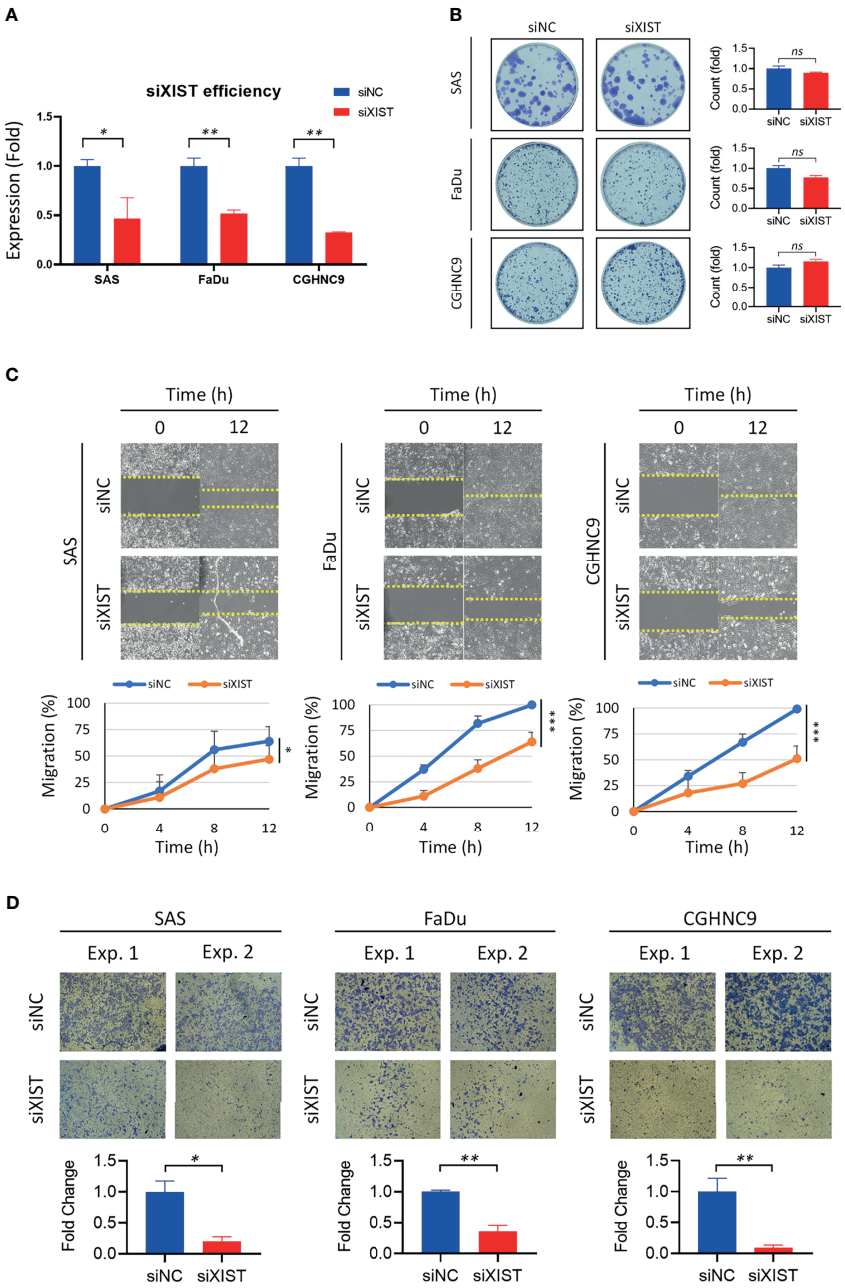


FIGURE 7 | Validation and functional analysis of XIST as an HNC panel biomarker. **(A)** siRNA efficiency of XIST silencing was examined with RT-qPCR. Optimal knockdown was achieved by transfection in SAS, FaDu, and CGHNC9 cell lines. **(B)** Colony formation ability was determined after successful silencing of XIST. No significant difference was observed between the three cancer cell lines when compared to the normal keratinocytes. **(C)** The wound-healing model was used for migration assay. SAS was partially inhibited by roughly 20%, while FaDu and CGHNC9 was inhibited by at least 60%. **(D)** Invasion ability was determined *via* Matrigel invasion assay. All three cell lines showed statistically significant inhibition rates in the XIST knockdown group. All functional experiments were performed in triplicates. (***p ≤ 0.001, **p ≤ 0.01, *p ≤ 0.05, t-test, ns = not significant).

lncRNA XIST in regards to X chromosome modulation (42). Although some early studies predicted an inverse correlation between these two molecules (43), newer studies began to disprove their correlation, focusing on their individual functions instead (44).

Numerous studies have linked XIST with multiple cancer types, such as colorectal (45), lung (46), and breast cancer (47). Various cancers such as thyroid (48) and osteoscaroma (49) have also shown XIST to act as an oncogene, which coincides with our findings in HNC. Many of these studies have also suggested that

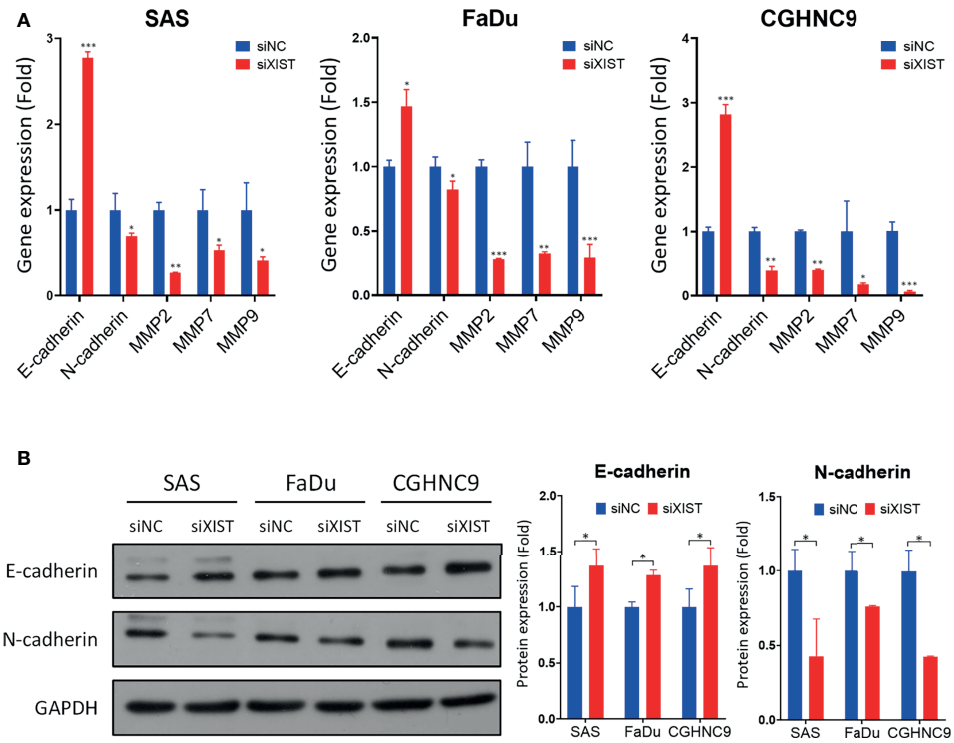


FIGURE 8 | Downstream molecular analysis of XIST-mediated migration and invasion dysregulation. **(A)** RT-qPCR of EMT-associated genes were performed in SAS, FaDu, and CGHNC9 cell lines post-XIST silencing. MMP and mesenchymal markers were downregulated, while the epithelial marker E-cadherin was upregulated. **(B)** Protein expression of EMT markers were analyzed through western blotting. Quantification was achieved using imageJ software. (***p ≤ 0.001, **p ≤ 0.01, *p ≤ 0.05, t-test).

XIST could act as a prognostic marker, given its various cancerous functions. Thus, we selected the lncRNA XIST for *in vitro* studies due to its performance in our screening results, along with its novelty in HNC. Additionally, various annotations from GO and KEGG, such as 'adherens junction', 'focal adhesion', and 'cell-cell adhesion', all strongly allude to the metastatic functions as well, which correlates with our functional analysis. Our results showed that XIST played essential roles in regulating cell migration and invasion, as silencing by siRNA significantly inhibited these functions in multiple HNC cells. These carcinogenic roles were also confirmed in previous cancer studies, such as liver (50), ovarian (51), and esophageal cancer (52). A review by Zhou et al. also highlighted the potential of XIST as a prognostic marker (53). Therefore, our findings implied that because XIST affected HNC progression through these metastatic functions, it can be used as a prognostic marker. Taken together, XIST can be used as an HNC marker that can predict prognosis, and can potentially be used as therapeutic target.

In conclusion, we established a systematic profiling method to screen for prognostic lncRNA markers in HNC. Our results from the *in silico* and *in vitro* combination analysis suggests that the panel of 31-lncRNA signatures may contribute to HNC tumorigenesis, and can provide valuable prognostic data. Additionally, XIST demonstrated carcinogenic functions in HNC, implicating its ability as a prognostic marker. Overall, our findings greatly contribute to the knowledge of HNC and prognosis, and can potentially be expanded to applications in precision medicine.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

S-JT was responsible for experimental analysis, acquisition of data, and the writing of the manuscript. A-JC and S-JT were responsible for the conceptualization and design. A-JC and JC were responsible for resources, project administration, supervision and funding acquisition. S-JT, A-JC, and G-RY were responsible for the review of the submitted manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Ministry of Science and Technology (Most-107-2314-B-182A-062-MY3), and Chang Gung Memorial Hospital – Linkou Medical Center (CMRPG3K1571).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.731752/full#supplementary-material

REFERENCES

1. Wu C, Gleysteen J, Teraphongphom NT, Li Y, Rosenthal E. In-Vivo Optical Imaging in Head and Neck Oncology: Basic Principles, Clinical Applications and Future Directions. *Int J Oral Sci* (2018) 10(2):10. doi: 10.1038/s41368-018-0011-4
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
3. Lo Nigro C, Denaro N, Merlotti A, Merlano M. Head and Neck Cancer: Improving Outcomes With a Multidisciplinary Approach. *Cancer Manag Res* (2017) 9:363–71. doi: 10.2147/CMAR.S115761
4. Hawkes N. Cancer Survival Data Emphasise Importance of Early Diagnosis. *BMJ* (2019) 364:l408. doi: 10.1136/bmj.l408
5. Gerstner AOH. Early Detection in Head and Neck Cancer - Current State and Future Perspectives. *GMS Curr Topics Otorhinolaryngol Head Neck Surg* (2008) 7:Doc06–Doc.
6. Balas MM, Johnson AM. Exploring the Mechanisms Behind Long Noncoding RNAs and Cancer. *Non-Coding RNA Res* (2018) 3(3):108–17. doi: 10.1016/j.ncrna.2018.03.001
7. Jarroux J, Morillon A, Pinskaya M. History, Discovery, and Classification of lncRNAs. In: Rao MRS, editor. *Long Non Coding RNA Biology*. Singapore: Springer Singapore (2017). p. 1–46.
8. Browne CJ, Godino A, Salery M, Nestler EJ. Epigenetic Mechanisms of Opioid Addiction. *Biol Psychiatry* (2020) 87(1):22–33. doi: 10.1016/j.biopsych.2019.06.027
9. Postnikova OA, Rogozin IB, Samuel W, Nudelman G, Babenko VN, Poliakov E, et al. Volatile Evolution of Long Non-Coding RNA Repertoire in Retinal Pigment Epithelium: Insights From Comparison of Bovine and Human RNA Expression Profiles. *Genes* (2019) 10(3):205. doi: 10.3390/genes10030205
10. Luo X, Qiu Y, Jiang Y, Chen F, Jiang L, Zhou Y, et al. Long non-Coding RNA Implicated in the Invasion and Metastasis of Head and Neck Cancer: Possible Function and Mechanisms. *Mol Cancer* (2018) 17(1):14. doi: 10.1186/s12943-018-0763-7
11. Rajagopal T, Talluri S, Akshaya RL, Dunna NR. HOTAIR LncRNA: A Novel Oncogenic Propellant in Human Cancer. *Clin Chim Acta* (2020) 503:1–18. doi: 10.1016/j.cca.2019.12.028
12. Yao F, Wang Q, Wu Q. The Prognostic Value and Mechanisms of lncRNA UCA1 in Human Cancer. *Cancer Manag Res* (2019) 11:7685–96. doi: 10.2147/CMAR.S200436
13. Wang Y, Wang S, Ren Y, Zhou X. The Role of lncRNA Crosstalk in Leading Cancer Metastasis of Head and Neck Squamous Cell Carcinoma. *Front Oncol* (2020) 10:561833–. doi: 10.3389/fonc.2020.561833
14. Wu M, Shang X, Sun Y, Wu J, Liu G. Integrated Analysis of Lymphocyte Infiltration-Associated lncRNA for Ovarian Cancer *via* TCGA, GTEx and GEO Datasets. *PeerJ* (2020) 8:e8961–e. doi: 10.7717/peerj.8961
15. Wang P, Jin M, Sun C-H, Yang L, Li Y-S, Wang X, et al. A three-lncRNA Expression Signature Predicts Survival in Head and Neck Squamous Cell Carcinoma (HNSCC). *Biosci Rep* (2018) 38(6):BSR20181528. doi: 10.1042/BSR20181528
16. Zhang G, Fan E, Zhong Q, Feng G, Shuai Y, Wu M, et al. Identification and Potential Mechanisms of a 4-lncRNA Signature That Predicts Prognosis in Patients With Laryngeal Cancer. *Hum Genomics* (2019) 13(1):36–. doi: 10.1186/s40246-019-0230-6
17. Zhong G, Lou W, Yao M, Du C, Wei H, Fu P. Identification of Novel mRNA-miRNA-lncRNA Competing Endogenous RNA Network Associated With Prognosis of Breast Cancer. *Epigenomics* (2019) 11(13):1501–18. doi: 10.2217/epi-2019-0209
18. Lin T, Fu Y, Zhang X, Gu J, Ma X, Miao R, et al. A Seven-Long Noncoding RNA Signature Predicts Overall Survival for Patients With Early Stage Non-Small Cell Lung Cancer. *Aging (Albany NY)* (2018) 10(9):2356–66. doi: 10.18632/aging.101550
19. Liu H, Zhang Q, Lou Q, Zhang X, Cui Y, Wang P, et al. Differential Analysis of lncRNA, miRNA and mRNA Expression Profiles and the Prognostic Value of lncRNA in Esophageal Cancer. *Pathol Oncol Res* (2020) 26(2):1029–39. doi: 10.1007/s12253-019-00655-8
20. Liu G, Zheng J, Zhuang L, Lv Y, Zhu G, Pi L, et al. A Prognostic 5-lncRNA Expression Signature for Head and Neck Squamous Cell Carcinoma. *Sci Rep* (2018) 8(1):15250. doi: 10.1038/s41598-018-33642-1
21. Cao W, Liu JN, Liu Z, Wang X, Han ZG, Ji T, et al. A three-lncRNA Signature Derived From the Atlas of ncRNA in Cancer (TANRIC) Database Predicts the Survival of Patients With Head and Neck Squamous Cell Carcinoma. *Oral Oncol* (2017) 65:94–101. doi: 10.1016/j.oraloncology.2016.12.017
22. Ni J, Liu S, Qi F, Li X, Yu S, Feng J, et al. Screening TCGA Database for Prognostic Genes in Lower Grade Glioma Microenvironment. *Ann Transl Med* (2020) 8(5):209–. doi: 10.21037/atm.2020.01.73
23. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* (2018) 173(2):400–16.e11. doi: 10.1016/j.cell.2018.02.052
24. Yu K, Chen B, Aran D, Charalel J, Yau C, Wolf DM, et al. Comprehensive Transcriptomic Analysis of Cell Lines as Models of Primary Tumors Across 22 Tumor Types. *Nat Commun* (2019) 10(1):3574. doi: 10.1038/s41467-019-11415-2
25. Boone DR, Micci M-A, Taglialatela IG, Hellmich JL, Weisz HA, Bi M, et al. Pathway-Focused PCR Array Profiling of Enriched Populations of Laser Capture Microdissected Hippocampal Cells After Traumatic Brain Injury. *PloS One* (2015) 10(5):e0127287–e. doi: 10.1371/journal.pone.0127287
26. Zhang B, Wang Q, Miao T, Yu B, Yuan P, Kong J, et al. Whether Alzheimer's Diseases Related Genes Also Differently Express in the Hippocampus of Ts65Dn Mice? *Int J Clin Exp Pathol* (2015) 8(4):4120–5.
27. Lu Y-C, Cheng A-J, Lee L-Y, You G-R, Li Y-L, Chen H-Y, et al. MiR-520b as a Novel Molecular Target for Suppressing Stemness Phenotype of Head-Neck Cancer by Inhibiting CD44. *Sci Rep* (2017) 7(1):2042. doi: 10.1038/s41598-017-02058-8
28. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia (New York NY)* (2017) 19(8):649–58. doi: 10.1016/j.neo.2017.05.002
29. Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H. Starbase V2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and Protein-RNA Interaction Networks From Large-Scale CLIP-Seq Data. *Nucleic Acids Res* (2014) 42(Database issue):D92–D7. doi: 10.1093/nar/gkt1248
30. Zhou RS, Zhang EX, Sun QF, Ye ZJ, Liu JW, Zhou DH, et al. Integrated Analysis of lncRNA-miRNA-mRNA ceRNA Network in Squamous Cell Carcinoma of Tongue. *BMC Cancer* (2019) 19(1):779. doi: 10.1186/s12885-019-5983-8
31. Carlevaro-Fita J, Lanzós A, Feuerbach L, Hong C, Mas-Ponte D, Pedersen JS, et al. Cancer LncRNA Census Reveals Evidence for Deep Functional Conservation of Long Noncoding RNAs in Tumorigenesis. *Commun Biol* (2020) 3(1):56. doi: 10.1038/s42003-019-0741-7
32. Cossu AM, Mosca L, Zappavigna S, Misso G, Bocchetti M, De Micco F, et al. Long Non-Coding RNAs as Important Biomarkers in Laryngeal Cancer and Other Head and Neck Tumours. *Int J Mol Sci* (2019) 20(14):3444. doi: 10.3390/ijms20143444
33. Barraclough H, Simms L, Govindan R. Biostatistics Primer: What a Clinician Ought to Know: Hazard Ratios. *J Thorac Oncol* (2011) 6(6):978–82. doi: 10.1097/JTO.0b013e31821b10ab
34. Yao Y, Chen X, Lu S, Zhou C, Xu G, Yan Z, et al. Circulating Long Noncoding RNAs as Biomarkers for Predicting Head and Neck Squamous Cell

Carcinoma. *Cell Physiol Biochem* (2018) 50(4):1429–40. doi: 10.1159/000494605
35. Li ZX, Zhu QN, Zhang HB, Hu Y, Wang G, Zhu YS. MALAT1: A Potential Biomarker in Cancer. *Cancer Manag Res* (2018) 10:6757–68. doi: 10.2147/CMAR.S169406
36. Wu C, Yang J, Li R, Lin X, Wu J, Wu J. LncRNA WT1-AS/miR-494-3p Regulates Cell Proliferation, Apoptosis, Migration and Invasion *via* PTEN/PI3K/AKT Signaling Pathway in Non-Small Cell Lung Cancer. *Onco Targets Ther* (2021) 14:891–904. doi: 10.2147/OTT.S278233
37. Zhang Y, Na R, Wang X. LncRNA WT1-AS Up-Regulates P53 to Inhibit the Proliferation of Cervical Squamous Carcinoma Cells. *BMC Cancer* (2019) 19 (1):1052. doi: 10.1186/s12885-019-6264-2
38. Zhang Y, Fan LJ, Zhang Y, Jiang J, Qi XW. Long Non-Coding Wilms Tumor 1 Antisense RNA in the Development and Progression of Malignant Tumors. *Front Oncol* (2020) 10:35. doi: 10.3389/fonc.2020.00035
39. Guo JC, Yang YJ, Zheng JF, Zhang JQ, Guo M, Yang X, et al. Silencing of Long Noncoding RNA HOXA11-AS Inhibits the Wnt Signaling Pathway *via* the Upregulation of HOXA11 and Thereby Inhibits the Proliferation, Invasion, and Self-Renewal of Hepatocellular Carcinoma Stem Cells. *Exp Mol Med* (2019) 51(11):1–20. doi: 10.1038/s12276-019-0328-x
40. Zhao X, Li X, Zhou L, Ni J, Yan W, Ma R, et al. LncRNA HOXA11-AS Drives Cisplatin Resistance of Human LUAD Cells *via* Modulating miR-454-3p/Stat3. *Cancer Sci* (2018) 109(10):3068–79. doi: 10.1111/cas.13764
41. Wei C, Zhao L, Liang H, Zhen Y, Han L. Recent Advances in Unraveling the Molecular Mechanisms and Functions of HOXA11–AS in Human Cancers and Other Diseases (Review). *Oncol Rep* (2020) 43(6):1737–54. doi: 10.3892/or.2020.7552
42. Weakley SM, Wang H, Yao Q, Chen C. Expression and Function of a Large Non-Coding RNA Gene XIST in Human Cancer. *World J Surg* (2011) 35 (8):1751–6. doi: 10.1007/s00268-010-0951-0
43. Gayen S, Maclary E, Buttigieg E, Hinten M, Kalantry S. A Primary Role for the Tsix lncRNA in Maintaining Random X-Chromosome Inactivation. *Cell Rep* (2015) 11(8):1251–65. doi: 10.1016/j.celrep.2015.04.039
44. Habieb A, Matboli M, El-Tayeb H, El-Asmar F. Potential Role of lncRNA-TSIX, miR-548-a-3p, and SOGA1 mRNA in the Diagnosis of Hepatocellular Carcinoma. *Mol Biol Rep* (2019) 46(4):4581–90. doi: 10.1007/s11033-019-04810-x
45. Yang X, Zhang S, He C, Xue P, Zhang L, He Z, et al. METTL14 Suppresses Proliferation and Metastasis of Colorectal Cancer by Down-Regulating Oncogenic Long Non-Coding RNA XIST. *Mol Cancer* (2020) 19(1):46–. doi: 10.1186/s12943-020-1146-4
46. Liu J, Yao L, Zhang M, Jiang J, Yang M, Wang Y. Downregulation of LncRNA-XIST Inhibited Development of Non-Small Cell Lung Cancer by Activating miR-335/SOD2/ROS Signal Pathway Mediated Pyroptotic Cell Death. *Aging (Albany NY)* (2019) 11(18):7830–46. doi: 10.18632/aging.102291
47. Salama EA, Adbeltawab RE, El Tayebi HM. XIST and TSIX: Novel Cancer Immune Biomarkers in PD-L1-Overexpressing Breast Cancer Patients. *Front Oncol* (2020) 9:1459–. doi: 10.3389/fonc.2019.01459
48. Liu H, Deng H, Zhao Y, Li C, Liang Y. LncRNA XIST/miR-34a Axis Modulates the Cell Proliferation and Tumor Growth of Thyroid Cancer Through MET-PI3K-AKT Signaling. *J Exp Clin Cancer Res* (2018) 37(1):279–. doi: 10.1186/s13046-018-0950-9
49. Deng C, Hu X, Wu K, Tan J, Yang C. Long non-Coding RNA XIST Predicting Advanced Clinical Parameters in Cancer: A Meta-Analysis and Case Series Study in a Single Institution. *Oncol Lett* (2019) 18(3):2192–202. doi: 10.3892/ol.2019.10592
50. Ning D, Chen J, Du P, Liu Q, Cheng Q, Li X, et al. The Crosstalk Network of XIST/miR-424-5p/OGT Mediates RAF1 Glycosylation and Participates in the Progression of Liver Cancer. *Liver Int* (2021) 41:1933–44. doi: 10.1111/liv.14904
51. Jiang R, Zhang H, Zhou J, Wang J, Xu Y, Zhang H, et al. Inhibition of Long Non-Coding RNA XIST Upregulates microRNA-149-3p to Repress Ovarian Cancer Cell Progression. *Cell Death Dis* (2021) 12(2):145–. doi: 10.1038/s41419-020-03358-0
52. Chen Z, Hu X, Wu Y, Cong L, He X, Lu J, et al. Long non-Coding RNA XIST Promotes the Development of Esophageal Cancer by Sponging miR-494 to Regulate CDK6 Expression. *Biomed Pharmacother* (2019) 109:2228–36. doi: 10.1016/j.biopha.2018.11.049
53. Zhou Q, Hu W, Zhu W, Zhang F, Lin-lin L, Liu C, et al. Long Non Coding RNA XIST as a Prognostic Cancer Marker – A Meta-Analysis. *Clin Chim Acta* (2018) 482:1–7. doi: 10.1016/j.cca.2018.03.016

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 19 October 2021
doi: 10.3389/fonc.2021.734655

CDCA7 Facilitates Tumor Progression by Directly Regulating *CCNA2* Expression in Esophageal Squamous Cell Carcinoma

Hongyi Li[1,2†], Yongjia Weng[1,2†], Shaojie Wang[1,2], Fang Wang[1,2], Yanqiang Wang[1,2], Pengzhou Kong[1,2], Ling Zhang[1,2], Caixia Cheng[3], Heyang Cui[1,2], Enwei Xu[4], Shuqing Wei[5], Dinghe Guo[1,2], Fei Chen[1,2], Yanghui Bi[6], Yongsheng Meng[7], Xiaolong Cheng[1,2] and Yongping Cui[1,2*]*

[1] Department of Pathology & Shanxi Key Laboratory of Carcinogenesis and Translational Research of Esophageal Cancer, Shanxi Medical University, Taiyuan, China, [2] Key Laboratory of Cellular Physiology, Ministry of Education, Shanxi Medical University, Taiyuan, China, [3] Department of Pathology, the First Hospital, Shanxi Medical University, Taiyuan, China, [4] Department of Pathology, Shanxi Province Cancer Hospital, Taiyuan, China, [5] Department of Thoracic Surgery (I), Shanxi Province Cancer Hospital, Taiyuan, China, [6] The Science Research Center, Shanxi Bethone Hospital, Taiyuan, China, [7] Tumor Biobank, Shanxi Province Cancer Hospital, Taiyuan, China

Background: *CDCA7* is a copy number amplified gene identified not only in esophageal squamous cell carcinoma (ESCC) but also in various cancer types. Its clinical relevance and underlying mechanisms in ESCC have remained unknown.

Methods: Tissue microarray data was used to analyze its expression in 179 ESCC samples. The effects of *CDCA7* on proliferation, colony formation, and cell cycle were tested in ESCC cells. Real-time PCR and Western blot were used to detect the expression of its target genes. Correlation of *CDCA7* with its target genes in ESCC and various SCC types was analyzed using GSE53625 and TCGA data. The mechanism of *CDCA7* was studied by chromatin immunoprecipitation (ChIP), luciferase reporter assays, and rescue assay.

Results: The overexpression of *CDCA7* promoted proliferation, colony formation, and cell cycle in ESCC cells. *CDCA7* affected the expression of cyclins in different cell phases. GSE53625 and TCGA data showed *CCNA2* expression was positively correlated with *CDCA7*. The knockdown of *CCNA2* reversed the malignant phenotype induced by *CDCA7* overexpression. Furthermore, *CDCA7* was found to directly bind to *CCNA2*, thus promoting its expression.

Conclusions: Our results reveal a novel mechanism of *CDCA7* that it may act as an oncogene by directly upregulating *CCNA2* to facilitate tumor progression in ESCC.

Keywords: *CDCA7*, cell cycle, *CCNA2*, copy number amplification, ESCC (esophageal squamous cell carcinoma)

OPEN ACCESS

Edited by:
Bin Qiao,
Zhengzhou University, China

Reviewed by:
Jian Zhang,
Texas A&M Health Science Center, United States
Jiheng Xu,
New York University, United States

***Correspondence:**
Yongping Cui
cuiyp@sxmu.edu.cn
Xiaolong Cheng
chengxl@sxmu.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *03 August 2021*
Accepted: *27 September 2021*
Published: *19 October 2021*

Citation:
Li H, Weng Y, Wang S, Wang F, Wang Y, Kong P, Zhang L, Cheng C, Cui H, Xu E, Wei S, Guo D, Chen F, Bi Y, Meng Y, Cheng X and Cui Y (2021) CDCA7 Facilitates Tumor Progression by Directly Regulating CCNA2 Expression in Esophageal Squamous Cell Carcinoma. Front. Oncol. 11:734655. doi: 10.3389/fonc.2021.734655

BACKGROUND

Esophageal cancer which accounts for 11% of diagnosed cancers was the fourth most common cancer type. In China, the dominant histologic type of esophageal cancer is esophageal squamous cell carcinoma (ESCC) which causes more than 175,000 deaths every year (1). The 5-year survival rate of ESCC which ranges from 22% to 30% still tends to be low because of the limitation of technical developments for early diagnosis and treatment (2). However, the advent and progression of next-generation sequencing (NGS) in recent years has given us some achievements on ESCC (3–7).

In our previous WGS analysis of 31 ESCC tumor tissues and matched adjacent non-tumor tissues, we identified some genes with copy number variation, including cell division cycle-associated 7 gene (*CDCA7*), that was amplified in 5 out of 31 ESCC patients (8). *CDCA7* is located on 2q31.1. It is characterized as a c-Myc and E2F responsive gene that participates in neoplastic transformation (9, 10). It has been reported that the expression of *CDCA7* is elevated in a high fraction of human lung, colon, ovary, rectum, stomach, and uterus cancer types, suggesting that *CDCA7* may play a crucial role in cancer development (11–13). A recent study showed that the high expression of *CDCA7* predicted poorer disease-free survival in patients with triple-negative breast cancer (TNBC) and was associated with metastatic relapse status (14). One research in lung adenocarcinoma reported that *CDCA7* promoted lung adenocarcinoma proliferation through regulating the cell cycle, while its mechanism has not been completely elucidated yet (15). Meanwhile, *CDCA7* as a DNA-binding protein can function as a transcription regulator to mediate the tumor-promoting effect (9).

CCNA2, which is synthesized at the beginning of S-phase (16, 17), binds and activates cyclin-dependent kinases (CDK) CDK2 and CDK1, the catalytic partners of CCNA2. The CDK2/CCNA2 complex is the machinery that drives the progression of S-phase. In the S-phase of the cell cycle, the CCNA2–CDK complex can phosphorylate key substances in the process of DNA replication, such as CDC6. This phosphorylation is crucially important for the initiation of DNA replication. It is possible that CCNA2–CDK contributes to tumorigenesis by the phosphorylation of oncoproteins and the increased expression level of *CCNA2* accelerates cell proliferation once the tumor has formed (18–20). Increased expression of *CCNA2* has been observed in various types of cancer such as lung, breast, liver, cervical, and others (18, 21–24). The expression level of *CCNA2* is closely related with cell proliferation; thus, it is used as a proliferation marker for the molecular diagnosis of cancer (18). Meanwhile, the expression of *CCNA2* appears to be of prognostic value for the prediction of survival and early relapse in many types of cancer (18, 25).

In our study, we analyzed the copy number amplification data from The Cancer Genome Atlas (TCGA) database in various types of tumors and the correlation between *CDCA7* expression level and clinical variables in ESCC using the mRNA expression data from the GEO database. Furthermore, we verified that *CDCA7* has as a tumor-promoting role in ESCC, and elaborated on its potential mechanisms of carcinogenesis. Our results show that CDCA7 may bind to *CCNA2* to upregulate its expression. Therefore, increased CCNA2 promotes the proliferation of ESCC cells, thus promoting tumor growth. Our study provides useful clues for more effective therapeutic strategies against ESCC.

METHODS

Clinical Samples

The copy number data were obtained from our study. The tumor and the matched adjacent non-tumor samples were recruited from Shanxi Cancer Hospital of Shanxi Medical University. The patients were without preoperative chemotherapy, radiotherapy, and other treatments before operation, and written consent was obtained from all of them. Hematoxylin and eosin (H&E) staining was used to diagnose these tissues, and the diagnosis was performed by at least two pathologists independently. The ESCC individuals were staged according to the American Joint Commission for Cancer (AJCC)/International Union Against Cancer (UICC) TNM staging system (eighth edition). The study was approved by the Institutional Reviewing Board (IRB) and the Research Committee of Shanxi Medical University.

Cell Lines and Cell Culture

ESCC cell lines KYSE150, KYSE180, KYSE450, and TE-1 and immortal embryonic esophageal epithelium cell lines NE3 and HET-1A used in the research were purchased from the Cell Bank of Type Culture Collection of the Chinese Academy of Sciences. The cell line 293T was from our lab. The cell lines KYSE150, KYSE180, KYSE450, and TE-1 were cultured in HyClone™ RPMI-1640 medium, and the cell lines HET-1A and 293T were cultured in HyClone™ DMEM/High Glucose medium (GE Healthcare Life Sciences, HyClone Laboratories, Logan, UT, USA). The culture was with 10% fetal bovine serum (FBS; Gibco; Thermo Fisher Scientific, Inc., Waltham, MA, USA). The cell line NE3 was cultured in a 1:1 mixture of EpiLife medium (Cascade Biologics, Inc., Portland, OR, USA) and defined keratinocyte serum-free medium (dKSFM; Gibco; Thermo Fisher Scientific, Inc., Waltham, MA, USA). All of the cell lines were cultured at 37°C, 5% CO_2. The culture medium was replaced according to the cell state. Subculture was carried out when the cell fusion was about 80%–90%.

Overexpression and Knockdown of *CDCA7* in ESCC Lines

SiRNAs or plasmids were transfected into the cells at the logarithmic growth phase using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA) according to the instructions of the manufacturer. For knockdown of endogenous *CDCA7*, we used vectors containing the sequence 5′-GCCCTCAGAGAATT CTGTGACTGAT-3′ (*CDCA7*-si1) and 5′-CATCCGTGAC CCTTCCGCATATAAT-3′ (*CDCA7*-si2). These shRNAs were cloned into the vector pHBLV-U6-Scramble-ZsGreen-Puro vector. For stable overexpression, the coding sequence (CDS) region of the *CDCA7* gene was cloned into pHBLV-CMV-MCS-

3FLAG-EF1-ZsGreen-T2A-PURO. The recombinant plasmids and the packaging plasmids (Hanbio Biotechnology Co., Ltd., Shanghai, China) were co-transfected into 293T cells. The lentivirus supernatant was used to infect the KYSE150, KYSE450, and KYSE180 cell lines. The negative control was the corresponding empty vectors. The KYSE150 and KYSE450 knockdown stable cell lines and the KYSE180 overexpression stable cell line were screened out for 7–14 days with 1.0, 1.0, and 0.8 μg/ml puromycin (Invitrogen; Thermo Fisher Scientific, Inc.), respectively. The efficiency of knockdown and overexpression was determined by real-time PCR and Western blot assay. We used small interference RNA (siRNA) for *CCNA2* knockdown, and the siRNA sequence information is as follows: *CCNA2*-si1, 5′-CTATGGACATGTCAATTGT-3′; *CCNA2*-si2, 5′-GAGTGTTAATGAAGTACCA-3′. The CDS of *CDCA7* and *CCNA2* genes was cloned into the pcDNA3.1 vector with a V5 tag and 6*His tag.

MTT Assay

The MTT assay was performed using a 96-well plate with 5×10^3 transfected cells each well and cultured for 24–120 h. A 20-μl MTT solution (5 mg/ml) was added to a 200-μl culture medium each well for 4 h at 37°C. The MTT formazan crystals that remained after removing the medium were then solubilized in dimethyl sulfoxide (DMSO) for 15–20 min. The absorbance was measured by a spectrophotometer at 490 nm to show the relative number of surviving cells in each well indirectly.

Colony Formation Assay

A total of 1,000 cells/well were seeded into six-well plates and incubated at 37°C and 5% CO_2 for 10–15 days. Polyformaldehyde (4%) was used to fix these cells and 1% crystal violet was used to stain these cells subsequently. The numbers of colonies containing more than 10 cells were counted.

Flow Cytometry Analysis

Cells collected were fixed with 70% alcohol and stored overnight at −20°C. Propidium iodide (PI) was used to stain the collected cells according to the instructions of the manufacturer. The stained cells were analyzed using a flow cytometer (BD Company, USA).

Immunofluorescence

KYSE150, KYSE180, and KYSE450 cells were transfected with CDCA7-V5 plasmid and empty vector, respectively. Formaldehyde (4%) was used to fix the cells for 10 min. BSA (1%) was used to incubated the cells for 1 h to block non-specific protein–protein interactions after permeabilized by 0.1% Triton X-100. The cells were incubated with the primary antibody rabbit anti-V5 (Abcam, Cambridge, UK, 2 μg/ml) overnight at 4°C. Alexa Fluor® 594 goat anti-rabbit IgG antibody (Thermo Fisher, Carlsbad, USA, 1:1,000) was used for 30 min at room temperature after washing four times in PBS. DAPI at a concentration of 0.5 μg/ml was used to stain the cell nuclei.

Chromatin Immunoprecipitation Sequencing Assay

KYSE150 cells were transfected with the V5-tagged CDCA7 plasmid for the chromatin immunoprecipitation (ChIP) assay. The assay was performed according to the instructions of the manufacturer (Millipore, Burlington, MA, USA). The DNA fragments were enriched by anti-V5 antibody (Abcam, Cambridge, UK), and the isotype IgG (Abcam, Cambridge, UK) was used as a negative control. CHIP-seq was performed by Novogene (Beijing, China). Screening and quality control of the CHIP-seq were based on standard protocol. The sequences of primers used for the amplification of *CCNA2* genome regions containing a putative CDCA7 binding site are listed in **Table S1**.

Western Blot

The cells were lysed for 1 h with RIPA buffer containing protease and phosphatase inhibitors (Thermo Fisher Scientific) on ice. The components of RIPA buffer are as follows: 1% Triton X-100, 50 mM Tris–HCl, pH 7.6, 150 mM NaCl, 1% sodium deoxycholate, and 0.1% SDS. The lysates were centrifuged at 12,000*g* at 4°C for 30 min, and the total protein concentrations of supernatant were determined by the Bradford method. Fifty micrograms of protein was separated by 10% SDS-PAGE and then transferred onto polyvinylidene fluoride (PVDF) membranes (Millipore, USA). The membrane was incubated with special antibodies, including CDCA7, CCND1, CCNA2, CCNE1, and GAPDH, at 4°C overnight. The IRDye 800CW secondary antibody (Abcam, Cambridge, UK) was used to detect the blot. A relative amount of protein was normalized to GAPDH level. The antibodies used in this experiment are shown as follows: CDCA7 (Sigma, USA), CCND1 (Proteintech, Rosemont, IL, USA), CCNA2 (Proteintech, Rosemont, IL, USA), CCNE1 (Proteintech, Rosemont, IL, USA), and GAPDH (Proteintech, Rosemont, IL, USA).

RNA Extraction and Real-Time PCR

Total RNA of ESCC cells was purified using RNAiso plus (Takara, Dalian, China). Two micrograms of total RNA was used for complementary DNA (cDNA) synthesis using a PrimeScript® RT reagent kit with gDNA Eraser (Takara). TB Green® Premix Ex Taq® II kit (Takara) was used in real-time PCR according to the instruction of the manufacturer. All real-time PCR reactions were performed in triplicate with an Applied Biosystems Step One Plus (ABI, Foster City, CA, USA). The relative expression levels of the target genes were normalized to endogenous GAPDH. Quantification of the expression levels of target genes was calculated using the $2^{-\Delta\Delta Ct}$ formula. The primers synthesized by Thermo Fisher are listed in **Table S2**.

Dual-Luciferase Reporter Assay

According to the results of ChIP-seq, we cloned the CDCA7 DNA-binding fragment with *CCNA2* from the KYSE150 cell line genomic DNA. The cloned DNA fragment was constructed into the reporter plasmid of pGL3-promoter (Promega, Madison, WI, USA). Then, we divided this DNA fragment into four segments and constructed them into the pGL3-promoter vector. The different DNA fragments were cloned using the primers listed in **Table S3**. Six motif sequences obtained from ChIP-seq were also constructed into pGL3-promoter vector. Cells (3×10^4) were cultured in triplicate in 24-well plates for 12–24 h. Then, the pGL3 reconstruction reporter plasmids were transiently co-

transfected with the pRL-TK plasmid into KYSE150 and KYSE150 knockdown cells using Lipofectamine 2000 reagent (Invitrogen, Carlsbad, CA). After transfection for 48 h, luciferase and Renilla signals were measured according to the instruction of Dual-Luciferase Reporter Assay Kit (Promega, Madison, USA).

Mouse Xenograft Assay

The effects of *CDCA7* on tumorigenesis and growth *in vivo* were detected *via* mouse xenograft assay. We used 20 5- to 6-week-old female NU-Foxn1^{nu} nude mice (Vital River Laboratory Animal Technology Co., Ltd., Beijing, China) for the mouse xenograft assay. A total of 3×10^6 KYSE150-NC cells or *CDCA7*-knockdown stable KYSE150 cells were used to inject into the right or left oxter of female NU-Foxn1^{nu} nude mice, respectively. Tumor size and weight were determined with calipers and balance twice a week. The mice were executed and the tumors were removed after 28 days. The formula $V = (W^2 \times L)/2$ was used to calculate the tumor volume. V is the tumor volume, W is the tumor width, and L is the tumor length. Tumor size was presented as mean ± standard deviation (SD).

Immunohistochemistry

The isolated xenograft tumor tissues were fixed using formalin and embedded by paraffin for immunohistochemical staining. Briefly, xylene and a series of grades of alcohol were used to deparaffinize and rehydrate these sections, and the sections were then soaked with 3% H_2O_2 15 min. Sodium citrate buffer (pH 6.0) or Tris-EDTA buffer (pH 9.0) were used for antigen retrieval for 4 or 3 min in a pressure cooker, followed by incubation with primary antibody at 4°C overnight. The slides were incubated with second antibody at 37°C for 20 min after washing with PBS and then stained with DAB and counterstained with hematoxylin. The expression of CDCA7, Ki-67, and CCNA2 was quantitatively analyzed with Aperio Cytoplasma 2.0 software by immunohistochemistry. The antibodies used in this experiment are shown as follows: CDCA7 (Sigma, USA, 1:200 dilution), Ki-67 (Abcam, Cambridge, UK), and CCNA2 (Proteintech, Rosemont, IL, USA, 1:2,000 dilution).

Bioinformatics and Data Analysis

The mRNA expression data and the clinical information of 179 ESCC tissues and paired non-tumor tissues by microarray analysis were from a previous study by Li et al. (26) and downloaded from the GEO database (GSE53625). The copy number data were obtained from our other study.

The copy number data of *CDCA7* in varied cancer types, including ESCC, ECA, lung squamous cell carcinoma (LUSC), head and neck squamous cell carcinoma (HNSC) collected from the TCGA database, were downloaded *via* cBioPortal for Cancer Genomics (https://www.cbioportal.org/) (27, 28). The expression data of *CDCA7* and *CCNA2* in different cancer types such as ESCC, LUSC, HNSC were downloaded from TCGA *via* Xena Browser (https://xenabrowser.net/heatmap/).

Statistical Analyses

Each of the experiment in the study was performed in triplicate, and data were presented as the mean ± SEM. Statistical Package for Social Science for Windows (SPSS, version 20.0; IBM Inc., USA) was used to analyze the experimental data. The means of two groups and more than two groups were compared using Student's *t*-test and one-way ANOVA, respectively. *P*-value of <0.05 was considered to be statistically significant. GraphPad Prism software was used to analyze the correlations between *CDCA7* and *CCNA2* using non-parametric correlation (Spearman).

RESULTS

CDCA7 Was Frequently Amplified in ESCC

In our previous study, *CDCA7* was identified as one of the copy number amplification genes in ESCC (8). Here, we analyzed the copy number amplification data from TCGA through cBioPortal and found that the copy number amplification of *CDCA7* existed in various kinds of tumor. Its alteration frequency was much higher in ESCC than in other tumors (**Figure 1A**). Furthermore, we analyzed the mRNA expression data of 179 pairs of ESCC tumors and adjacent normal tissues *via* microarray analysis. The mRNA expression data and the clinical information of 179 ESCC tissues and paired non-tumor tissues by microarray analysis were from a previous study by Li et al. and downloaded the from the GEO database (GSE53625) (26). We observed that *CDCA7* showed statistically higher expression levels in most of the individuals compared with that of normal tissues (**Figure 1B**). After analyzing the copy number amplification and expression of *CDCA7* in 95 ESCC patients in the TCGA database, we found that there was a correlation between the expression of *CDCA7* and the copy number amplification, indicating that the *CDCA7* copy number amplification may cause to increase its expression (**Figure 1C**).

The cohort of 179 patients was divided into two groups according to the expression level of *CDCA7*. The top 25% of patients were defined as the patients with a higher level (named as $CDCA7_{high}$) and the remaining 75% were defined as the patients with a lower level (named as $CDCA7_{low}$) according to the expression level of *CDCA7* from high to low. Then, we analyzed the correlation between the expression of *CDCA7* and the clinical variables in ESCC. The results in **Table 1** show that the expression of *CDCA7* was related to the grade of ESCC patients ($P = 0.0083$). The patients with $CDCA7_{high}$ had a poor grade compared with the $CDCA7_{low}$ patients. Furthermore, the patients with $CDCA7_{high}$ had a worse survival than those with $CDCA7_{low}$ (log rank $P = 0.01$, **Figure 1D**) using Kaplan–Meier survival analysis. The multivariate analysis showed that TNM stage [hazard ratio (HR) = 2.662, 95% CI: 1.593–4.449, $P < 0.001$], location (lower vs. upper) (HR = 2.718, 95% CI: 1.336–5.530, $P = 0.006$), age (HR = 1.657, 95% CI: 1.074–2.558, $P = 0.022$), and *CDCA7* expression (HR = 1.999, 95% CI: 1.241–3.218, $P = 0.004$) were independent predictive factors for overall survival (**Figure 1E**). Furthermore, *CDCA7* was related with the survival status in patients in the male group ($P = 0.001$), age <60 group ($P = 0.02$), drinking group ($P < 0.001$), smoking group ($P = 0.014$), T1+T2 group ($P = 0.019$), N0+N1 group ($P = 0.026$), and TNM stage = III group ($P = 0.025$) (**Figures S1, S2**). Hence, we speculate the copy number amplification and high expression

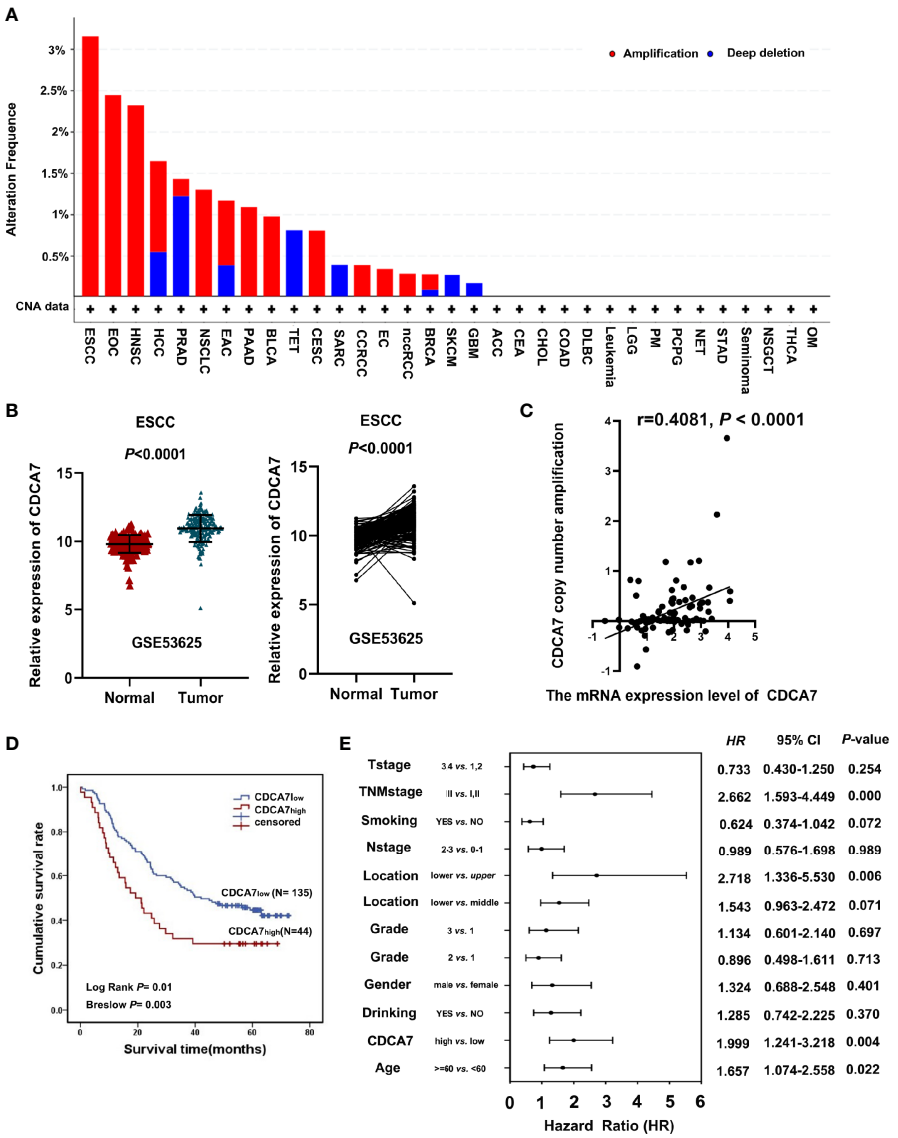


FIGURE 1 | *CDCA7* expression predicts the prognosis of esophageal squamous cell carcinoma (ESCC) patients. **(A)** The *CDCA7* copy number amplification of various kinds of tumor in the TCGA database. **(B)** ESCC tumor tissues had a high *CDCA7* expression compared with its non-tumor tissues using non-paired *t*-test and paired *t*-test; $P < 0.001$. **(C)** The correlation analysis between *CDCA7* copy number amplification and expression ($r = 0.4081$, $P < 0.0001$). **(D)** The patients with $CDCA7_{low}$ had better survival than those with $CDCA7_{high}$ (log rank $P = 0.01$) using Kaplan–Meier survival analysis. **(E)** Multivariate analysis showed that TNM stage [hazard ratio (HR) = 2.662, 95% CI: 1.593–4.449, $P < 0.001$), location (lower vs. upper) (HR = 2.718, 95% CI: 1.336–5.530, $P = 0.006$), age (HR = 1.657, 95% CI: 1.074–2.558, $P = 0.022$), and *CDCA7* expression were independent predictive factors for overall survival (HR = 1.999, 95% CI: 1.241–3.218, $P = 0.004$).

level of *CDCA7* may promote the occurrence and development of ESCC.

CDCA7 Promotes Cell Proliferation, Colony Formation, and Cell Cycle of ESCC Cells

To verify the biological roles of *CDCA7* in ESCC, we first analyzed the mRNA and protein expression levels in immortal embryonic esophageal epithelium cell lines NE3 and HET-1A and ESCC cell lines including KYSE150, KYSE180, KYSE450, and TE-1 *via* quantitative real-time PCR (q-RTPCR) and Western blot (**Figure S3**). In all these cell lines, we selected KYSE150 and KESE450 as relatively high endogenous *CDCA7* level cell lines for knockdown experiments. KYSE180 was selected as low endogenous *CDCA7* level cell line for overexpression. The efficiency of overexpression and knockdown were verified by Western blot, respectively (**Figures 2A, B**). Then, we detected the changes in cell phenotypes, including proliferation, colony formation, and cell cycle. The results showed that *CDCA7* silencing significantly inhibited the ability of cell proliferation and colony formation in

TABLE 1 | Correlation analysis between *CDCA7* copy number in ESCC and clinicopathological variables.

Clinical features	Total (n = 179)	$CDCA7_{High}$ (n = 44)	$CDCA7_{Low}$ (n = 135)	P-value
Age				
<60	91	24	67	0.571
≥60	88	20	68	
Gender				
Female	33	10	23	0.398
Male	146	34	112	
Location				
Upper	20	5	15	0.794
Middle	97	22	75	
Lower	62	17	45	
Smoking				
Never	65	15	50	0.724
Yes	114	29	85	
Drinking				
Never	73	17	56	0.739
Yes	106	27	79	
Grade				
Well	32	6	26	**0.008**
Moderately	98	18	80	
Poorly	49	20	29	
T stage				
1 + 2	39	12	27	0.310
3 + 4	140	32	108	
LN stage				
N0–N1	145	34	111	0.467
N2–N3	34	10	24	
TNM stage				
1 + 2	87	23	64	0.576
3 + 4	92	21	71	

In bold: P<0.05 was considered to be statistically significant.

KYSE150 and KESE450 (**Figures 2C, E**), while overexpression of *CDCA7* increased the ability of cell proliferation and colony formation markedly (**Figures 2D, F**). Meanwhile, the results of flow cytometry indicated that *CDCA7* overexpression decreased the proportion of G1-phase cells and increased the proportion of G2 + M-phase cells (**Figure 2H**). On the contrary, *CDCA7* silencing significantly increased the proportion of G1-phase cells and decreased the proportion of G2 + M-phase cells (**Figure 2G**).

To confirm the oncogenic role of *CDCA7 in vivo*, we established a subcutaneous transplantation tumor model in female NU-Foxn1nu nude mice using stable *CDCA7*-knockdown KYSE150 and KYSE150 cells. Four weeks later, tumors were stripped after the mice were sacrificed. The tumor growth rate of the KYSE150 group was significantly faster than that of the *CDCA7*-knockdown group (**Figure 2I**). The results found that the mean tumor volume of the *CDCA7*-knockdown group and the control group was 158.74 ± 24.83 and 488.41 ± 35.84 mm^3, respectively (t-test, $P < 0.001$, **Figure 2I**). The mean tumor weight of the control group and the *CDCA7*-knockdown group was 209.61 ± 108.84 and 54.23 ± 19.39 mg, respectively (t-test, $P < 0.001$, **Figure 2I**).

Cyclins Were Identified as *CDCA7* Targets by ChIP-Sequencing

Immunofluorescence assay was performed in KYSE150, KYSE450, and KYSE180 cells to affirm whether *CDCA7* expresses in the nucleus as CDCA7 was found to be a DNA-binding protein and can serve as a transcription regulator (9, 11, 14). The results showed that CDCA7 was located in both the cytoplasm and the nucleus (**Figure S4**).

Since CDCA7 may act as a transcription regulator, chromatin immunoprecipitation sequencing (ChIP-seq) technology was applied to screen a wide range of DNA fragments interacting with CDCA7. Genome-wide mapping of CDCA7-binding profile by ChIP-seq identified 14,930 binding events ($P < 10^{-3}$), associated with 11,897 unique genes following a nearest gene annotation. As shown in **Figure 3A**, most (12,462/14,930) of the binding events occur at a distance about 2,000 bp from the transcriptional start site (TSS) of genes, which is generally considered to be the gene promoter region and activation region. These results suggested that CDCA7 may play a role as a transcription factor or transcription regulatory factor. Next, we performed a KEGG pathway enrichment analysis on the 11,897 unique genes which were associated with the DNA fragments obtained from ChIP-seq. The KEGG pathway enrichment analysis showed that target genes were enriched in the pathways including pathways in cancer, cell cycle pathway, PI3K–Akt signaling pathway, MAPK signaling pathway, Ras signaling pathways, and Hippo signaling pathways, which may contribute to ESCC cell proliferation and tumorigenesis (**Figure 3B**). Among the DNA fragments obtained from ChIP-seq, some of them were located in the promoter region of the cell cycle related genes, i.e., 3,000 bp before the transcription initiation site. These genes include *CCND1*, *CCNE1*, *CCNA2*, etcetera (**Figure 3C**). Bdg files, such as the CDCA7_V5.bdg and Control.bdg shown in **Figure 3C**, are the corresponding bedgraph format track files provided by the company, which are convenient to view the position distribution of reads on the genome under different resolution conditions. CDCA7_V5.bdg represents the DNA fragments that can bind to CDCA7 detected by the anti-V5 antibody, and Control.bdg is the DNA fragments of the input group. When we opened the bdg files in the UCSC database and compared CDCA7_V5.bdg with Control.bdg, the location of the peak is the binding site of CDCA7 with the three cyclins. As shown in **Figure 3C**, CDCA7 binds with *CCND1* at the position of −3,732 to −2,502 bp, binds with *CCNA2* at the position of −74 to 734 bp, and binds with *CCNE1* at the positon of −53 to 925 bp from each transcription start site, respectively. At the same time, when the binding of H3K4Me1 and H3K21Ac was displayed on the genome in the UCSC database, we found that apart from *CCNE1*, CDCA7, and H3K4Me1, H3K27Ac shared the same binding position in *CCND1* and *CCNA2* genomes. These findings once again suggested that CDCA7 may play a role as a transcription factor or transcription regulator to regulate the expression of cyclins.

CCNA2 May Be the Downstream Target Gene of *CDCA7*

Since the results of ChIP-seq showed that CDCA7 may act as a transcription regulator to regulate the expression of cyclins, we detected the mRNA and protein levels of the three cyclins in *CDCA7* overexpression and knockdown stable cell lines. We

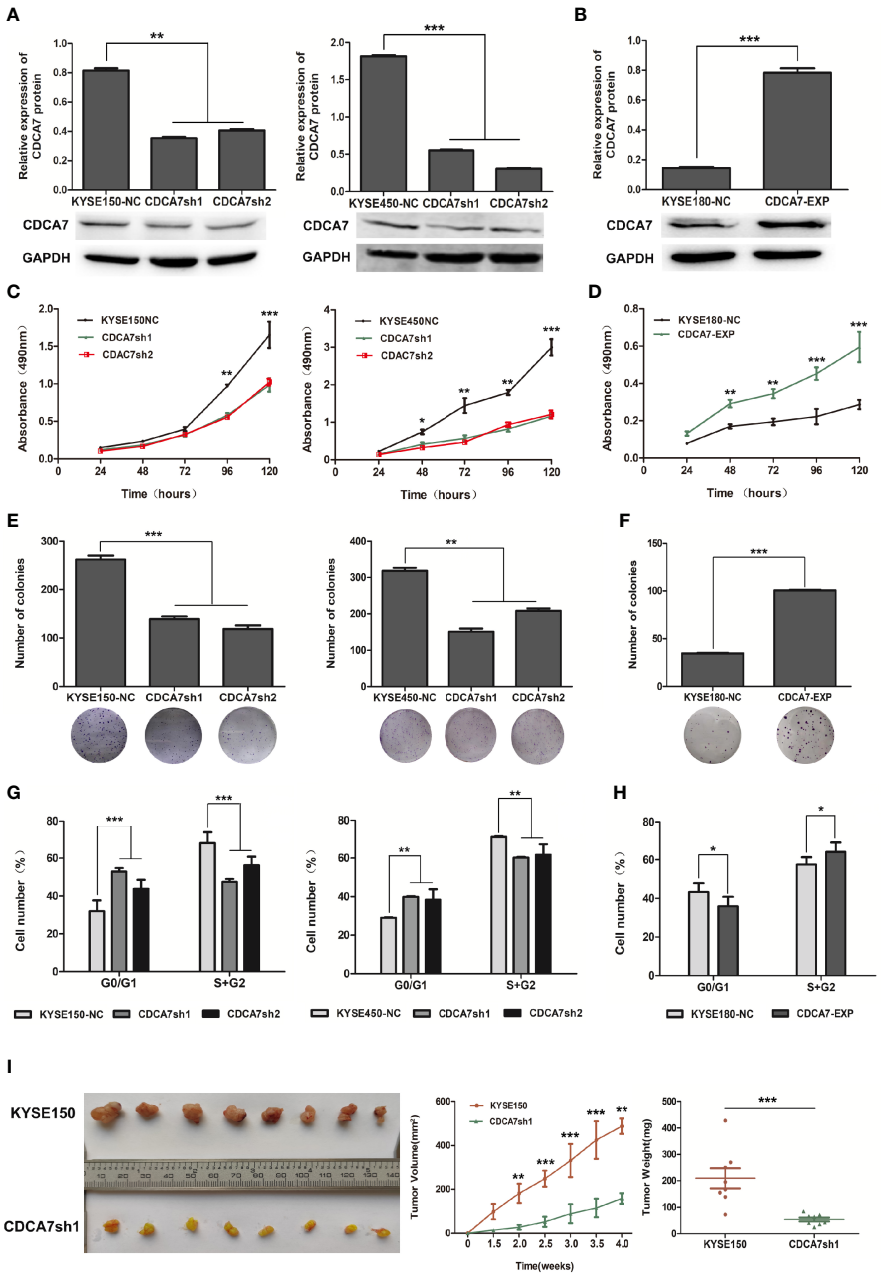


FIGURE 2 | The effect of *CDCA7* gene in ESCC cell lines. **(A)** The *CDCA7* knockdown efficiency in KYSE150 and KYSE450 cells. **(B)** The *CDCA7* overexpression efficiency in KYSE180 cells. **(C)** *CDCA7* knockdown inhibited the proliferation of ESCC Cells. **(D)** *CDCA7* overexpression promoted the proliferation of ESCC cells. **(E)** *CDCA7* knockdown inhibited the ability of colony formation in ESCC cells. **(F)** *CDCA7* overexpression promoted the ability of colony formation in ESCC cells. **(G)** *CDCA7* knockdown inhibited the cell cycle of ESCC cells. **(H)** *CDCA7* overexpression promoted the cell cycle of ESCC cells. **(I)** Tumor growth was inhibited significantly in the CDCA7-knockdown group compared with the control group in vivo. Left: tumor tissues in the CDCA7-knockdown group and the control group; right: tumor weight and tumor growth curve. ($0.01 < P \leq 0.05$, *; $0.001 < P \leq 0.01$, **; $P \leq 0.001$,***).

found that the expression level of the *CCNA2* was significantly decreased when *CDCA7* was knocked down and vice versa (**Figures 4A, B**). Meanwhile, we verified the correlation between *CDCA7* and the three cyclins using the mRNA data of 96 ESCC tissues in the TCGA database. The results (**Figure 4C**) showed that there was a weak positive correlation between *CDCA7* and *CCND1* ($r = 0.2121$, $P = 0.038$) and a strong positive correlation between *CDCA7* and *CCNA2* ($r = 0.6527$, $P < 0.0001$), while there was no correlation between *CDCA7* and *CCNE1* ($r = -0.0528$, $P = 0.6116$). Next, we analyzed the mRNA expression data of ESCC ($n = 358$) in GSE53625. Based on the mRNA expression data of ESCC ($n = 358$), *CDCA7* was positively

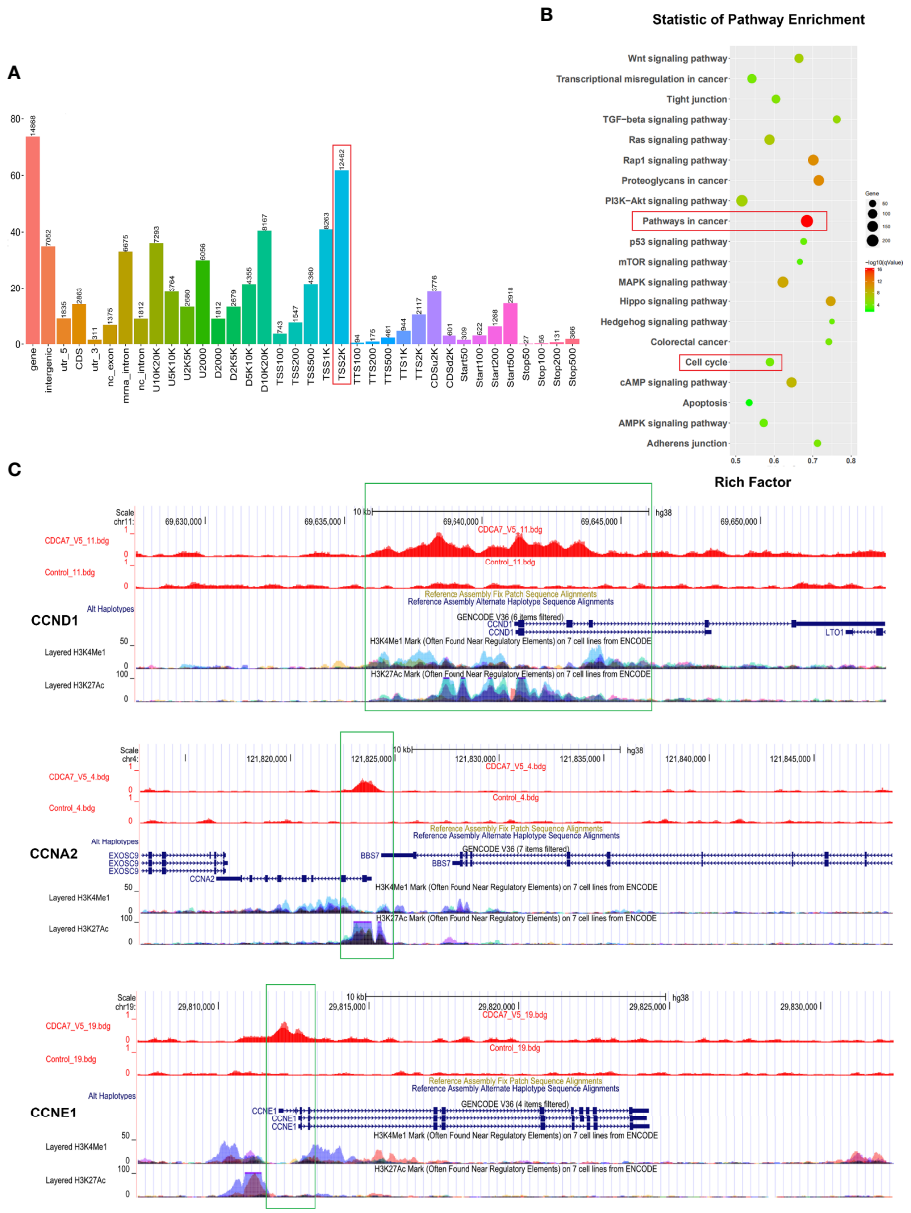


FIGURE 3 | Cyclins may be the downstream genes regulated by CDCA7 as a transcription regulator. **(A)** Most of the binding sites of CDCA7 with DNA located at the position of 2,000 bp before the transcription initiation site of the genes. **(B)** The pathway enrichment analysis of the DNA fragments obtained from ChIP-sequencing. **(C)** CDCA7 binds with *CCND1* at the position of –3,732 to –2,502 bp, binds with *CCNA2* at the position of –74 to 734 bp, and binds with *CCNE1* at the positon of –53 to 925 bp from each transcription start site, respectively.

correlated with *CCNA2* (r = 0.7047, P < 0.0001). Interestingly, when the mRNA expression data of other types of SCC in the TCGA database were analyzed, we found that *CDCA7* was positively correlated with *CCNA2* in LUSC (n = 501, r = 0.4995, P < 0.0001) and HNSC (n = 502, r = 0.5771, P < 0.0001) (**Figure 4D**). Immunohistochemistry was further used to detect the tumor tissue stripped from nude mice with anti-CDCA7, anti-CCNA2, and anti-Ki-67 antibodies. The results showed that the staining intensity of CCNA2 and Ki-67 in the *CDCA7* knockdown group was obviously weaker than that of the KYSE150 group (**Figure 4E**). The *H*-score of CCNA2 and Ki-67 in *CDCA7*-knockdown group (181.344 ± 17.549 and 7.84 ± 0.200) was significantly lower than that in the control group (91.19 ± 9.07 and 58.59 ± 0.626) (t-test, P < 0.001, **Figure 4E**).

These results revealed that the *CDCA7* probably affected the cell cycle progression, occurrence, and development of cancers through regulating the expression of *CCNA2*.

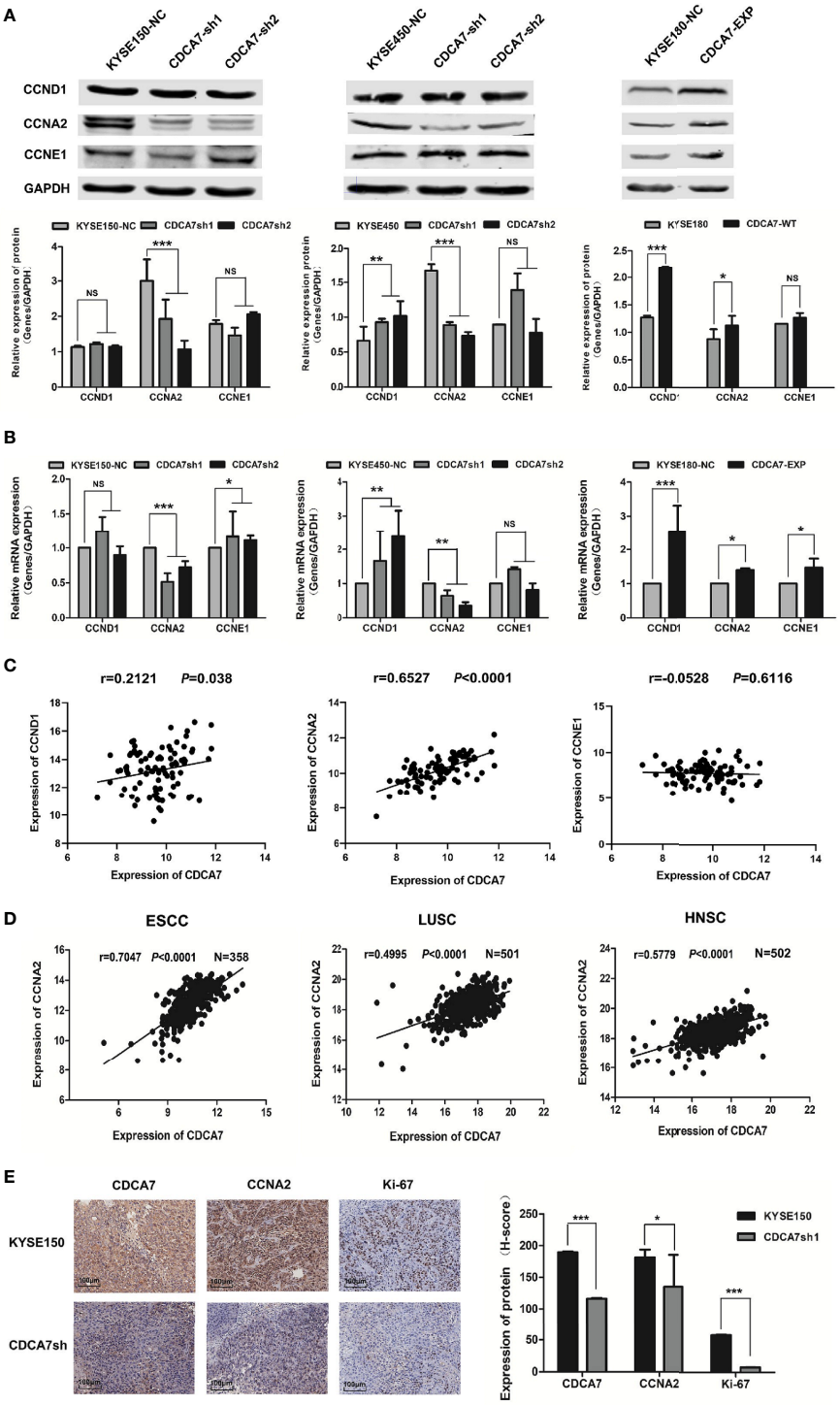


FIGURE 4 | *CCNA2* may be the downstream target gene of CDCA7. **(A)** The protein expression levels of *CCND1*, *CCNA2*, and *CCNE1* in *CDCA7* knockdown cell lines and *CDCA7* overexpression cell lines. **(B)** The mRNA expression levels of *CCND1*, *CCNA2*, and *CCNE1* in *CDCA7* knockdown cell lines and *CDCA7* overexpression cell lines. **(C)** *CDCA7* was positively correlated with *CCND1* expression ($r = 0.2121$, $P = 0.038$) and *CCNA2* expression ($r = 0.6527$, $P < 0.0001$), while it was not correlated with *CCNE1* expression ($r = -0.0528$, $P = 0.6116$). **(D)** The correlation of *CDCA7* and *CCNA2* expression in ESCC, LUSC, and HNSC; correlation coefficient (r) and *P*-values were shown in the figures. $P < 0.05$ was considered statistically significant. **(E)** IHC assay showed *CDCA7*, *CCNA2*, and *Ki-67* expression in *CDCA7* knockdown xenograft tumor tissue and the control group tissue. Scale bar = 100 μm ($P > 0.05$, NS; $0.01 < P \leq 0.05$, *; $0.001 < P \leq 0.01$,**; $P \leq 0.001$,***).

CDCA7 Regulates *CCNA2* Expression Through Binding to the Target Regions of *CCNA2*

The binding region of CDCA7 on *CCNA2* started from position −74 to 734 bp relative to the TSS, and we constructed the −90 to 809 bp into the pGL3 promoter vector for the dual-luciferase assay subsequently. The reason why we expanded the region of the DNA fragment is the high GC content of the DNA sequence near 734 bp and there is no way to design a suitable pair of PCR primers. To explore the binding domain of CDCA7, the interval from −83 to 809 bp was divided into four segments randomly. The four segments were −90 to 130, 113–292, 275–476, and 456–809 bp. Dual-luciferase assay indicated that the 456–809-bp region of the *CCNA2* was the core element regulated by CDCA7 (**Figure 5A**). To understand the molecular mechanism for the activity of CDCA7 in regulating gene transcription, a *de novo* search for DNA-enriched motifs was performed within the binding fragments and five predicted motifs were obtained. Each of the motifs corresponded to some transcription factor at different degrees (**Figure 5B**). To demonstrate whether CDCA7 acts its role as a transcription factor or a transcription regulator factor by binding with DNA through these motifs, we constructed the sequences of the motifs into pGL3-promoter vectors for dual-luciferase reporting experiments, the results revealed that motif-1 (5′-TAGACAAGAGTT-3′), motif-2 (5′-GTGATCAGTGCAGA-3′), motif-3 (5′-CTGGAACAGCAC-3′), motif-4 (5′-GTGTGTGTGTGT-3′), and motif-5 (5′-AGTAGTAGTA-3′) might be the functional binding sites of CDCA7 with DNA (**Figure 5D**). Next, we compared these motifs with the sequences from −74 to 734 bp and found five DNA

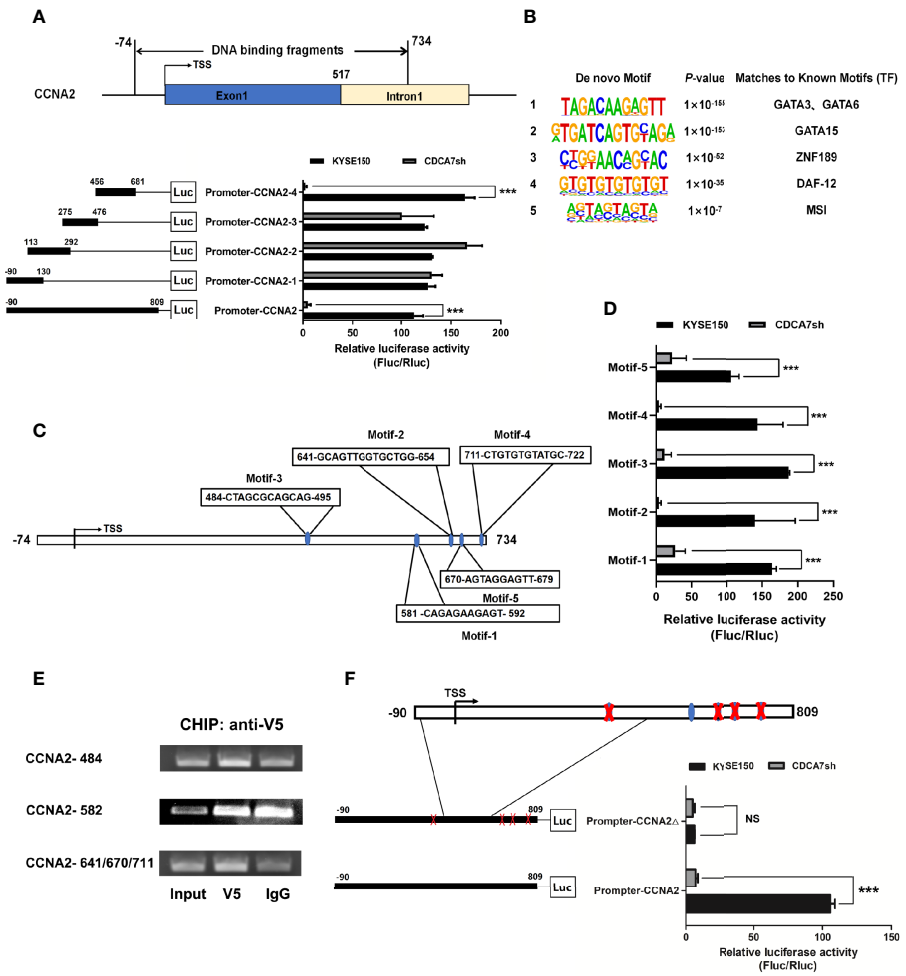


FIGURE 5 | CDCA7 directly binds to *CCNA2* and increases its transcription activity in ESCC cells. **(A)** CDCA7 binds with *CCNA2* from −74 to 734 bp in the ChIP-seq and luciferase reporter assays showed that 456 to 809 bp of *CCNA2* were the core regions regulated by *CDCA7*. **(B)** Five predicted motifs were analyzed from the DNA-binding fragments obtained from the ChIP-seq. **(C)** Five DNA fragments in the 456–809-bp region were highly similar to motif-1, motif-2, motif-3, motif-4, and motif-5. **(D)** Luciferase reporter assays showed that CDCA7 may regulate the transcription of the target genes through the five motifs. **(E)** ChIP-PCR showed that CDCA7 binds to the *CCNA2* at the position of 484–495, 641–654, 670–679, and 711–722 bp in ESCC cells. **(F)** Luciferase reporter assays showed that CDCA7 could not activate the expression of downstream reporter gene when the four binding sites of 484–495, 641–654, 670–679, and 711–722 bp were knocked out. (P > 0.05, NS; P ≤ 0.001,***).

fragments highly similar to the five motifs located in the 456–809-bp region (**Figure 5C**). We inferred that the five DNA segments may be the functional binding sites of CDCA7 with *CCNA2*. In order to further research whether CDCA7 directly binds to *CCNA2* through these sites, ChIP-PCR was performed. We found that CDCA7 bound to the DNA fragment from 484 to 495, 641 to 654, 670 to 679, and 711 to 722 bp in the *CCNA2* genomic region (**Figure 5E**). To identify whether CDCA7 regulates *CCNA2* through these binding sites, we constructed a dual-luciferase reporter plasmid with the segments of *CCNA2* from −90 and 809 bp which deleted the 484–495, 641–654, 670–679, and 711–722 bp. The results showed that CDCA7 could not activate the expression of downstream reporter gene when the four binding sites were knocked out (**Figure 5F**). It proved again that CDCA7 may regulate the transcription and expression of *CCNA2* by binding with these four binding sites.

CDCA7 Promotes Cell Cycle Through Regulating *CCNA2*

To confirm whether *CDCA7* promotes cell cycle through *CCNA2*, we carried out the interference and rescue experiment of *CCNA2*. The results showed that forced overexpression of *CCNA2* in *CDCA7* knockdown ESCC cells (**Figure 6A**) was performed, and a series of phenotype changes had been identified in ESCC cells. The results showed that *CCNA2* overexpression promoted the proliferation and colony formation (**Figure 6B**) induced by *CDCA7* knockdown in KYSE450 cells. Meanwhile, when we silenced its expression in *CDCA7* overexpression ESCC cells (**Figure 6C**), the results of cell phenotype experiments showed that knockdown of *CCNA2* can reduce cell proliferation and colony formation ability (**Figure 6D**) induced by *CDCA7* overexpression in KYSE180 cells. These results suggested that the acceleration effect of *CDCA7* on cell cycle may depend on its transcription regulation of *CCNA2*, and *CCNA2* inhibition may partially reverse the cell proliferation progression induced by *CDCA7* overexpression.

The results indicated that *CDCA7* gene might act as a tumor promoter in ESCC and its copy number amplification or increased expression may accelerate the cell cycle process and promote cell proliferation by binding to the genome of *CCNA2* functional domain and increasing its expression in ESCC. When *CDCA7* is knocked down or decreased, its transcription regulation effect is attenuated, and the cell cycle process and the cell proliferation of ESCC are inhibited as the expression of *CCNA2* is depressed. Furthermore, the mechanism that *CDCA7* acts as an oncogene possibly through regulation of cell proliferation might be applied in various types of SCC.

DISCUSSION

Previous reports showed that overexpression of *CDCA7* predicts poor prognosis and tumor progression in human breast cancer, lung adenocarcinoma and lymphoma, colorectal cancer, and pancreatic diseases (29–33). In this study, we uncovered the potential prognostic value of *CDCA7*, one of the copy number altered genes, for ESCC patients; revealed the tumor-promoting role of *CDCA7* gene; and explored its possible mechanism in ESCC for the first time. *CDCA7* was highly expressed in not only ESCC but also SCC in transcriptome sequencing data. The Kaplan–Meier survival analysis showed that patients with high expression level of *CDCA7* had poor prognosis. This result reminded us that *CDCA7* may be used as a candidate target to guide the individual diagnosis and a biomarker to establish a technical system for the molecular classification of ESCC.

Further functional studies reveal that CDCA7 may exert its oncogenic roles *via* directly binding to the position of 484–495, 641–654, 670–679, and 711–722 bp from the transcription start site of *CCNA2*. The data of TCGA and GSE53625 further confirmed the positive correlation between *CDCA7* and *CCNA2* in ESCC, indicating that the high expression level of *CDCA7* may be an important driving event in the occurrence and development in ESCC.

CCNA2, which is one of the two A-type cyclins and ubiquitously expressed in cultured cells, has been reported to be upregulated in a variety of cancers (34–37). CCNA2 is considered to be the critical S-phase cyclin in mammalian cells (18, 38). *CCNA2* is expressed at the beginning of the S-phase (16, 39) and existed in both the S- and G2-phases. Once synthesized, it binds with its catalytic partners, the cyclin-dependent kinases (CDK) CDK2 and CDK1, and activates its catalytic activity. The CDK2/CCNA2 complex promotes DNA replication through localizing to replication foci in the nucleus (17, 40). The complexes phosphorylate the proteins which play important roles in DNA synthesis and thus drive the S-phase progression (16, 18, 19, 41–43). In addition, a second function of CCNA2 is involved in the entry of cells into mitosis since it also is expressed at the G2-phase (44). The accumulation of CCNA2 is rate-limiting for S-phase entry, so overexpression of *CCNA2* can induce cultured cell early entry into the S-phase under normal circumstances (45, 46). Indeed, inhibition of CCNA2 function by p21Cip1 during the G2-phase or injection of anti-CCNA2 antibodies into cultured fibroblasts both can block the process of cells into mitosis (41, 47).

It is known to all that disorder of the cell cycle process is one of the causes of many cancers (48–52). Cancer cells lose many of the inhibitory controls in the cell cycle because of the inactivation or mutation of suppressor genes and overexpression or amplification of oncogenes (53). The aberrant transcription of upregulation of cyclins and CDKs can result in uncontrolled cell cycle progression and mitosis. Our study showed that *CCNA2* was a direct downstream target gene of CDCA7, and its expression may be activated by CDCA7 on both the transcription level and the protein level. Therefore, the copy number amplification or increase of *CDCA7* may lead to a high level expression of *CCNA2* to accelerate the cell cycle process. This may be a mechanism and indicate the important role of *CDCA7* in ESCC. Therefore, we speculated that patients with high expression of *CDCA7* could be treated with cell cycle-specific agents (CCSA) since the expression of *CCNA2* and the number of cells in the proliferative phase are correspondingly increased. This study provides a theoretical and experimental

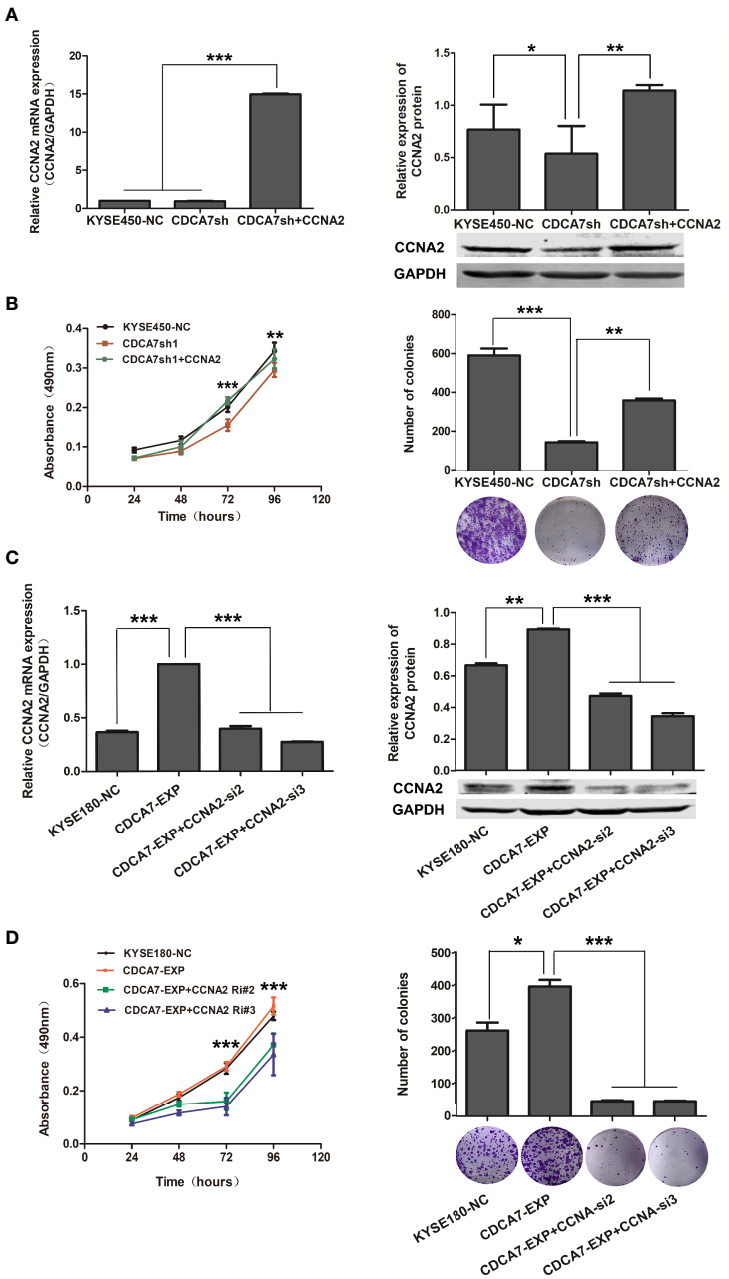


FIGURE 6 | *CDCA7* promotes S-phases *via* transcriptionally regulating *CCNA2* expression. **(A)** *CCNA2* overexpression in KYSE450 cell line with *CDCA7* knockdown; the mRNA and protein expression levels of *CCNA2* and *GAPDH* were detected by qRT-PCR and Western blot. *GAPDH* was used as a loading control. **(B)** *CCNA2* overexpression promoted the proliferation induced by *CDCA7* knockdown. **(C)** *CCNA2* knockdown in KYSE180 cell line with *CDCA7* overexpression; the mRNA and protein expression levels of *CCNA2* and *GAPDH* were detected by Western blot. *GAPDH* was used as a loading control. **(D)** *CCNA2* knockdown inhibited the proliferation induced by *CDCA7* overexpression. ($0.01 < P \leq 0.05$, *;$0.001 < P \leq 0.01$,**,$P \leq 0.001$,***).

foundation for the research and development of drug targets for clinical treatment of ESCC in China.

In summary, our study shows that *CDCA7*, a copy number amplification gene in ESCC, may act as a tumor promoter *via* regulating *CCNA2* directly and accelerate the cell cycle process of ESCC cells. The copy number amplification may lead to tumorigenesis and progression of ESCC. Our findings provide a new insight into the molecular mechanisms involved in ESCC development. However, there are still some deficiencies in our research process. Whether the high expression level of *CDCA7* is more sensitive to CCSA as we expected needs further experimental verification. Meanwhile, further in-depth research

is needed to clarify the mechanism of ECSS carcinogenesis, to develop the prognostic method, and to identify feasible therapeutic targets which could be used to overcome the disease.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found here: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53625, https://xenabrowser.net/, and https://www.cbioportal.org/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Reviewing Board (IRB) and the Research Committee of Shanxi Medical University. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by Institutional Reviewing Board (IRB) and the Laboratory Animal Welfare Committee of Shanxi Medical University.

AUTHOR CONTRIBUTIONS

YC and XC designed the study. HL, YJW, SJW, FW, CC, EX, SQW, DG, FC, YB, and YM acquired the data. HL, YJW, SJW, and HC analyzed the data. HL prepared the manuscript. PK, YQW, and LZ edited the manuscript. YC reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The work was supported by funds from the National Natural Science Foundation of China (81602175, 81773150, 81972613), the Fund for Shanxi "1331 Project" and "1331 Project" Key Subjects Construction, the Fund for "Sanjin Scholars," the Foundation for Youths of Shanxi Province (201901D211349, 201901D211345), and the Natural Science Foundation of Shanxi Province (201801D121306).

ACKNOWLEDGMENTS

This work uses data that have been provided from a previous study by Li et al. and can be downloaded from the GEO database (GSE53625).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.734655/full#supplementary-material

REFERENCES

1. Chen W, Zheng R, Baade P, Zhang S, Zeng H, Bray F, et al. Cancer Statistics in China, 2015. *CA: Cancer J For Clin* (2016) 66(2):115–32. doi: 10.3322/caac.21338
2. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global Surveillance of Trends in Cancer Survival 2000-14 (CONCORD-3): Analysis of Individual Records for 37 513 025 Patients Diagnosed With One of 18 Cancers From 322 Population-Based Registries in 71 Countries. *Lancet (London England)* (2018) 391(10125):1023–75. doi: 10.1016/S0140-6736(17)33326-3
3. Zhang L, Zhou Y, Cheng C, Cui H, Cheng L, Kong P, et al. Genomic Analyses Reveal Mutational Signatures and Frequently Altered Genes in Esophageal Squamous Cell Carcinoma. *Am J Hum Genet* (2015) 96(4):597–611. doi: 10.1016/j.ajhg.2015.02.017
4. Song Y, Li L, Ou Y, Gao Z, Li E, Li X, et al. Identification of Genomic Alterations in Oesophageal Squamous Cell Cancer. *Nature* (2014) 509 (7498):91–5. doi: 10.1038/nature13176
5. Lin D, Hao J, Nagata Y, Xu L, Shang L, Meng X, et al. Genomic and Molecular Characterization of Esophageal Squamous Cell Carcinoma. *Nat Genet* (2014) 46(5):467–73. doi: 10.1038/ng.2935
6. Gao Y, Chen Z, Li J, Hu X, Shi X, Sun Z, et al. Genetic Landscape of Esophageal Squamous Cell Carcinoma. *Nat Genet* (2014) 46(10):1097–102. doi: 10.1038/ng.3076
7. Sawada G, Niida A, Uchi R, Hirata H, Shimamura T, Suzuki Y, et al. Genomic Landscape of Esophageal Squamous Cell Carcinoma in a Japanese Population. *Gastroenterology* (2016) 150(5):1171–82. doi: 10.1053/j.gastro.2016.01.035
8. Cheng C, Zhou Y, Li H, Xiong T, Li S, Bi Y, et al. Whole-Genome Sequencing Reveals Diverse Models of Structural Variations in Esophageal Squamous Cell Carcinoma. *Am J Hum Genet* (2016) 98(2):256–74. doi: 10.1016/j.ajhg.2015.12.013
9. Prescott J, Osthus R, Lee L, Lewis B, Shim H, Barrett J, et al. A Novel C-Myc-Responsive Gene, JPO1, Participates in Neoplastic Transformation. *J Biol Chem* (2001) 276(51):48276–84. doi: 10.1074/jbc.M107357200
10. Haggerty T, Zeller K, Osthus R, Wonsey D, Dang C. A Strategy for Identifying Transcription Factor Binding Sites Reveals Two Classes of Genomic C-Myc Target Sites. *Proc Natl Acad Sci USA* (2003) 100(9):5313–8. doi: 10.1073/pnas.0931346100
11. Goto Y, Hayashi R, Muramatsu T, Ogawa H, Eguchi I, Oshida Y, et al. JPO1/CDCA7, a Novel Transcription Factor E2F1-Induced Protein, Possesses Intrinsic Transcriptional Regulator Activity. *Biochim Biophys Acta* (2006) 1759:60–8. doi: 10.1016/j.bbaexp.2006.02.004
12. Gill R, Gabor T, Couzens A, Scheid M. The MYC-Associated Protein CDCA7 Is Phosphorylated by AKT to Regulate MYC-Dependent Apoptosis and Transformation. *Mol Cell Biol* (2013) 33(3):498–513. doi: 10.1128/MCB.00276-12
13. Osthus R, Karim B, Prescott J, Smith B, McDevitt M, Huso D, et al. The Myc Target Gene JPO1/CDCA7 Is Frequently Overexpressed in Human Tumors and has Limited Transforming Activity *In Vivo*. *Cancer Res* (2005) 65 (13):5620–7. doi: 10.1158/0008-5472.CAN-05-0536
14. Ye L, Li F, Song Y, Yu D, Xiong Z, Li Y, et al. Overexpression of CDCA7 Predicts Poor Prognosis and Induces EZH2-Mediated Progression of Triple-Negative Breast Cancer. *Int J Of Cancer* (2018) 143(10):2602–13. doi: 10.1002/ijc.31766
15. Wang H, Ye L, Xing Z, Li H, Lv T, Liu H, et al. CDCA7 Promotes Lung Adenocarcinoma Proliferation via Regulating the Cell Cycle. *Pathol Res Pract* (2019) 215(11):152559. doi: 10.1016/j.prp.2019.152559
16. Girard F, Strausfeld U, Fernandez A, Lamb N. Cyclin A Is Required for the Onset of DNA Replication in Mammalian Fibroblasts. *Cell* (1991) 67(6):1169–79. doi: 10.1016/0092-8674(91)90293-8

17. Jackman M, Kubota Y, den Elzen N, Hagting A, Pines J. Cyclin A- and Cyclin E-Cdk Complexes Shuttle Between the Nucleus and the Cytoplasm. *Mol Biol Cell* (2002) 13(3):1030–45. doi: 10.1091/mbc.01-07-0361
18. Yam C, Fung T, Poon R. Cyclin A in Cell Cycle Control and Cancer. *Cell Mol Life Sci CMLS* (2002) 59(8):1317–26. doi: 10.1007/s00018-002-8510-y
19. den Elzen N, Pines J. Cyclin A Is Destroyed in Prometaphase and can Delay Chromosome Alignment and Anaphase. *J Cell Biol* (2001) 153(1):121–36. doi: 10.1083/jcb.153.1.121
20. Wang Y, Prives C. Increased and Altered DNA Binding of Human P53 by S and G2/M But Not G1 Cyclin-Dependent Kinases. *Nature* (1995) 376 (6535):88–91. doi: 10.1038/376088a0
21. Ohashi R, Gao C, Miyazaki M, Hamazaki K, Tsuji T, Inoue Y, et al. Enhanced Expression of Cyclin E and Cyclin A in Human Hepatocellular Carcinomas. *Anticancer Res* (2001) 21:657–62.
22. Volm M, Koomägi R, Mattern J, Stammler G. Cyclin A is Associated With an Unfavourable Outcome in Patients With Non-Small-Cell Lung Carcinomas. *Br J Cancer* (1997) 75(12):1774–8. doi: 10.1038/bjc.1997.302
23. Bukholm I, Bukholm G, Nesland J. Over-Expression of Cyclin A is Highly Associated With Early Relapse and Reduced Survival in Patients With Primary Breast Carcinomas. *Int J Cancer* (2001) 93(2):283–7. doi: 10.1002/ijc.1311
24. Kanai M, Shiozawa T, Xin L, Nikaido T, Fujii S. Immunohistochemical Detection of Sex Steroid Receptors, Cyclins, and Cyclin-Dependent Kinases in the Normal and Neoplastic Squamous Epithelia of the Uterine Cervix. *Cancer* (1998) 82(9):1709–19. doi: 10.1002/(SICI)1097-0142(19980501)82:9<1709::AID-CNCR18>3.0.CO;2-8
25. Yasmeen A, Berdel W, Serve H, Müller-Tidow C–. And A-Type Cyclins as Markers for Cancer Diagnosis and Prognosis. *Expert Rev Mol Diagn* (2003) 3 (5):617–33. doi: 10.1586/14737159.3.5.617
26. Li J, Chen Z, Tian L, Zhou C, He M, Gao Y, et al. LncRNA Profile Study Reveals a three-lncRNA Signature Associated With the Survival of Patients With Oesophageal Squamous Cell Carcinoma. *Gut* (2014) 63(11):1700–10. doi: 10.1136/gutjnl-2013-305806
27. Cerami E, Gao J, Dogrusoz U, Gross B, Sumer S, Aksoy B, et al. The Cbio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov* (2012) 2(5):401–4. doi: 10.1158/2159-8290.CD-12-0095
28. Gao J, Aksoy B, Dogrusoz U, Dresdner G, Gross B, Sumer S, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the Cbioportal. *Sci Signaling* (2013) 6(269):pl1. doi: 10.1126/scisignal.2004088
29. Chhatriya B, Mukherjee M, Ray S, Saha B, Lahiri S, Halder S, et al. Transcriptome Analysis Identifies Putative Multi-Gene Signature Distinguishing Benign and Malignant Pancreatic Head Mass. *J Trans Med* (2020) 18(1):420. doi: 10.1186/s12967-020-02597-1
30. Li S, Huang J, Qin M, Zhang J, Liao C. High Expression of CDCA7 Predicts Tumor Progression and Poor Prognosis in Human Colorectal Cancer. *Mol Med Rep* (2020) 22(1):57–66. doi: 10.3892/mmr.2020.11089
31. Lewis B, Shim H, Li Q, Wu C, Lee L, Maity A, et al. Identification of Putative C-Myc-Responsive Genes: Characterization of Rcl, a Novel Growth-Related Gene. *Mol Cell Biol* (1997) 17(9):4967–78. doi: 10.1128/MCB.17.9.4967
32. Martín-Cortázar C, Chiodo Y, Jiménez R, Bernabé M, Cayuela M, Iglesias T, et al. CDCA7 Finely Tunes Cytoskeleton Dynamics to Promote Lymphoma Migration and Invasion. *Haematologica* (2020) 105(3):730–40. doi: 10.3324/haematol.2018.215459
33. Jiménez-P R, Martín-Cortázar C, Kourani O, Chiodo Y, Cordoba R, Domínguez-Franjo M, et al. CDCA7 Is a Critical Mediator of Lymphomagenesis That Selectively Regulates Anchorage-Independent Growth. *Haematologica* (2018) 103(10):1669–78. doi: 10.3324/haematol.2018.188961
34. Wang J, Chenivesse X, Henglein B, Bréchot C. Hepatitis B Virus Integration in a Cyclin A Gene in a Hepatocellular Carcinoma. *Nature* (1990) 343 (6258):555–7. doi: 10.1038/343555a0
35. Pines J, Hunter T. Human Cyclin A is Adenovirus E1A-Associated Protein P60 and Behaves Differently From Cyclin B. *Nature* (1990) 346(6286):760–3. doi: 10.1038/346760a0
36. Ravnik S, Wolgemuth D. The Developmentally Restricted Pattern of Expression in the Male Germ Line of a Murine Cyclin A, Cyclin A2, Suggests Roles in Both Mitotic and Meiotic Cell Cycles. *Dev Biol* (1996) 173(1):69–78. doi: 10.1006/dbio.1996.0007
37. Wolgemuth D. Function of the A-Type Cyclins During Gametogenesis and Early Embryogenesis. *Results Problems Cell Differ* (2011) 53:391–413. doi: 10.1007/978-3-642-19065-0_17
38. Hochegger H, Takeda S, Hunt T. Cyclin-Dependent Kinases and Cell-Cycle Transitions: Does One Fit All? *Nat Rev Mol Cell Biol* (2008) 9(11):910–6. doi: 10.1038/nrm2510
39. Erlandsson F, Linnman C, Ekholm S, Bengtsson E, Zetterberg A. A Detailed Analysis of Cyclin A Accumulation at the G(1)/S Border in Normal and Transformed Cells. *Exp Cell Res* (2000) 259(1):86–95. doi: 10.1006/excr.2000.4889
40. Cardoso M, Leonhardt H, Nadal-Ginard B. Reversal of Terminal Differentiation and Control of DNA Replication: Cyclin A and Cdk2 Specifically Localize at Subnuclear Sites of DNA Replication. *Cell* (1993) 74 (6):979–92. doi: 10.1016/0092-8674(93)90721-2
41. Pagano M, Pepperkok R, Verde F, Ansorge W, Draetta G. Cyclin A Is Required at Two Points in the Human Cell Cycle. *EMBO J* (1992) 11 (3):961–71. doi: 10.1002/j.1460-2075.1992.tb05135.x
42. Zindy F, Lamas E, Chenivesse X, Sobczak J, Wang J, Fesquet D, et al. Cyclin A Is Required in S Phase in Normal Epithelial Cells. *Biochem Biophys Res Commun* (1992) 182(3):1144–54. doi: 10.1016/0006-291X(92)91851-G
43. Fisher D. Control of DNA Replication by Cyclin-Dependent Kinases in Development. *Results Problems Cell Differ* (2011) 53:201–17. doi: 10.1007/978-3-642-19065-0_10
44. Swenson K, Farrell K, Ruderman J. The Clam Embryo Protein Cyclin A Induces Entry Into M Phase and the Resumption of Meiosis in Xenopus Oocytes. *Cell* (1986) 47(6):861–70. doi: 10.1016/0092-8674(86)90801-9
45. Resnitzky D, Hengst L, Reed S. Cyclin A-Associated Kinase Activity Is Rate Limiting for Entrance Into S Phase and Is Negatively Regulated in G1 by p27Kip1. *Mol Cell Biol* (1995) 15(8):4347–52. doi: 10.1128/MCB.15.8.4347
46. Chibazakura T, Kamachi K, Ohara M, Tane S, Yoshikawa H, Roberts J. Cyclin A Promotes S-Phase Entry via Interaction With the Replication Licensing Factor Mcm7. *Mol Cell Biol* (2011) 31(2):248–55. doi: 10.1128/MCB.00630-10
47. Furuno N, den Elzen N, Pines J. Human Cyclin A is Required for Mitosis Until Mid Prophase. *J Cell Biol* (1999) 147(2):295–306. doi: 10.1083/jcb.147.2.295
48. Malumbres M. Cyclin-Dependent Kinases. *Genome Biol* (2014) 15(6):122. doi: 10.1186/gb4184
49. Kaplon J, van Dam L, Peeper D. Two-Way Communication Between the Metabolic and Cell Cycle Machineries: The Molecular Basis. *Cell Cycle (Georgetown Tex)* (2015) 14(13):2022–32. doi: 10.1080/15384101.2015.1044172
50. Moncada S, Higgs E, Colombo S. Fulfilling the Metabolic Requirements for Cell Proliferation. *Biochem J* (2012) 446(1):1–7. doi: 10.1042/BJ20120427
51. Kalucka J, Missiaen R, Georgiadou M, Schoors S, Lange C, De Bock K, et al. Metabolic Control of the Cell Cycle. *Cell Cycle (Georgetown Tex)* (2015) 14 (21):3379–88. doi: 10.1080/15384101.2015.1090068
52. Aarts M, Linardopoulos S, Turner N. Tumour Selective Targeting of Cell Cycle Kinases for Cancer Treatment. *Curr Opin Pharmacol* (2013) 13(4):529–35. doi: 10.1016/j.coph.2013.03.012
53. Icard P, Fournel L, Wu Z, Alifano M, Lincet H. Interconnection Between Metabolism and Cell Cycle in Cancer. *Trends Biochem Sci* (2019) 44(6):490–501. doi: 10.1016/j.tibs.2018.12.007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 29 November 2021
doi: 10.3389/fonc.2021.741626

Precise Identification of Recurrent Somatic Mutations in Oral Cancer Through Whole-Exome Sequencing Using Multiple Mutation Calling Pipelines

*Li-Han Lin[1], Chung-Hsien Chou[2], Hui-Wen Cheng[1], Kuo-Wei Chang[2,3] and Chung-Ji Liu[1,4]**

[1] Department of Medical Research, MacKay Memorial Hospital, Taipei, Taiwan, [2] Institute of Oral Biology, School of Dentistry, National Yang Ming Chiao Tung University, Taipei, Taiwan, [3] Department of Stomatology, Taipei Veterans General Hospital, Taipei, Taiwan, [4] Department of Oral and Maxillofacial Surgery, Taipei MacKay Memorial Hospital, Taipei, Taiwan

OPEN ACCESS

Edited by:
Die Wang, Hudson Institute of Medical Research, Australia

Reviewed by:
Beifang Niu, Chinese Academy of Sciences (CAS), China
Hong Zheng, Stanford University, United States
Yuchao Zhang, Shanghai Institutes for Biological Sciences (CAS), China

***Correspondence:**
Chung-Ji Liu
cjliu3229@gmail.com

Specialty section:
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology

Received: *15 July 2021*
Accepted: *11 November 2021*
Published: *29 November 2021*

Citation:
Lin L-H, Chou C-H, Cheng H-W, Chang K-W and Liu C-J (2021) Precise Identification of Recurrent Somatic Mutations in Oral Cancer Through Whole-Exome Sequencing Using Multiple Mutation Calling Pipelines. Front. Oncol. 11:741626. doi: 10.3389/fonc.2021.741626

Understanding the genomic alterations in oral carcinogenesis remains crucial for the appropriate diagnosis and treatment of oral squamous cell carcinoma (OSCC). To unveil the mutational spectrum, in this study, we conducted whole-exome sequencing (WES), using six mutation calling pipelines and multiple filtering criteria applied to 50 paired OSCC samples. The tumor mutation burden extracted from the data set of somatic variations was significantly associated with age, tumor staging, and survival. Several genes (*MUC16*, *MUC19*, *KMT2D*, *TTN*, *HERC2*) with a high frequency of false positive mutations were identified. Moreover, known (*TP53*, *FAT1*, *EPHA2*, *NOTCH1*, *CASP8*, and *PIK3CA*) and novel (*HYDIN*, *ALPK3*, *ASXL1*, *USP9X*, *SKOR2*, *CPLANE1*, *STARD9*, and *NSD2*) genes have been found to be significantly and frequently mutated in OSCC. Further analysis of gene alteration status with clinical parameters revealed that canonical pathways, including clathrin-mediated endocytotic signaling, NFκB signaling, PEDF signaling, and calcium signaling were associated with OSCC prognosis. Defining a catalog of targetable genomic alterations showed that 58% of the tumors carried at least one aberrant event that may potentially be targeted by approved therapeutic agents. We found molecular OSCC subgroups which were correlated with etiology and prognosis while defining the landscape of major altered events in the coding regions of OSCC genomes. These findings provide information that will be helpful in the design of clinical trials on targeted therapies and in the stratification of patients with OSCC according to therapeutic efficacy.

Keywords: mutation burden, oral cancer, somatic mutation, survival, whole-exome sequencing

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is one of the most common malignancies of the upper aerodigestive tract, with poor prognosis and high mortality rates. In 2020 alone, 377,713 new cases of OSCC were diagnosed worldwide, among whom 177,757 have died from their disease (1). OSCC generally develops as a result of multi-step carcinogenic processes (2). Meanwhile, approximately 4%–7.4% of the patients have been found to develop simultaneous tumors which are located in the head and neck region (3, 4). Moreover, multiple lesions may develop concurrently and over large mucosal areas, subsequently progressing into cancers. This may be the reason for the high recurrence of OSCC after treatment (5), as well as the increased incidence and mortality of OSCC worldwide (6). Therefore, understanding the genomic alterations which are associated with OSCC carcinogenesis is crucial for appropriate diagnosis and therapy.

Recent developments in high-throughput next-generation parallel sequencing technologies have facilitated the sensitive detection and quantification of genetic alterations in tumor biopsies. In line with this, whole-exome sequencing (WES) has provided new insights into the molecular basis of head and neck squamous cell carcinoma (HNSCC) progression (7, 8). WES data obtained from the Cancer Genome Atlas (TCGA) (9, 10) has further highlighted this molecular complexity by identifying novel significantly mutated genes (11). Therefore, to improve the diagnosis of individuals at risk and the treatment of patients, more sensitive and specific biomarkers for OSCC need to be established (10, 12). Previous exome sequencing studies on HNSCC have consistently revealed that *TP53*, *CDKN2A*, *PIK3CA*, *HRAS*, and *NOTCH1* were significantly mutated (7, 8). Another genomic analysis of OSCC in Taiwan revealed that the *CHUK* and *ELAVL1* genes were significantly and frequently mutated (13). Moreover, frequently and recurrently mutated genes, including *USP9X*, *MLL4*, *ARID2*, *UNC13C*, and *TRPM3*, have been reported in gingivo-buccal oral squamous cell carcinomas (14). The accumulation of somatic mutations within a cancer genome has revealed that certain oncogenic patterns are associated with the exposure to mutagens and with defects in DNA repair (15, 16).

Several packages for the analysis of genomic data with different algorithms have been applied to increase the accuracy of mutation detection. Nonetheless, significant discrepancies between the results of different algorithms have been observed, leading to difficulties in selecting candidate mutations for validation (17, 18). To date, no single study has been able to exhaustively address all possible relevant issues in variant calling. Hence, the current study sought to contribute towards addressing this concern by comparing six variant callers using data from 50 matched-paired OSCC samples which were sequenced on a whole-exome platform. Circularity in defining false-negative mutations was minimized by relying on previous high-quality work using independent data and variant calling methods. Selected data were then combined and validated using Sanger sequencing and Integrative Genomics Viewer (IGV) to greatly reduce false-positive calls while maintaining sensitivity for detecting genuine mutations. Thereafter, truly somatic mutations were analyzed using clinical data.

METHODS

Participants and Data Collection

50 patients with OSCC were enrolled in this study after providing informed consent. This study was approved by the institutional review board of MacKay Memorial Hospital (approval numbers: 12MMHIS178 and 15MMHIS104). Tumor specimens were collected from patients during OSCC surgery. Laser capture microdissection was performed to isolate relatively pure tumor cells for DNA extraction according to previously established protocols (19). 10 mL of whole blood were collected in Vacutainer tubes containing ethylenediaminetetraacetic acid as anticoagulant (Becton Dickinson, Franklin Lakes, NJ). Genomic DNA was extracted from blood or tumor specimens using the QIAamp DNA Blood Mini Kit according to the manufacturer's instructions (QIAGEN, Hilden, Germany).

Demographic data, including age, sex, clinical stage, perineural invasion, and lymphovascular invasion were retrospectively obtained from the patients' medical records. Clinical staging was performed according to the American Joint Committee on Cancer (AJCC 7th edition) guidelines for tumor, node, and metastasis TNM classification (20). None of the patients which were enrolled in this study had received adjuvant chemotherapy or radiotherapy before surgery. WES data from the TCGA-HNSCC dataset was collected and downloaded from the Genomic Data Commons portal (https://portal.gdc.cancer.gov/). Using the TCGA-HNSCC dataset, all cases where the primary site was located at the tongue, lip, mouth floor, tonsil, gums, palate, or oropharynx, were included. This dataset was called "TCGA-OSCC" dataset. In total, 387 OSCC patients with somatic mutation data were analyzed herein. Only mutations in the coding regions and in splicing sites were retained, whereas mutations in the introns, intergenic regions, and in untranslated regions (UTR) were filtered out.

Whole-Exome Sequencing

WES was performed with the SureSelect Human All Exon v6 + UTR Enrichment Kit (Agilent, Santa Clara, CA), followed by sequencing on a NextSeq500 DNA sequencer (Illumina, San Diego, CA). Downstream analysis was performed as previously described (21). The software used for WES analysis is listed in **Table S1**. Somatic mutation was called using our pipelines were shown in **Figure S1**. Sequencing data was aligned to the human genome (NCBI build GRCh38/UCSC hg38) using BWA-MEM. SAMtools was used for the file format conversion from SAM to BAM. Six different programs were used to call somatic mutations, including two callers (Muse and SomaticSniper) for single nucleotide variants (SNVs) (22, 23) and four callers (Mutect2, Strelka2, VarScan2, and VarDict) for both SNVs and short insertion and deletion variants (indels) (9, 23–26). These variant callers were run with default parameters and further filtering of the data was based on the following criteria:

(1) Mutations called by Muse, SomaticSniper, and Mutect2 were labeled as "PASS" in the FILTER column, (2) Strelka2 algorithm with WES default parameters was used to identify somatic mutations (27). Only variants labeled "PASS" were considered high-quality variants (28). (3) Mutations called by VarScan2 were labeled as "Somatic" in the somatic status column, and (4) mutations called by VarDict were labeled as "StrongSomatic" in the INFO column. Thereafter, mutations outside the targeted region were removed. The filtered mutations were considered somatic mutations. The command line calls somatic mutations as described in **Table S2**. All WES data in this manuscript were submitted to Short Reads Archive under the BioProject accession PRJNA749133 and SRA Run Selector project (https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA749133&o=acc_s%3Aa).

The somatic mutations were annotated using Ensembl Variant Effect Predictor (version 102, https://asia.ensembl.org/Homo_sapiens/Tools/VEP) (21). Potential mutational driver genes in OSCC were identified and annotated using the InToGene platform (https://www.intogen.org/search) and Bailey et al. datasets (29, 30).

Filtering Strategies

To reduce false-positive somatic mutations which might originate from germline mutations or might have been accidentally generated during sample preparation, DNA amplification, sequencing, and ambiguous mapping (31), the following filter flags were used to annotate and filter somatic mutations: (1) removing common polymorphisms (SNPs): minor allele frequency in the 1000 Genomes Project or The Genome Aggregation Database (gnomAD) > 1%; (2) removing Panel-of-Normal (PoN): A normal panel was created from 50 normal samples using GATK (Genome Analysis Toolkit) tool CreateSomaticPanelOfNormals; (3) removing oxodG artifacts: oxidation-damaged base 8-oxodG (8-oxoguanine) was identified and filtered using GATK tools CollectSequencingArtifactMetrics and FilterByOrientationBias; (4) removing strand bias, multiallelic site, and clustered events: strand bias is a type of sequencing bias wherein one DNA strand is favored over another by the variant. A multiallelic site is a genomic locus that contains two or more alternative (Alt) alleles. Clustered events are several variants that are clustered in a region of the genome. These artifacts were estimated and filtered using the GATK tools CalculateContamination, GetPileupSummaries, and FilterMutectCalls; and (v) Alt allele count filter: we filtered out mutations with Alt alleles in tumor derived date (T_Alt) < 4 and mutations with Alt alleles in normal data (N_Alt) ≥ 4 (**Figure 2A**) (32).

Merging Results From Multiple Callers

Somatic mutations were called using multiple callers and stored as VCF (Variant Call Format) files. Following variant calling and filtering, VCF files from six tools were merged according to each sample ID and genomic position (e.g., ID-chr1-123). The number of variants hits detected for each mutation in the merge files was then counted. Thereafter, mutations that were not identified by two or more variant callers were removed. The tumor mutation burden (TMB) was calculated using the number of non-synonymous mutations per mega-base (Mb) in the target region of SureSelect Human All Exon v6 + UTR (91.08 Megabases). The target region BED file is available online at the SureDesign website (https://earray.chem.agilent.com/suredesign/).

Mutation Validation

Sanger sequencing and IGV was performed to validate somatic mutations (33, 34). For Sanger sequencing, individual primer sets designed by Primer3 (version 0.4.0) are listed in **Table S3**. Polymerase chain reactions (PCRs) were performed using the KAPA LongRange HotStart PCR Kit (KAPA Biosystems, Wilmington, MA, USA). Amplicons were sequenced on an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA) with the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems).

The top 20 most frequently mutated genes [mutated in at least 7/50 (14%) patients] were selected and examined for false-positive rates using IGV. Mutations were considered "true-positive" based on the following criteria: (1) number of Alt alleles < 3 in normal cells and ≥ 3 in tumors; (2) both forward and reverse strands have at least one mutant allele; (3) number of mismatches within a 40 bp window ≤3; and (4) allelic configurations of the mutation are multiallelic variants (21).

Visualization of WES Data

The 200 most frequently mutated genes [mutated in at least 4/50 (8%) patients] in our data and in the TCGA-OSCC dataset were selected for the creation of a circular plot using Circos-0.69 (http://circos.ca/software/). Track 1 (inner circle) visualizes the mutation frequency of the genes in the TCGA-OSCC dataset (**Figure 4**), whereas Track 2 (outer cycle) illustrates the genomic profile from our WES data.

Pathway Analysis

The 200 most frequently mutated genes in our study were imported into the analysis pipeline of the Ingenuity Pathway Analysis (IPA, QIAGEN, CA, USA; http://www.ingenuity.com/products/pathways_analysis.html). IPA was used to examine which canonical pathways were enriched within our candidate genes. In IPA, Fisher's exact test was used to determine whether a canonical pathway was enriched within a data set. A −log [*P* value] > 1.3 (corresponding to a *P* value of <0.05) was set as the threshold for statistical significance. The United States Food and Drug Administration (FDA) Table of Pharmacogenomic Biomarkers in Drug Labels was used to match our candidate genes against FDA-approved drugs (https://www.fda.gov/).

Statistical Analysis

Data are presented as mean ± standard error of mean. The Chi-squared test, Fisher's exact test, and Mann–Whitney U test were used for statistical analysis. The receiver operating characteristic curve (ROC) analysis was used to identify the optimal thresholds of read-depth (DP) and minor allele frequency (MAF) in mutation calling. Overall survival (OS) was defined as the duration from the first date of diagnosis to death or last date of follow-up. Kaplan–Meier analysis was used to compare the OS

between two groups. *P* values less than 0.05 were considered to be statistically significant.

RESULTS

Patient Characteristics

In this study, we have collected and analyzed tumor specimens and matched blood samples from 50 patients with OSCC. 48 patients were male and 2 were female, with an average age of 59.6 years (range: 40–89 years). All patients were confirmed to have squamous cell carcinoma. The most common primary sites were the bucca (28%, 14/50) and the gingiva (24%, 12/50). The detailed clinical characteristics of our study subjects are described in **Table 1**.

Variant Calling in OSCC

Six different variant calling tools were used to identify somatic mutations in the 50 paired tumor/normal samples: Muse, Mutect2, SomaticSniper, Sterlka2, VarScan2, and VarDict (**Figure 1**). All six callers were used to identify somatic SNVs. Moreover, somatic indels were identified using Mutect2, Strelka2, VarScan2, and VarDict. A total of 163,069 somatic mutations were found by the six callers using the default parameters in the target region (coding and splicing region). Muse, Mutect2, SomaticSniper, Sterlka2, VarScan2, and VarDict detected 10,019, 79,933, 9037, 58,070, 3656, and 37,240 somatic mutations, respectively (**Figure 1**). To reduce false-positive mutations from variant calling, filter flags were used to annotate and assess somatic mutations (**Figure 2A** and **Table S4**). Somatic mutations marked with filter flags, including "common SNP", "oxodG" oxidative damage, "StrandBias", "Multiallelic site", "Clustered events", and "PoN" variants were then filtered out. Furthermore, somatic mutations that were lacking sufficient evidence to be called a somatic mutation, such as those with low Alt alleles in the tumor sample (T_Alt < 4 alleles) and high Alt alleles in normal samples (N_Alt ≥4 alleles), were filtered out. As result, the filtered (PASS only) data contained 8,730, 23,691, 3,574, 54,846, 2,683, and 23,545 somatic mutations detected by Muse, Mutect2, SomaticSniper, Sterlka2, VarScan2, and VarDict, respectively (**Figure 1** and **Table S4**). As depicted in **Figure S2A**, somatic mutations marked "PASS" filter flag were successfully verified by IGV and Sanger sequencing. However, the somatic mutations which have been labeled "N_Alt ≥4," "T_Alt < 4," "Multiallelic site," and "Clustered event" were not confirmed by Sanger sequencing (**Figures S2B–E**). Furthermore, our results show that all variant callers except Mutect2 and VarDict mis-detected dinucleotide mutations as SNVs. In total, 53 dinucleotide mutations were identified in our study (**Table S5**), all of which were confirmed by Sanger sequencing (**Figure S3**). Finally, the VCF files from each caller were merged (**Figure 1**). Somatic mutations that were only identified by a single caller were excluded. Overall, the ≥2 caller agreeing data set only recovered 8% of the variants in the unfiltered data set. Our study identified 13,730 somatic mutations in 7,729 unique genes, including 3,057 synonymous, 8,541 missense, 17 start loss, 854 stop gain, 9 stop loss, 3 stop retain, 969 splicing site, 53 dinucleotide, 183 frameshift, and 44 in frame mutations (**Figure 1**).

TABLE 1 | Association between tumor burden and clinical parameters.

Parameter	N	mean ± SEM	*P*-value
Age			
≤ 60	23	1.403 ± 0.383	0.034*
> 60	27	2.312 ± 0.664	
Gender			
Male	48	1.94 ± 0.416	0.488
Female	2	0.786 ± 0.006	
T stage			
T1-3	13	1.351 ± 0.442	0.177
T4	37	2.085 ± 0.518	
N stage			
N0	28	2.282 ± 0.678	0.464
N+	22	1.400 ± 0.283	
Clinical stage			
I-III	11	0.801 ± 0.169	0.021*
IV	39	2.202 ± 0.502	
Differentiation			
Well	36	2.166 ± 0.545	0.538
Moderate-poor	14	1.194 ± 0.235	
Perineural invasion			
No	33	1.182 ± 0.334	0.301
Yes	17	2.039 ± 1.006	
Lymphovascular invasion			
No	39	1.885 ± 0.471	0.399
Yes	11	1.926 ± 0.764	
HPV status			
p16 negative	46	1.989 ± 0.433	0.453
p16 positive	4	0.804 ± 0.173	

Statistical test for comparing two groups by Mann Whitney U-test.
**P < 0.05.*

Effect of Filtering Criteria in Different Callers

Figure 2 and **Table S6** summarize the process of filtering mutations in the six callers. These variant callers detected a different amount of SNVs and indels (**Figures 2B, D**). For SNVs, Mutect2, Strelka2, and VarDict produced the largest number of unfiltered SNVs (70,532, 58,033, and 36,060, respectively), while Muse, SomaticSniper, and VarScan2 called the smallest number of SNVs (10,019, 9,037, and 3,430, respectively) (**Figure 2B** and **Table S6**). Moreover, Mutect2 and SomaticSniper had high rates of SNVs that did not satisfy the filtering criteria (non-PASS SNVs) (68.2% and 60.5%, respectively), whereas Muse, Strelka2, VarScan2, and VarDict had low rates of non-PASS SNVs (12.9%, 5.5%, 25.7%, and 37.1%, respectively). These non-PASS SNVs mainly consisted of "PoN" and "T_Alt < 4" in Mutect2 (39.9% and 15.1%, respectively) and "PoN" and "N_Alt >4" in SomaticSniper (46.0% and 13.2%, respectively) (**Figure 2C**). For indels, Mutect2 found more unfiltered indels compared to Strelka2, VarScan2, and VarDict (9,401, 37, 226, and 1,180, respectively) (**Figure 2D** and **Table S6**). After filtering, high rates of non-PASS indels were found in Mutect2, whereas Strelka2, VarScan2, and VarDict had low rates of non-PASS indels (86.6%, 37.8%, 40.7% and 33.9%, respectively). These non-PASS indels mainly consisted of "PoN" and "T_Alt < 4" in Mutect2 (32.6% and 30.4%, respectively), "Clustered events" and

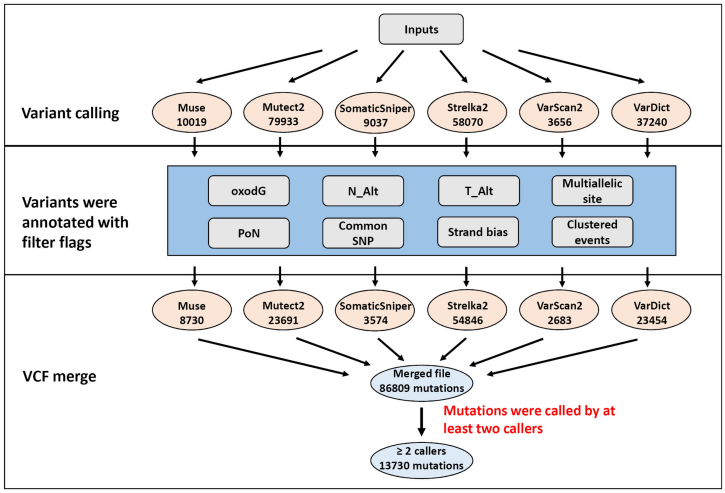


FIGURE 1 | Flowchart of the filtering strategy. Mutations were identified by multiple variant callers parametrized to identify potentially somatic mutations. The number of mutations for each filtering step is depicted in illustration.

"T_Alt < 4" in Strelka2 (13.5% and 10.8%, respectively), "PoN" in VarScan2 (24.8%), and "PoN" and "T_Alt < 4" in VarDict (17.4% and 14.7%, respectively) (**Figure 2E**). As such, approximately 23% (37,287/163,069) of unfiltered mutations consisted of PoN variants. Removing mutations marked with the "PoN" filter flag was a crucial step in reducing false-positive rates during WES analysis.

Muse and VarScan2 removed the lowest number of SNVs after filtering with filter flags and removing variants found by a single caller only (**Table S6**). Strelka2 removed the lowest number of indels after variant filtering. Mutect2 removed the largest number of both SNVs and indels after filtering with filter flags and removing variants called by a single caller only. Although SomaticSniper removed a large number of the SNVs selection after filtering with filter flags, it removed a smaller number of SNVs after removing variants called by a single caller. Strelka2 recovered 94.5% of the SNVs after filtering with filter flags but only recovered 19.5% of the same after removing variants called by a single caller. Moreover, Strelka2 found a very small number of indels only.

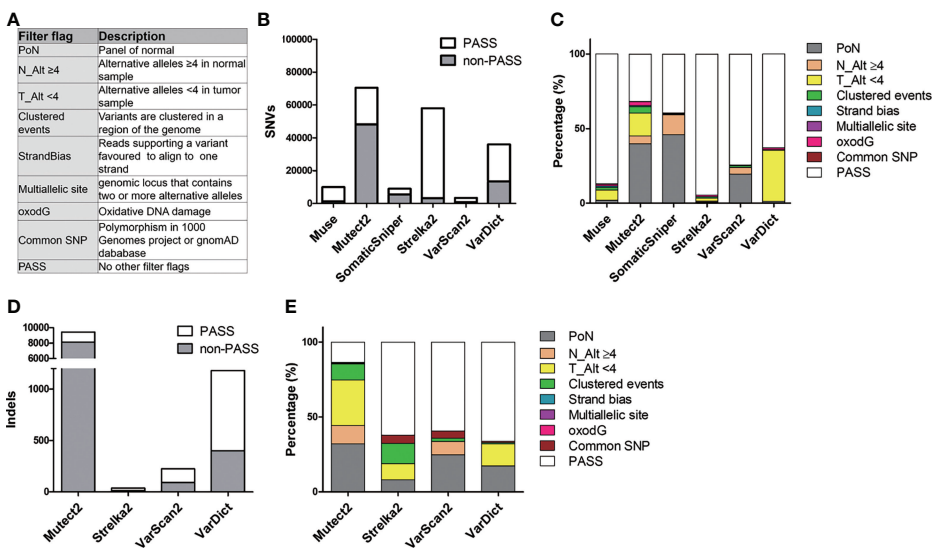


FIGURE 2 | Description of filter flag distribution in six callers. **(A)** Purpose of the filter flags. **(B)** The count of PASS and non-PASS SNVs for each caller. **(C)** The stacked bar chart illustrates the proportion of different filter flags in SNVs for each caller. **(D)** Representation of the amount of PASS and non-PASS indels. **(E)** Histogram of the proportion of different filter flags in indels.

Evaluation of Mutation Calling Performance

In order to evaluate the performance of variant calling, IGV was used to visualize and examine somatic mutation in the six callers. The 20 most frequently mutated genes [mutated in at least 7/50 (14%) patients] were utilized for this analysis. **Table S7** shows that the most frequent non-PASS genes during IGV examination were *MUC16*, *MUC19*, *KMT2D*, and *TTN*. To assess the reliability of the IGV examination, a number of mutations for each top 4 IGV non-PASS genes was selected for confirmation by Sanger sequencing. **Figure S4** shows that both IGV-passed and -non-passed mutations could be confirmed by Sanger sequencing. The results of the IGV evaluation were also consistent with those of Sanger sequencing (**Table S8**). In **Table S7**, *MUC16* and *MUC19* genes had a large number of false-positive mutations [2100/2121 (99.0%) and 1900/1906 (99.7%), respectively] that did not satisfy the IGV filtering criteria. The IGV screenshot (**Figure S5**) demonstrates that recurrent false-positive variants were observed at the *MUC16* and *MUC19* loci in our WES data. Therefore, *MUC16* and *MUC19* mutations were removed to reduce false-positive calls. A total of 2276 mutations were retained after removing mutations in *MUC16* and *MUC19*, which then were used to evaluate the validation statistics of the mutation callers.

Table S9 shows that filtered mutations had significantly higher rates of IGV-PASS compared to unfiltered mutations ($P < 0.001$). The filtered mutations were also associated with a significantly higher rate of IGV-PASS in Mutect2 and SomaticSniper (both $P < 0.001$ and $P < 0.001$). An increased number of callers agreeing on the mutations was associated with an increase in IGV-PASS rates. However, in ≥2, ≥3, ≥4, ≥5 and 6 caller agreeing groups, filtered mutations significantly increased IGV-PASS rates compared to unfiltered mutations. Both filtering with filter flags and selection with mutation calling times were important procedures to reduce false-positive calls.

The Effect of Read-Depth and Mutation Frequency on Variant Calling

Furthermore, this study examined whether DP and MAF were associated with false-positive variant calls. **Figure S6A** shows the MAF and DP for unfiltered mutations in different mutation calling times. Accordingly, our results show that mutations with high MAF (MAF ≥ 0.8) and high DP (DP ≥ 1000) were mainly distributed in the one caller agreeing group. After filtering with filter flags, the high MAF and high DP mutations were significantly reduced in the one caller group (**Figure S6B**). IGV examination data shows that mutations with high DP (DP ≥ 1000) were mainly distributed in the IGV non-PASS group (**Figure S6C**). The aforementioned data suggest that read depths of 1000 and a MAF of 0.8 can be selected to eliminate false-positives mutations.

Receiver operating characteristic (ROC) curve was used to evaluate the optimal cut-off value of DP and MAF in distinguishing IGV-PASS mutations (**Figures S6E, G**). The AUCs for DP and MAF were 0.478 (95% CI: 0.459–0.496) and 0.607 (95% CI: 0.585–0.628), respectively. At a threshold of 175 reads for DP, the false positive rate and false negative rate were 51% and 45.1% in separating IGV-PASS mutations patients from normal IGV non-PASS mutations. At a threshold of 0.182 for MAF, the false positive rate and false negative rate were 38.5 and 86.3%, respectively. The **Figures S6E, G** showed the relationship between false positive rate and false negative rate with DP or MAF, respectively.

Estimation of Tumor Mutational Burden in OSCC

TMB was calculated as the number of somatic mutations in the coding region per Mb. In the TCGA-OSCC dataset, 71,890 non-synonymous mutations were identified in 387 patients (185.76 mutations per patient). In our unfiltered data, a total of 163,069 mutations were identified in 50 patients (3261.38 mutations per patient). Our findings show that the mean TMB was 35.81 mutations/Mb per patient. After filtering procedures, 13,730 mutations were retained (274.6 mutations per patient), with the TMB decreasing to 3.01 mutations/Mb per patient in filtered mutations. As depicted in **Table 1**, high TMB was significantly associated with old age ($P = 0.034$) and advanced clinical stage ($P = 0.021$). The median TMB (0.96 mutations/Mb per patient) was considered the cut-off point for assessing outcomes in OSCC. Patients with higher TMB exhibited a poorer outcome compared to those with lower TMB ($P = 0.041$, **Figure 3A**). In TCGA-OSCC dataset, patients with higher TMB also had a poorer outcome compared to those with lower TMB ($P = 0.026$, **Figure 3B**).

Genomic Profiling in OSCC

CIRCOS plots were used to illustrate the mutational landscape in our study and in the TCGA-OSCC dataset (**Figure 4**). The outer cycle illustrates the genomic profile from our WES data, with the most frequently mutated genes being *TP53* (62%), *FAT1* (40%), *NOTCH1* (28%), and *TTN* (26%) (**Figure 4**). The inner circle illustrates the mutation frequency of the genes in the TCGA-OSCC dataset, with the most mutations observed in *TP53* (68%), *TTN* (42%) *FAT1* (26%), and *CDKN2A* (22%). The mutational landscape in our study appears to be similar to that in the TCGA-OSCC dataset.

We found five novel mutated genes in at least six different patients (≥10%) included herein, which were not detected in the TCGA-OSCC dataset (**Table S10**). The mutation frequencies of *SKOR2*, *CPLANE1*, *CCDC168*, *STARD9*, and *NSD2* were 14%, 12%, 12%, 12%, and 10%, respectively. The InToGene platform and Bailey et al. data sets were used to predict potential mutational driver genes in OSCC. A total of 53 recognized mutational driver genes were found therein, among whom *TP53* and *FAT1* had high mutation rates (**Table S11**). **Table 2** shows the differences in the mutation frequency distribution between our most frequent genes (MAF ≥ 10%) and those in the TCGA-OSCC dataset. Accordingly, among our most frequent genes, 73 genes were identified to have high MAF, whereas only 1 had low MAF. Several of these genes we found were tumor suppressor genes, including *FAT1*, *EPHA2*, *ASXL1*, *PTPRT*, *USP9X*, *IGF2R*, *SPTBN1*, and *PLCB3* (39–43, 45, 47, 48),

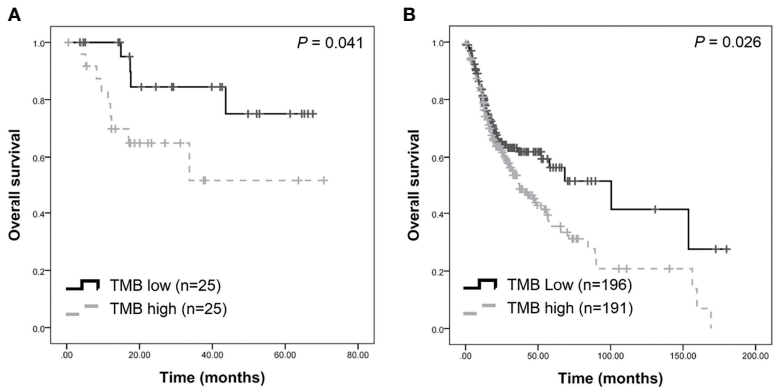


FIGURE 3 | Survival analysis of tumor burden (TMB) in patients with OSCC. Higher TMB were associated with poorer survival in our dataset **(A)** and TCGA-OSCC dataset **(B)**.

whereas *ARHGEF10L* and *NSD2* function as oncogenes (44, 55). *TRRAP* has been found to be involved in the regulation of stemness in ovarian cancer stem cells (35), while *DNMT1* and *EPHB1* can regulate tumor progression (38, 46). Other genes have been reported to been involved in the control of proliferation, invasion, and apoptosis in cancer cells (36, 37, 49–54, 56–59).

Furthermore, this study examined the relationship between non-synonymous mutation status and clinical parameters in the 20 most frequently mutated genes (**Figure 5** and **Table S12**). Accordingly, somatic mutation in *ASXL1* was significantly associated with older age (P = 0.010) (**Table S12**). In addition, *FAT1* mutations were significantly associated with advanced clinical stage (P = 0.033) and marginally significantly associated with T stage (P = 0.050). Histological grade was significantly associated with *PKD1L1* mutations (P = 0.044) and was marginally significantly associated with *FAT3* mutations (P = 0.087). The presence of *TTN* mutations was significantly associated with perineural invasion (P = 0.038). Mutation of *FMN2* was marginally significantly associated with T stage (P = 0.093) and histological grade mutations (P = 0.087). Survival analysis revealed that patients with *CASP8* (log-rank P = 0.004), *CUBN* (log-rank P = 0.025), and *USP9X* mutations (log-rank P = 0.018) had significantly reduced OS rates (**Figures 6A–C**). **Figure S7** showed that distribution of the 20 most frequently mutated genes and clinical features in the TCGA-OSCC patients. In the TCGA-HNSCC dataset, *USP9X* mutation was found to be associated with poor OS (log-rank P = 0.010, **Figure S8C**). *CASP8* mutation was marginal significantly associated with OS (log-rank P = 0.070, **Figure S8C**). However, there were no significant differences in OS between *CUBN* mutation group and non-*CUBN* mutation group (log-rank P =0.885, **Figure S8B**).

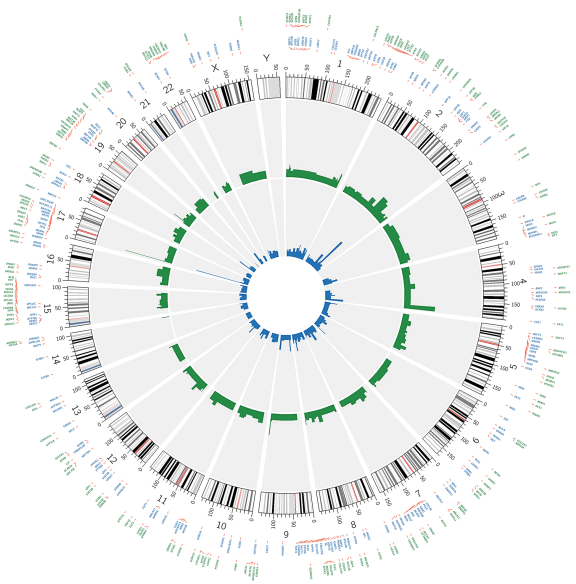


FIGURE 4 | Circos plot showing the 200 most frequently mutated genes in OSCC patients. The inner circle visualizes the mutation frequency of genes in the TCGA OSCC dataset. The outer cycle illustrates the genomic profile which was observed in our study.

Molecular Pathway Analysis

Pathway enrichment analysis of the 200 most frequently mutated genes was performed using the IPA. Accordingly, upstream regulator analysis in OSCC (**Figure S9**) showed that *AKT*, *TP53*, and *ERK* were the most predicted upstream regulators controlling different gene clusters in OSCC. **Table 3** shows that four canonical pathways had a P value < 0.05 using IPA, including the clathrin-mediated endocytosis signaling, NFκB signaling, PEDF signaling, and the calcium signaling pathways. Furthermore, we evaluated the association between OS and mutations of genes in canonical pathways. Among the patients included herein, 27 (54%) had a mutation in the NFκB signaling-related gene set, including *CARD10*, *CASP8*, *EP300*, *FGFR1*, *IGF2R*, and *PIK3CA* (**Figure 7A**). Mutations in *CARD10*, *CASP8*, *EP300*, *FGFR1*, *IGF2R*, and *PIK3CA* were observed in 5 (10%), 11 (22%), 6 (12%), 4 (8%), 6(12%), and 7 (14%) patients,

TABLE 2 | List of most frequently mutated genes in our study showed a statistically significant difference in mutation rates compared to TCGA-OSCC dataset.

Genes	Our MAF	TCGA MAF	Driver gene	Involved in cancer[ref]	Genes	Our MAF	TCGA MAF	Driver gene	Involved in cancer[ref]
Our MAF > TCGA MAF					Our MAF > TCGA MAF				
FAT1	0.4	0.26	Y	Y (32)	*SBNO2*	0.10	0.02	N	
RYR2	0.2	0.09	N		*FBN3*	0.10	0.03	N	
FMN2	0.18	0.06	N	Y (35)	*TNRC6B*	0.10	0.02	N	Y (36)
ABCA13	0.16	0.07	N		*CFAP74*	0.10	0.01	N	
EPHA2	0.16	0.06	Y	Y (33)	*MGA*	0.10	0.03	N	
HYDIN	0.16	0.04	N		*CFAP46*	0.10	0.02	N	
ALPK3	0.14	0.01	N		*MAGEL2*	0.10	0.02	N	
ASXL1	0.14	0.03	Y	Y (34)	*ABCB5*	0.10	0.03	N	Y (37)
EPPK1	0.14	0.06	N	Y (38)	*COL24A1*	0.10	0.03	N	
PKD1L1	0.14	0.03	N		*SPTBN1*	0.10	0.03	N	Y (39)
PTPRT	0.14	0.04	Y	Y (40)	*CLTCL1*	0.10	0.02	N	
SKOR2	0.14	#N/A	N		*NSD2*	0.10	#N/A	N	Y (41)
SORCS3	0.14	0.05	N		*PLCB3*	0.10	0.02	N	Y (42)
USP9X	0.14	0.04	Y	Y (43)	*DNHD1*	0.10	0.02	N	
WDFY3	0.14	0.04	N		*KIF26A*	0.10	0.01	N	
DOCK10	0.12	0.02	N		*DNMT1*	0.10	0.02	N	Y (44)
CPLANE1	0.12	#N/A	N		*COL6A5*	0.10	0.01	N	
CCDC168	0.12	#N/A	N		*DCHS1*	0.10	0.03	N	
ITPR2	0.12	0.04	N		*C2CD3*	0.10	0.02	N	
TRRAP	0.12	0.05	Y	Y (45)	*GOLGA3*	0.10	0.02	N	
LAMA5	0.12	0.04	N	Y (46)	*COL20A1*	0.10	0.01	N	
FBN2	0.12	0.04	Y		*MICAL3*	0.10	0.03	N	
STARD9	0.12	#N/A	N		*BOC*	0.10	0.03	N	
DNAH1	0.12	0.04	N		*ADAMTS13*	0.10	0.01	N	
CACNA1A	0.12	0.03	N		*CARD10*	0.10	0.02	N	Y (37)
IGF2R	0.12	0.03	N	Y (47)	*NBEAL1*	0.10	0.03	N	
ARHGEF10L	0.12	0.01	N	Y (48)	*UBR1*	0.10	0.02	N	
KIF1A	0.12	0.02	N	Y (49)	*DNAH6*	0.10	0.01	N	
SCN10A	0.12	0.03	N		*ZNF407*	0.10	0.02	N	Y (50)
TRPC6	0.12	0.02	N	Y (51)	*URB2*	0.10	0.01	N	
EP400	0.12	0.04	N		*LYST*	0.10	0.03	N	
MYOF	0.12	0.02	N	Y (52)	*ROCK2*	0.10	0.02	N	Y (53)
SBF1	0.10	0.02	N	Y (54)	*OTOF*	0.10	0.03	N	
PPIP5K2	0.10	0.02	N		*EPHB1*	0.10	0.03	N	Y (55)
UNC80	0.10	0.01	N		*AUTS2*	0.10	0.02	N	
OTOF	0.10	0.03	N		*DNAH14*	0.10	0.00	N	
Our MAF < TCGA MAF									
TTN	0.26	0.47	N						

Fisher's exact test and chi-square were used as a test for statistical significance.
MAF, minor allele frequency.
Ref, references were cited using superscript numerals.

respectively. Patients with mutations in the NFκB signaling-related gene set had poorer prognosis compared to those without such mutations (**Figure 6D**). However, no significant correlation between mutation status in NFκB signaling-related gene set and OS was found in TCGA-OSCC dataset (**Figure S8D**).

Mutations in the calcium signaling-related gene set were observed in 27 (54%) patients (**Figure 7B**). Mutations in the calcium signaling-related gene set, including *ADCY2*, *PLCB1*, *PLCB3*, *ITPR1*, *ITPR2*, and *PIK3CA*, were observed in 4 (8%), 4 (8%), 5 (10%), 4 (8%), 6 (12%), and 7 (14%) patients, respectively. Mutations in the calcium signaling-related gene set were associated with poor outcomes in OSCC (**Figure 6E**). In TCGA-OSCC dataset, there was no significant correlation of mutation status in calcium signaling-related gene set with OS (**Figure S8E**).

Thereafter, identified FDA-approved drugs associated with our candidate genes. Accordingly, seven candidate genes were found which may be targeted by FDA-approved drugs. They are involved in the regulation of the NFκB signaling-related and calcium signaling-related pathways (**Figure 7C**). These genes include *PIK3CA* and six receptor tyrosine kinases, namely *FGFR1*, *FGFR2*, *FGFR3*, *EGFR*, *ALK*, and *ROS1* (**Figure 7C**).

DISCUSSION

False-positive calls are a major problem in the detection of somatic mutations. One of the most effective filters has encoded the expected distribution of alternate allele read counts at every genomic position, based on a large panel of

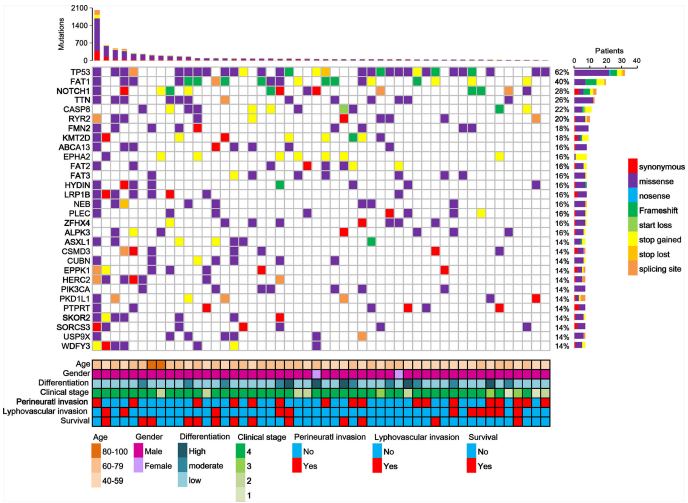


FIGURE 5 | Distribution of the 20 most frequently mutated genes in the OSCC patients. Each column represents an individual OSCC patient, and each row denotes a gene and clinical features. Clinical features and mutation types are color coded as indicated.

8000 TCGA normals (PoN) (31). For each genomic position, the "PoN" encodes the distribution of alt read counts across all TCGA normals. This filter tags a somatic variant call if its observed read count is consistent with the "PoN" based on a likelihood test. This allows calls with several supporting reads to be retained when they occur at a site with low allele-fraction (AF) sequencing noise in the "PoN". To remove high AF artifacts, all somatic calls at a site with recurrently high AF across the "PoN" are removed, whereas those with several supporting reads at the same locus are retained.

Studies have discovered that somatic mutations caused by several carcinogenic and mutagenic chemicals may induce cancer development (60). Somatic, but not germ line, mutations have been found to cause the change from normal cells to cancer cells. and thus being responsible for malignancies (61). Therefore, identifying and removing germ line mutations

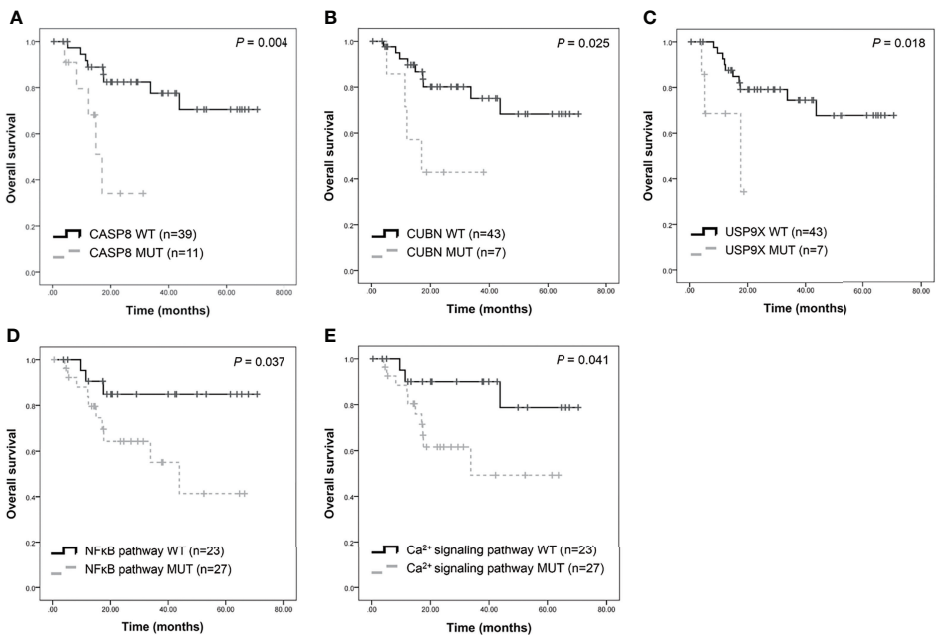


FIGURE 6 | Survival analysis of candidate genes in patients with OSCC. Kaplan-Meier plots of overall survival based on *CASP8* mutation **(A)**, *CUBN* mutation **(B)**, *USP9X* mutation **(C)**, NFκB **(D)**, and calcium signaling pathway **(E)** status.

TABLE 3 | The canonical pathways of the top 200 frequently mutated genes in HNSCC.

Ingenuity Canonical Pathways	-log (P-value)	Molecules
Clathrin-mediated Endocytosis Signaling	0.856	*APOB*, *CLTCL1*, *EPHB2*, *PIK3CA*, *USP9X*
NFκB Signaling	1.39	*CARD10*, *CASP8*, *EP300*, *FGFR1*, *IGF2R*, *PIK3CA*
PEDF Signaling	1.53	*CASP8*, *PIK3CA*, *ROCK2*, *TP53*
Calcium signaling	1.8	*ADCY2*, *CASP8*, *ITPR1*, *ITPR2*, *PIK3CA*, *PLCB1*, *PLCB3*

PEDF, Pigment epithelium-derived factor.

from the available data is necessary in order to identify true somatic mutations. Aside from comparing paired normal and cancer tissue, another method to differentiate germ line mutations from somatic ones in cancerous tissue is to establish a normal panel. A somatic variant call is tagged by this filter if its observed read count is consistent with the "PoN" based on a likelihood test. Likewise, a common germ line site would have recurring high allelic fractions across the "PoN". Moreover, a call at that site with a similarly high AF will be flagged.

To remove germ line events or high AF artifacts, approximately 23% of the hits found by mutation calling were initially removed. Different callers had diverse results, with 46.0% and 39.0% of the SNVs not satisfying the filtering criteria in SomaticSniper and Mutect2, respectively. In addition, the Broad Panel of normals flagged almost 30% of the calls in the full set, which were also removed following TCGA data release policies.

There was indeed highly confident evidence for true mutations in our data. Although the ability to call a variant depends on several factors, two key factors include the coverage (or DP) at a site and the frequency of the alternative (i.e. non-reference) allele frequency (18). However, the repeat-rich sequences which are present within centromeric regions and acrocentric short arms, are often fully represented in whole-genome short-read data sets and contribute to inappropriate alignments and high DP signals that localize into a small number of assembled homologous regions (62). Consequently, these regions often provide artifactual peak false-positive calls (62). Efforts to mitigate these mapping errors frequently involve providing an additional 'decoy' database or a collection of useful sequences which are missing from the human genome that can help to ensure proper alignment (63). "oxodG" and "Clustered Events" can reduce false-positives (64), with almost 1.9% of the calls in our study having been screened out.

After filtering with "PoN", "OxodG", "Clustered Events", "StrandBias", " Multiallelic site", and "Common SNP", coverage should no longer be the major factor for false-positive calling. After filtering, IGV non-PASS mutations did not have higher rates of low DP (<100 DP) compared to IGV-PASS mutations (**Figure S6B**). After removing the low AF noise, even an AF of lower than 1% could be identified and validated by the IGV as true mutation.

The possible impact of amplification errors and content bias related to the library method used should nonetheless be considered. Given that potential sources of error may be addressed through assay design, these should be considered early in the design phase of test development (65). Moreover, in cases with rearrangements, isolated neighboring regions may originate from genomic areas which are very distant from the intended or predicted targets. The fragment sizes resulting from shearing and other fragmentation mechanisms will have a

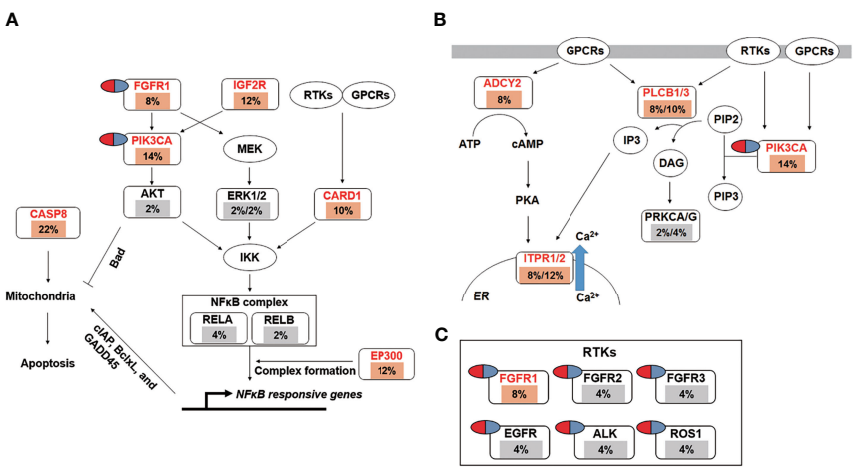


FIGURE 7 | The landscape of major altered genes and significantly altered pathways in OSCC. The NFκB signaling **(A)** and calcium signaling pathways **(B)** were identified by IPA canonical pathway analysis. Pill illustrations show FDA approved drugs which are targeting receptor tyrosine kinase families and *PIK3CA*. Genes were assigned to each pathway and represented with their mutation frequencies (orange font: top 200 frequently mutated genes [≥8%] which were imported into the IPA analysis; gray font: genes with <8% mutation frequencies which were not included in IPA analysis).

considerable influence on the outcome of the analysis. Shorter fragments will be captured with higher specificity than longer fragments, as the former will contain a lower proportion of off-target sequences. On the other hand, longer reads are expected to map to the reference sequence with less accordance than shorter reads (66). As such, several large genes, including *TTN*, *MUC19*, *MUC16*, and *KMT2D*, had more false-positive results that could be screened by our study.

Despite removing the normal panel SNPs and repeat-rich sequences, false-positive mutations still are present (31). Thus, combining the results of several mutation callers is important to reduce false-positives while maintaining sensitivity (18), which is helping with the identification of true somatic mutations. However, this approach is raising some questions, such as how many mutation callers should be utilized. Analyzing the number of true and false detections in any combination of mutation callers across all replicates suggests that the combination of at least of two callers has a significantly better performance compared to individual callers (18, 31). Fortunately, intersecting mutation callers did not diminish the amount of the identified true mutations. Karimnez found only 1 out of 343 gold-standard mutations was missing, when intersectioning five programs (18). Comparing our intersectioning results with those of other WES studies and the TCGA database yielded differences in the main mutated genes and in the mutated allele frequencies (**Table 2**). We conducted IGV and Sanger sequencing to demonstrate that our data was more precise in detecting true mutations (**Table S8**).

TMB is defined as the total number of somatic gene coding errors, base substitutions, insertions, or deletions per megabase of tumor tissue (67). According to the studies conducted previously, the estimated TMB value for each sample is defined as the total mutation frequency divided by the length of the combined human exons (38 Mb) (68–70). We determined the TMB in patients with tumor and blood samples sufficient for WES (71, 72). Different algorithms have been found to produce varying results regarding the tumor burden. Accordingly, in several studies, the TMB of OSCC ranged from 4 to 104 mutations (7, 8, 10, 73, 74), whereas others have found a TMB of 2.079 (68), 2.96 (69), and 4.7 (70), mutations when the mutation count was divided by the length of the complete human exon data (38 Mb). We therefore suggest to use the target region's true length when counting the mutation burden (i.e., the SureSelect Human All Exon v6 + UTR 91.08 Mb in the current study) to indicate the true burden. Given that the total size of the WES was not uniform across various studies, using 38 MB as the denominator to calculate the TMB from the TCGA database will incur bias distortion. As such, our data evaluates the mutation burden more precisely, which then can be correlated with clinical parameters, including disease stage and OS rate of OSCC. Interestingly, age has been found to be only marginally related to TMB. Nonetheless, more samples may be collected and analyzed for further elucidation.

Several studies have identified landscape mutations in HNSCC (7, 8, 13, 14, 34, 75–80). When somatic variant callers were first compared, a surprisingly large number of unique calls was identified for each method (17). Recently, a few papers have used a combination of Mutect2 and Strelka to screen for mutations (13, 77, 80). Accordingly, Nisa used Mutect and VarScan2 to identify mutation patterns between metastatic and recurrent HNSCC (79). Nonetheless, there are still differences in the callers, parameters, and filters which were used in the different projects (31). Ideally, future variant calling and filtering efforts should use robust benchmarking to evaluate various combinations of callers, filters, and parameters and determine which callers and filters are optimal for OSCC (31). The study presented here used multiple pipelines to call mutations. The combination of all six programs was able to identify 98.6% of the actual mutations but resulted in a 50% loss of point mutations. After filtering and combination, two different programs were found to achieve a specificity of 88.7%.

Using the intersection of multiple mutation calling pipelines helped to identify true novel somatic mutations, including *SKOR2*, *CCDC168*, *STARD9*, and *CFAP74*, with allele frequencies of 0.14, 0.12, 0.12, and 0.1. Still, the verification of this pathway requires further evaluation.

The effectiveness of different therapeutic modalities is largely dependent on the mutational profile of a tumor, given that genetic alterations are likely to confer new oncogenic potential to cancer cells (81). The precise targeting of these alterations, together with a modifications of the treatment regimen, decreases therapeutic resistance, possibly saving countless patients from morbidity and mortality. Modern research has unveiled a new mutational landscape for oral cancer and factors contributing to the resistance and therapeutic efforts (2). The p53 and RB pathways are playing a key role in cell cycle control and have been found to be frequently abrogated in HPV−negative tumors, which may then be followed by the activation of PI3K pathway (2). No mutually exclusive or concordant mutations exist with those in *PIK3CA*. FAT1–NOTCH1–AJUBA pathway alterations that impact β-catenin signaling might form another route through which carcinogenesis is triggered. The role of *EGFR* and other growth factor receptors and their ligands warrants further study with regard to inappropriate models by genome editing, although some data suggest that amplification or mutation of these proteins may serve as an alternative to *CCND1* amplification (2). However, several genes are mutated at only very low frequencies with often unclear functional consequences. Using an intersectional analysis for mutation calling, the mutation pattern is not completely like reported before. *APOB*, *CLTCL1*, *EPHB2*, *PIK3CA*, and *USP9X* are involved with clathrin-mediated endocytosis (CME) signaling. CME is responsible for the uptake of transmembrane receptors and transporters, remodeling the plasma membrane composition in response to environmental changes, and regulating cell surface signaling (40). Studies have found that EGF-dependent cell proliferation is enhanced in CME-defective cells (82). Mutant *USP9X* has been found to be associated with poor prognosis. As well, mutations in *ADCY2*, *CASP8*, *ITPR1*, *ITPR2*, *PIK3CA*, *PLCB1*, and *PLCB3* (which are involved in calcium signaling) are associated with poor prognosis. These genes may be targeted by new compounds to find out if the can improve OSCC prognosis.

A confident caller with a low false-positive profile is better suited for the discovery of driver genes, given that the removal of false-positive noise is helping researchers with the identification

of significantly recurring patterns. Once the significant driver genes have been identified, a second pass over the mutation data set can identify calls of lower confidence that could provide additional examples of the gene of interest. Our results show that several genes, including *ASXL*, *FAT1*, *PKD1L1*, *TTN*, *FAT3*, and *FMN2*, are associated with clinicopathological parameters.

Although the MC3 program produced high-quality calls within each tumor-specific analysis group, differences in the callers, parameters, and filters used still were present from one project to another (31). Ellrott et al. used multiple genomic pipelines to identify mutation calling of tumor exomes by building standardized genomic analysis pipelines which can be massively deployed to tens of thousands of samples. However, care should be taken when analyzing a wide variety of cohorts (31).

CONCLUSIONS

After performing the identification of mutations with an array of six mutation callers adopted by different analysis centers, this study demonstrates that consensus calling outperformed single algorithms both regarding sensitivity and validation status. Finally, the use of consistent methods for calling enhances the utility of this resource in future endeavors to compare the molecular makeup across different studies. The results of this effort provide the integral components which are necessary for future studies in somatic variant calling.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, BankIt2445141.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Mackay Memorial Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conception: C-JL. Lab work: H-WC and C-HC. Interpretation or analysis of data: L-HL, C-JL, and C-HC. Preparation of the manuscript: L-HL, K-WC, and C-JL. Revision for important intellectual content: C-JL. Supervision: C-JL and K-WC. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from MacKay Memorial Hospital (MMH-E-105-12 and MMH-E-108-12) and the Ministry of Science and Technology, Taiwan (MOST 108-2314-B-195 -002 -MY2 and MOST 105-2314-B-195-005-MY3).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.741626/full#supplementary-material

Supplementary Figure 1 | Somatic mutation calling pipelines. Aligned data were analyzed by six differently callers to generate VCF files. VCF files were further filtered out sequence context artefacts and germline variants.

Supplementary Figure 2 | Validation of filter flags by Sanger sequencing and IGV. Variants that were marked with (A) "PASS", (B) "N_Alt", (C) "T_Alt", (D) "Multiallelic site", and (E) "Cluster event" were examined by direct sequencing and IGV (screenshot). Except for PASS variants, variants which had been marked with other filter flags were considered to be false positive variants.

Supplementary Figure 3 | Confirmation by Sanger sequencing of dinucleotide mutations called by Mutect2. Sanger sequencing and IGV screenshot of dinucleotide mutations in **(A)** *FAN1* chr15:30925801_GG>TT and **(B)** *RAI1* chr17:17798486_GG>CA.

Supplementary Figure 4 | Validation of IGV filtering results by Sanger sequencing. IGV PASS variants were consistent with those of the Sanger sequencing in **(A)** *MUC16*, **(C)** *MUC19*, and **(E)** *KMT2D*. IGV non-PASS variants were not detected by Sanger sequencing in **(B)** *MUC16*, **(D)** *MUC19*, and **(F)** *KMT2D*. Solid arrows indicate the positions of the mutations. Open arrows indicate the positions of polymorphisms.

Supplementary Figure 5 | Recurrent false positives in *MUC16* and *MUC19*. In the zoomed-out pane, IGV visualizes that both in normal and tumor samples, a large number of variants were detected in (A) the *MUC16* and (B) the *MUC19* locus. These observations were not only made in in patients 1074 and 700 but also in other patients.

Supplementary Figure 6 | Association between mutation frequency and read depth with false positive variant calls. Distribution of mutation frequency and read depth in (A) unfiltered data, (B) filtered data, and (C) IGV examination data. Receiver operating characteristic (ROC) curve was performed to evaluate the threshold value of DP **(D)** and MAF **(F)** in distinguishing IGV-PASS mutations. Distribution of false positive rate (blue) and false negative rate (red) base on DP **(D)** or MAF **(F)** values.

Supplementary Figure 7 | Distribution of the 20 most frequently mutated genes in the TCGA-OSCC patients. Each column represents an individual OSCC patient, and each row denotes a gene and clinical features. Clinical features and mutation types are color coded as indicated. Data were extracted from the TCGA database (https://portal.gdc.cancer.gov/).

Supplementary Figure 8 | Survival analysis of candidate genes in patients with TCGA-OSCC. Kaplan-Meier plots of overall survival based on *CASP8* mutation **(A)**, *CUBN* mutation **(B)**, *USP9X* mutation **(C)**, NFκB **(D)**, and calcium signaling pathway **(E)** status.

Supplementary Figure 9 | Regulator Effect networks identified by IPA in top 200 frequently mutated genes. We imported the 200 most frequently mutated genes into IPA, these genes were mutated in at least 7 patients (highlighted in blue). **(A)** *Akt* **(B)**, *TP53*, and **(C)** *ERK1/2* were located in the core that is the most important factor in our dataset. The red frame indicates that the genes were associated with survival outcomes in the Kaplan-Meier analysis.

REFERENCES

1. *Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019*. World Health Organization (WHO) - World Health Organization (2020).
2. Leemans CR, Snijders PJF, Brakenhoff RH. The Molecular Landscape of Head and Neck Cancer. *Nat Rev Cancer* (2018) 18:269–82. doi: 10.1038/nrc.2018.11
3. Boute P, Page C, Biet A, Cuvelier P, Strunski V, Chevalier D. Epidemiology, Prognosis and Treatment of Simultaneous Squamous Cell Carcinomas of the Oral Cavity and Hypopharynx. *Eur Ann Otorhinolaryngol Head Neck Dis* (2014) 131:283–7. doi: 10.1016/j.anorl.2013.10.003
4. Gluckman JL. Synchronous Multiple Primary Lesions of the Upper Aerodigestive System. *Arch Otolaryngol* (1979) 105:597–8. doi: 10.1001/archotol.1979.00790220031007
5. Braakhuis BJ, Tabor MP, Leemans CR, van der Waal I, Snow GB, Brakenhoff RH. Second Primary Tumors and Field Cancerization in Oral and Oropharyngeal Cancer: Molecular Techniques Provide New Insights and Definitions. *Head Neck* (2002) 24:198–206. doi: 10.1002/hed.10042
6. Gupta S, Kong W, Peng Y, Miao Q, Mackillop WJ. Temporal Trends in the Incidence and Survival of Cancers of the Upper Aerodigestive Tract in Ontario and the United States. *Int J Cancer* (2009) 125:2159–65. doi: 10.1002/ijc.24533
7. Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, et al. Exome Sequencing of Head and Neck Squamous Cell Carcinoma Reveals Inactivating Mutations in NOTCH1. *Science* (2011) 333:1154–7. doi: 10.1126/science.1206923
8. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The Mutational Landscape of Head and Neck Squamous Cell Carcinoma. *Science* (2011) 333:1157–60. doi: 10.1126/science.1208130
9. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nat Genet* (2013) 45:1113–20. doi: 10.1038/ng.2764
10. Cancer Genome Atlas N. Comprehensive Genomic Characterization of Head and Neck Squamous Cell Carcinomas. *Nature* (2015) 517:576–82. doi: 10.1038/nature14129
11. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational Landscape and Significance Across 12 Major Cancer Types. *Nature* (2013) 502:333–9. doi: 10.1038/nature12634
12. Schmidt H, Kulasinghe A, Kenny L, Punyadeera C. The Development of a Liquid Biopsy for Head and Neck Cancers. *Oral Oncol* (2016) 61:8–11. doi: 10.1016/j.oraloncology.2016.07.014
13. Su SC, Lin CW, Liu YF, Fan WL, Chen MK, Yu CP, et al. Exome Sequencing of Oral Squamous Cell Carcinoma Reveals Molecular Subgroups and Novel Therapeutic Opportunities. *Theranostics* (2017) 7:1088–99. doi: 10.7150/thno.18551
14. C. India Project Team of the International Cancer Genome. Mutational Landscape of Gingivo-Buccal Oral Squamous Cell Carcinoma Reveals New Recurrently-Mutated Genes and Molecular Subgroups. *Nat Commun* (2013) 4:2873. doi: 10.1038/ncomms3873
15. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of Mutational Processes in Human Cancer. *Nature* (2013) 500:415–21. doi: 10.1038/nature12477
16. Alexandrov LB, Ju YS, Haase K, Van Loo P, Martincorena I, Nik-Zainal S, et al. Mutational Signatures Associated With Tobacco Smoking in Human Cancer. *Science* (2016) 354:618–22. doi: 10.1126/science.aag0299
17. Kim SY, Speed TP. Comparing Somatic Mutation-Callers: Beyond Venn Diagrams. *BMC Bioinformatics* (2013) 14:189. doi: 10.1186/1471-2105-14-189
18. Karimnezhad A, Palidwor GA, Thavorn K, Stewart DJ, Campbell PA, Lo B, et al. Accuracy and Reproducibility of Somatic Point Mutation Calling in Clinical-Type Targeted Sequencing Data. *BMC Med Genomics* (2020) 13:156. doi: 10.1186/s12920-020-00803-z
19. Liu CJ, Lin SC, Chen YJ, Chang KM, Chang KW. Array-Comparative Genomic Hybridization to Detect Genomewide Changes in Microdissected Primary and Metastatic Oral Squamous Cell Carcinomas. *Mol Carcinog* (2006) 45:721–31. doi: 10.1002/mc.20213
20. Brandwein-Gensler M, Smith RV. Prognostic Indicators in Head and Neck Oncology Including the New 7th Edition of the AJCC Staging System. *Head Neck Pathol* (2010) 4:53–61. doi: 10.1007/s12105-010-0161-y
21. Lin SC, Lin LH, Yu SY, Kao SY, Chang KW, Cheng HW, et al. FAT1 Somatic Mutations in Head and Neck Carcinoma Are Associated With Tumor Progression and Survival. *Carcinogenesis* (2018) 39:1320–30. doi: 10.1093/carcin/bgy107
22. Fan Y, Xi L, Hughes DS, Zhang J, Zhang J, Futreal PA, et al. MuSE: Accounting for Tumor Heterogeneity Using a Sample-Specific Error Model Improves Sensitivity and Specificity in Mutation Calling From Sequencing Data. *Genome Biol* (2016) 17:178. doi: 10.1186/s13059-016-1029-6
23. Larson DE, Harris CC, Chen K, Koboldt DC, Abbott TE, Dooling DJ, et al. SomaticSniper: Identification of Somatic Point Mutations in Whole Genome Sequencing Data. *Bioinformatics* (2012) 28:311–7. doi: 10.1093/bioinformatics/btr665
24. Ewing AD, Houlahan KE, Hu Y, Ellrott K, Caloian C, Yamaguchi TN, et al. Combining Tumor Genome Simulation With Crowdsourcing to Benchmark Somatic Single-Nucleotide-Variant Detection. *Nat Methods* (2015) 12:623–30. doi: 10.1038/nmeth.3407
25. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: Somatic Mutation and Copy Number Alteration Discovery in Cancer by Exome Sequencing. *Genome Res* (2012) 22:568–76. doi: 10.1101/gr.129684.111
26. Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, et al. VarDict: A Novel and Versatile Variant Caller for Next-Generation Sequencing in Cancer Research. *Nucleic Acids Res* (2016) 44:e108. doi: 10.1093/nar/gkw227
27. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Kallberg M, et al. Strelka2: Fast and Accurate Calling of Germline and Somatic Variants. *Nat Methods* (2018) 15:591–4. doi: 10.1038/s41592-018-0051-x
28. Perez-Amado CJ, Tovar H, Gomez-Romero L, Beltran-Anaya FO, Bautista-Pina V, Dominguez-Reyes C, et al. Mitochondrial DNA Mutation Analysis in Breast Cancer: Shifting From Germline Heteroplasmy Toward Homoplasmy in Tumors. *Front Oncol* (2020) 10:572954. doi: 10.3389/fonc.2020.572954
29. Martinez-Jimenez F, Muinos F, Sentis I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, et al. A Compendium of Mutational Cancer Driver Genes. *Nat Rev Cancer* (2020) 20:555–72. doi: 10.1038/s41568-020-0290-x
30. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* (2018) 173:371–85.e18. doi: 10.1016/j.cell.2018.02.060
31. Ellrott K, Bailey MH, Saksena G, Covington KR, Kandoth C, Stewart C, et al. Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. *Cell Syst* (2018) 6:271–81.e7. doi: 10.1016/j.cels.2018.03.002
32. Karlsson J, Nilsson LM, Mitra S, Alsen S, Shelke GV, Sah VR, et al. Molecular Profiling of Driver Events in Metastatic Uveal Melanoma. *Nat Commun* (2020) 11:1894. doi: 10.1038/s41467-020-15606-0
33. Sequeira I, Rashid M, Tomas IM, Williams MJ, Graham TA, Adams DJ, et al. Genomic Landscape and Clonal Architecture of Mouse Oral Squamous Cell Carcinomas Dictate Tumour Ecology. *Nat Commun* (2020) 11:5671. doi: 10.1038/s41467-020-19401-9
34. Vettore AL, Ramnarayanan K, Poore G, Lim K, Ong CK, Huang KK, et al. Mutational Landscapes of Tongue Carcinoma Reveal Recurrent Mutations in Genes of Therapeutic and Prognostic Relevance. *Genome Med* (2015) 7:98. doi: 10.1186/s13073-015-0219-2
35. Kang KT, Kwon YW, Kim DK, Lee SI, Kim KH, Suh DS, et al. TRRAP Stimulates the Tumorigenic Potential of Ovarian Cancer Stem Cells. *BMB Rep* (2018) 51:514–9. doi: 10.5483/BMBRep.2018.51.10.042
36. De S, Cipriano R, Jackson MW, Stark GR. Overexpression of Kinesins Mediates Docetaxel Resistance in Breast Cancer Cells. *Cancer Res* (2009) 69:8035–42. doi: 10.1158/0008-5472.CAN-09-1224
37. He Y, Kan W, Li Y, Hao Y, Huang A, Gu H, et al. A Potent and Selective Small Molecule Inhibitor of Myoferlin Attenuates Colorectal Cancer Progression. *Clin Transl Med* (2021) 11:e289. doi: 10.1002/ctm2.289
38. Ning X, Shi Z, Liu X, Zhang A, Han L, Jiang K, et al. DNMT1 and EZH2 Mediated Methylation Silences the microRNA-200b/a/429 Gene and Promotes Tumor Progression. *Cancer Lett* (2015) 359:198–205. doi: 10.1016/j.canlet.2015.01.005
39. Julien SG, Dube N, Hardy S, Tremblay ML. Inside the Human Cancer Tyrosine Phosphatome. *Nat Rev Cancer* (2011) 11:35–49. doi: 10.1038/nrc2980

40. Mettlen M, Chen PH, Srinivasan S, Danuser G, Schmid SL. Regulation of Clathrin-Mediated Endocytosis. *Annu Rev Biochem* (2018) 87:871–96. doi: 10.1146/annurev-biochem-062917-012644
41. Zhi X, Lin L, Yang S, Bhuvaneshwar K, Wang H, Gusev Y, et al. betaII-Spectrin (SPTBN1) Suppresses Progression of Hepatocellular Carcinoma and Wnt Signaling by Regulation of Wnt Inhibitor Kallistatin. *Hepatology* (2015) 61:598–612. doi: 10.1002/hep.27558
42. Zhu C, Ji X, Zhang H, Zhou Q, Cao X, Tang M, et al. Deubiquitylase USP9X Suppresses Tumorigenesis by Stabilizing Large Tumor Suppressor Kinase 2 (LATS2) in the Hippo Pathway. *J Biol Chem* (2018) 293:1178–91. doi: 10.1074/jbc.RA117.000392
43. Yeddula N, Xia Y, Ke E, Beumer J, Verma IM. Screening for Tumor Suppressors: Loss of Ephrin Receptor A2 Cooperates With Oncogenic KRas in Promoting Lung Adenocarcinoma. *Proc Natl Acad Sci USA* (2015) 112: E6476–85. doi: 10.1073/pnas.1520110112
44. Tang J, Liu C, Xu B, Wang D, Ma Z, Chang X. ARHGEF10L Contributes to Liver Tumorigenesis Through RhoA-ROCK1 Signaling and the Epithelial-Mesenchymal Transition. *Exp Cell Res* (2019) 374:46–68. doi: 10.1016/j.yexcr.2018.11.007
45. Xiao W, Hong H, Kawakami Y, Kato Y, Wu D, Yasudo H, et al. Tumor Suppression by Phospholipase C-Beta3 *via* SHP-1-Mediated Dephosphorylation of Stat5. *Cancer Cell* (2009) 16:161–71. doi: 10.1016/j.ccr.2009.05.018
46. Wei W, Wang H, Ji S. Paradoxes of the EphB1 Receptor in Malignant Brain Tumors. *Cancer Cell Int* (2017) 17:21. doi: 10.1186/s12935-017-0384-z
47. Wu X, Bekker-Jensen IH, Christensen J, Rasmussen KD, Sidoli S, Qi Y, et al. Tumor Suppressor ASXL1 Is Essential for the Activation of INK4B Expression in Response to Oncogene Activity and Anti-Proliferative Signals. *Cell Res* (2015) 25:1205–18. doi: 10.1038/cr.2015.121
48. Liu SB, Zhou LB, Wang HF, Li G, Xie QP, Hu B. Loss of IGF2R Indicates a Poor Prognosis and Promotes Cell Proliferation and Tumorigenesis in Bladder Cancer *via* AKT Signaling Pathway. *Neoplasma* (2020) 67:129–36. doi: 10.4149/neo_2019_190206N108
49. Yamada K, Ono M, Perkins ND, Rocha S, Lamond AI. Identification and Functional Characterization of FMN2, a Regulator of the Cyclin-Dependent Kinase Inhibitor P21. *Mol Cell* (2013) 49:922–33. doi: 10.1016/j.molcel.2012.12.023
50. Zarringhalam K, Tay Y, Kulkarni P, Bester AC, Pandolfi PP, Kulkarni RV. Identification of Competing Endogenous RNAs of the Tumor Suppressor Gene PTEN: A Probabilistic Approach. *Sci Rep* (2017) 7:7755. doi: 10.1038/s41598-017-08209-1
51. Ma D, Pan Z, Chang Q, Zhang JJ, Liu X, Hua N, et al. KLF5-Mediated Eppk1 Expression Promotes Cell Proliferation in Cervical Cancer *via* the P38 Signaling Pathway. *BMC Cancer* (2021) 21:377. doi: 10.1186/s12885-021-08040-y
52. Gordon-Weeks A, Lim SY, Yuzhalin A, Lucotti S, Vermeer JAF, Jones K, et al. Tumour-Derived Laminin Alpha5 (LAMA5) Promotes Colorectal Liver Metastasis Growth, Branching Angiogenesis and Notch Pathway Inhibition. *Cancers (Basel)* (2019) 11(5):630. doi: 10.3390/cancers11050630
53. Guo Q, Grimmig T, Gonzalez G, Giobbie-Hurder A, Berg G, Carr N, et al. ATP-Binding Cassette Member B5 (ABCB5) Promotes Tumor Cell Invasiveness in Human Colorectal Cancer. *J Biol Chem* (2018) 293:11166–78. doi: 10.1074/jbc.RA118.003187
54. Peng L, He K, Cao Z, Bi L, Yu D, Wang Q, et al. CARD10 Promotes the Progression of Renal Cell Carcinoma by Regulating the NFkappaB Signaling Pathway. *Mol Med Rep* (2020) 21:329–37. doi: 10.3892/mmr.2019.10840
55. Huang Z, Wu H, Chuai S, Xu F, Yan F, Englund N, et al. NSD2 Is Recruited Through Its PHD Domain to Oncogenic Gene Loci to Drive Multiple Myeloma. *Cancer Res* (2013) 73:6277–88. doi: 10.1158/0008-5472.CAN-13-1000
56. Chigurupati S, Venkataraman R, Barrera D, Naganathan A, Madan M, Paul L, et al. Receptor Channel TRPC6 Is a Key Mediator of Notch-Driven Glioblastoma Growth and Invasiveness. *Cancer Res* (2010) 70:418–27. doi: 10.1158/0008-5472.CAN-09-2654
57. Tan X, Chen S, Wu J, Lin J, Pan C, Ying X, et al. PI3K/AKT-Mediated Upregulation of WDR5 Promotes Colorectal Cancer Metastasis by Directly Targeting ZNF407. *Cell Death Dis* (2017) 8:e2686. doi: 10.1038/cddis.2017.111
58. Li M, Ke J, Wang Q, Qian H, Yang L, Zhang X, et al. Upregulation of ROCK2 in Gastric Cancer Cell Promotes Tumor Cell Proliferation, Metastasis and Invasion. *Clin Exp Med* (2017) 17:519–29. doi: 10.1007/s10238-016-0444-z
59. Li W, Ouyang Z, Zhang Q, Wang L, Shen Y, Wu X, et al. SBF-1 Exerts Strong Anticervical Cancer Effect Through Inducing Endoplasmic Reticulum Stress-Associated Cell Death *via* Targeting Sarco/Endoplasmic Reticulum Ca(2+)-ATPase 2. *Cell Death Dis* (2014) 5:e1581. doi: 10.1038/cddis.2014.538
60. Loeb LA, Harris CC. Advances in Chemical Carcinogenesis: A Historical Review and Prospective. *Cancer Res* (2008) 68:6863–72. doi: 10.1158/0008-5472.CAN-08-2852
61. Stratton MR, Campbell PJ, Futreal PA. The Cancer Genome. *Nature* (2009) 458:719–24. doi: 10.1038/nature07943
62. Eichler EE, Clark RA, She X. An Assessment of the Sequence Gaps: Unfinished Business in a Finished Human Genome. *Nat Rev Genet* (2004) 5:345–54. doi: 10.1038/nrg1322
63. Li H. Toward Better Understanding of Artifacts in Variant Calling From High-Coverage Samples. *Bioinformatics* (2014) 30:2843–51. doi: 10.1093/bioinformatics/btu356
64. Miga KH, Eisenhart C, Kent WJ. Utilizing Mapping Targets of Sequences Underrepresented in the Reference Assembly to Reduce False Positive Alignments. *Nucleic Acids Res* (2015) 43:e133. doi: 10.1093/nar/gkv671
65. Ma X, Shao Y, Tian L, Flasch DA, Mulder HL, Edmonson MN, et al. Analysis of Error Profiles in Deep Next-Generation Sequencing Data. *Genome Biol* (2019) 20:50. doi: 10.1186/s13059-019-1659-6
66. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn* (2017) 19:341–65. doi: 10.1016/j.jmoldx.2017.01.011
67. Schumacher TN, Kesmir C, van Buuren MM. Biomarkers in Cancer Immunotherapy. *Cancer Cell* (2015) 27:12–4. doi: 10.1016/j.ccell.2014.12.004
68. Jiang AM, Ren MD, Liu N, Gao H, Wang JJ, Zheng XQ, et al. Tumor Mutation Burden, Immune Cell Infiltration, and Construction of Immune-Related Genes Prognostic Model in Head and Neck Cancer. *Int J Med Sci* (2021) 18:226–38. doi: 10.7150/ijms.51064
69. Zhang L, Li B, Peng Y, Wu F, Li Q, Lin Z, et al. The Prognostic Value of TMB and the Relationship Between TMB and Immune Infiltration in Head and Neck Squamous Cell Carcinoma: A Gene Expression-Based Study. *Oral Oncol* (2020) 110:104943. doi: 10.1016/j.oraloncology.2020.104943
70. Eder T, Hess AK, Konschak R, Stromberger C, Johrens K, Fleischer V, et al. Interference of Tumour Mutational Burden With Outcome of Patients With Head and Neck Cancer Treated With Definitive Chemoradiation: A Multicentre Retrospective Study of the German Cancer Consortium Radiation Oncology Group. *Eur J Cancer* (2019) 116:67–76. doi: 10.1016/j.ejca.2019.04.015
71. Carbone DP, Reck M, Paz-Ares L, Creelan B, Horn L, Steins M, et al. First-Line Nivolumab in Stage IV or Recurrent Non-Small-Cell Lung Cancer. *N Engl J Med* (2017) 376:2415–26. doi: 10.1056/NEJMoa1613493
72. Cui J, Wang D, Nie D, Liu W, Sun M, Pei F, et al. Difference in Tumor Mutation Burden Between Squamous Cell Carcinoma in the Oral Cavity and Larynx. *Oral Oncol* (2021) 114:105142. doi: 10.1016/j.oraloncology.2020.105142
73. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational Landscape of Metastatic Cancer Revealed From Prospective Clinical Sequencing of 10,000 Patients. *Nat Med* (2017) 23:703–13. doi: 10.1038/nm.4333
74. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor Mutational Load Predicts Survival After Immunotherapy Across Multiple Cancer Types. *Nat Genet* (2019) 51:202–6. doi: 10.1038/s41588-018-0312-8
75. Farah CS, Jessri M, Bennett NC, Dalley AJ, Shearston KD, Fox SA. Exome Sequencing of Oral Leukoplakia and Oral Squamous Cell Carcinoma Implicates DNA Damage Repair Gene Defects in Malignant Transformation. *Oral Oncol* (2019) 96:42–50. doi: 10.1016/j.oraloncology.2019.07.005
76. Al-Hebshi NN, Li S, Nasher AT, El-Setouhy M, Alsanosi R, Blancato J, et al. Exome Sequencing of Oral Squamous Cell Carcinoma in Users of Arabian Snuff Reveals Novel Candidates for Driver Genes. *Int J Cancer* (2016) 139:363–72. doi: 10.1002/ijc.30068
77. Zhang H, Song Y, Du Z, Li X, Zhang J, Chen S, et al. Exome Sequencing Identifies New Somatic Alterations and Mutation Patterns of Tongue Squamous Cell Carcinoma in a Chinese Population. *J Pathol* (2020) 251:353–64. doi: 10.1002/path.5467
78. Su SC, Chang LC, Lin CW, Chen MK, Yu CP, Chung WH, et al. Mutational Signatures and Mutagenic Impacts Associated With Betel Quid Chewing in

Oral Squamous Cell Carcinoma. *Hum Genet* (2019) 138:1379–89. doi: 10.1007/s00439-019-02083-9

79. Nisa L, Barras D, Medova M, Aebersold DM, Medo M, Poliakova M, et al. Comprehensive Genomic Profiling of Patient-Matched Head and Neck Cancer Cells: A Preclinical Pipeline for Metastatic and Recurrent Disease. *Mol Cancer Res* (2018) 16:1912–26. doi: 10.1158/1541-7786.MCR-18-0056
80. Li H, Ngan HL, Liu Y, Chan HHY, Poon PHY, Yeung CK, et al. Comprehensive Exome Analysis of Immunocompetent Metastatic Head and Neck Cancer Models Reveals Patient Relevant Landscapes. *Cancers (Basel)* (2020) 12. doi: 10.3390/cancers12102935
81. Klionsky DJ, Abdel-Aziz AK, Abdelfatah S, Abdellatif M, Abdoli A, Abel S, et al. Guidelines for the Use and Interpretation of Assays for Monitoring Autophagy (4th Edition)(1). *Autophagy* (2021) 17:1–382. doi: 10.1080/15548627.2020.1797280
82. Vieira AV, Lamaze C, Schmid SL. Control of EGF Receptor Signaling by Clathrin-Mediated Endocytosis. *Science* (1996) 274:2086–9. doi: 10.1126/science.274.5295.2086

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership