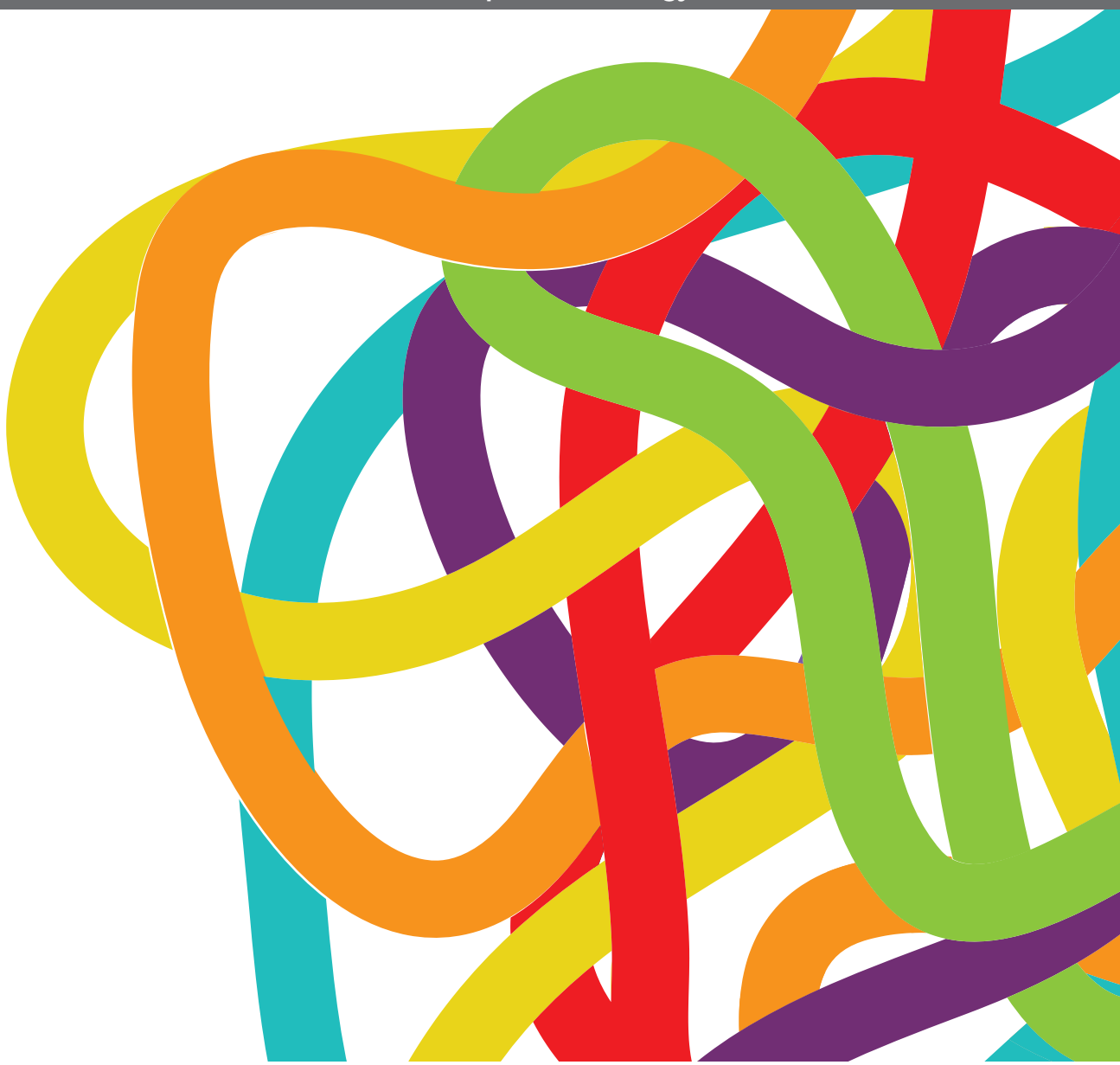
THE EVOLVING ROLE OF NEXT GENERATION SEQUENCING IN CANCER CARE

EDITED BY: Feng He, Bing Xu and Qinghua Xu
PUBLISHED IN: Frontiers in Oncology and
Frontiers in Cell and Developmental Biology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-824-0
DOI 10.3389/978-2-88976-824-0

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

THE EVOLVING ROLE OF NEXT GENERATION SEQUENCING IN CANCER CARE

Topic Editors:
Feng He, Shanghai University of Traditional Chinese Medicine, China
Bing Xu, Xiamen University, China
Qinghua Xu, The Canhelp Genomics Research Center, China

Citation: He, F., Xu, B., Xu, Q., eds. (2022). The Evolving Role Of Next Generation Sequencing In Cancer Care. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-824-0

Table of Contents

frontiers in Oncology

ORIGINAL RESEARCH
published: 23 March 2021
doi: 10.3389/fonc.2021.634732

Using lncRNA Sequencing to Reveal a Putative lncRNA-mRNA Correlation Network and the Potential Role of PCBP1-AS1 in the Pathogenesis of Cervical Cancer

Linhan Li[1†], Qisong Peng[2†], Min Gong[1], Ling Ling[1], Yingxue Xu[1] and Qiaoling Liu[1]*

[1] Department of Gynaecology and Obstetrics, Affiliated Jiangning Hospital of Nanjing Medical University, Nanjing, China,
[2] Department of Clinical Laboratory, Affiliated Jiangning Hospital of Nanjing Medical University, Nanjing, China

OPEN ACCESS

Edited by:
César López-Camarillo, Universidad Autónoma de la Ciudad de México, Mexico

Reviewed by:
Napoleon Navarro-Tito, Autonomous University of Guerrero, Mexico
Desi Shang, Harbin Medical University, China

***Correspondence:**
Qiaoling Liu
lql711030@163.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 28 November 2020
Accepted: 24 February 2021
Published: 23 March 2021

Citation:
Li L, Peng Q, Gong M, Ling L, Xu Y and Liu Q (2021) Using lncRNA Sequencing to Reveal a Putative lncRNA-mRNA Correlation Network and the Potential Role of PCBP1-AS1 in the Pathogenesis of Cervical Cancer. Front. Oncol. 11:634732. doi: 10.3389/fonc.2021.634732

Background/Aims: Long non-coding RNAs (lncRNAs) play important roles in many diseases and participate in posttranscriptional regulatory networks in tumors. However, the functions of major lncRNAs in cervical cancer are unclear. Therefore, the aim of this study was to construct a lncRNA-mRNA coexpression functional network and analyze lncRNAs that might contribute to the pathogenesis of cervical cancer.

Methods: Differentially expressed lncRNAs (DElncRNAs) and mRNAs (DEmRNAs) between three pairs of cervical cancer tissues and adjacent mucosa were identified by lncRNA microarray analysis. LncRNA-mRNA correlation analysis and functional enrichment were performed on the DEGs. From the correlation network, PCBP1-AS1 was selected as a candidate for further analysis. PCBP1-AS1 expression was examined by qPCR, and Kaplan–Meier survival, clinicopathology, GSEA, and immune infiltration analysis of PCBP1-AS1 were performed. The immune responses of PCBP1-AS1 expression in cervical cancer were analyzed using TIMER and western blot. PCBP1-AS1 was knocked down and overexpressed to evaluate its role in cell proliferation, migration, and invasion.

Results: A total of 130 lncRNAs were significantly differentially expressed in cervical cancer patient samples compared with control samples. Differentially expressed mRNAs in the lncRNA-mRNA interaction network were involved in the EMT process. Combined with the Kaplan–Meier survival analyses, the coexpression network revealed that PCBP1-AS1 was significantly associated with OS and clinicopathological parameters in cervical cancer patients. Moreover, PCBP1-AS1 expression was not only significantly increased in cervical cancer specimens but also associated with tumor stage, TNM, and invasion. GSEA revealed that PCBP1-AS1 is closely correlated with cell biological function *via* the p53 and notch signaling pathways. TIMER analysis revealed that the numbers of NK cells and M2 macrophages decreased when PCBP1-AS1 expression was high, which was consistent with the western blot results in clinical samples. Furthermore, *in vitro*

experiments showed that high expression of PCBP1-AS1 promoted cell proliferation, migration, and invasion.

Conclusions: Transcriptomic and lncRNA-mRNA correlation analyses revealed that PCBP1-AS1 plays a key role as an independent prognostic factor in patients with cervical cancer. The identification of PCBP1-AS1 as a new biomarker for cervical cancer could help explain how changes in the immune environment promote cervical cancer development.

Keywords: lncRNA-mRNA correlation network, long non-coding RNA, cervical cancer, lncRNA sequencing, proliferation

INTRODUCTION

Cervical cancer is one of the most common malignancies in female patients, and it has the highest mortality of all female reproductive system malignancies (1, 2). Moreover, its prevalence rate is rising among young women (3). Most patients suffering from cervical cancer are diagnosed at advanced stages, accompanied by invasion and distant metastasis (4, 5). At present, surgical resection and chemotherapy are the first-tier options of cervical cancer treatments, but tumor metastasis and recurrence still lead to poor prognosis (5–7). Although some progress has been made in research on the mechanism of cervical cancer, clinical applications are still limited, resulting in persistently high mortality in cervical cancer (8, 9). Therefore, the discovery of new mechanisms associated with cervical cancer for the identification of useful biomarkers as well as new specific therapeutic targets in cervical cancer is urgently needed.

In recent years, high-throughput transcriptome sequencing has become very common, revealing that up to 70% of the human genome is transcribed. However, the coding-protein transcripts are less than 2%, and most transcripts belong to non-protein-coding RNAs (ncRNAs) (10), including microRNAs (miRNAs), long non-coding RNAs (lncRNAs), and circular RNAs. Accumulating evidence has demonstrated that ncRNAs play crucial roles in the occurrence and progression of tumors (11–13). In the past decade, lncRNAs, defined as transcripts with a length of more than 200 nt, have been found to play key roles in multiple types of human tumorigenesis, metastasis, and chemotherapy resistance (14–16). Nevertheless, most of the functional lncRNAs in cervical cancer have yet to be identified. The mechanism of lncRNA function is associated with its target mRNAs. Therefore, lncRNA induced target mRNA transcription disorders was an effective strategy to identify key functional lncRNAs for cancer. For example, to search for candidate prostate cancer-related lncRNAs, lncRNA-mRNA bipartite networks, and lncRNA-mRNA coexpression networks have been constructed (17, 18). Although the mechanism of lncRNAs is not fully understood, they have already been considered potential biomarkers and therapeutic targets for many tumors (19, 20).

Here, we performed lncRNA-seq to investigate the expression levels of lncRNAs and mRNAs in six cervical cancer samples (three paired cervical cancer and adjacent mucosa) and constructed a lncRNA-mRNA coexpression network to identify the role of the candidate lncRNAs in the expression, prognosis, clinical pathology, immune infiltration, proliferation, migration, and invasion of cervical cancer and HeLa cells.

MATERIALS AND METHODS

Sample Collection and Preparation

A total of three pairs of cervical cancer and adjacent tissues were collected from three cervical cancer patients who underwent surgical operation from July 2019 to August 2012 in Department of Gynecology and Obstetrics, Affiliated Jiangning Hospital of Nanjing Medical University. After sequencing, we collected another 20 pairs of samples for data verification (**Supplementary Table 1**). Among them, 15 pairs (15 cervical cancer tissues and 15 controls) were collected from June 2020 to December 2020 for qPCR assay. Another five pairs of samples (five cervical cancer tissues and five controls) were collected between January 2021 and February 2021 to indirectly measure the number of immune cells in cervical cancer samples.

Specimens were frozen in liquid nitrogen immediately after operation and stored at −80°C until extraction. All samples were confirmed by histopathological examination. This study was approved by the ethics committee of the hospital. Informed consent to collection and use of the biological samples was obtained from each patient.

LncRNA Microarray

Total RNA was isolated using a RNeasy mini kit (Qiagen, Germany) and analyzed by 1% agarose gel electrophoresis (Bio-Rad, USA) to ensure that no degradation occurred. The RNA libraries were constructed using the TruSeq RNA Sample Preparation Kit (Illumina, USA). After purification, libraries were quantified using a Qubit 8000 (Life Technologies, USA) and validated with an Agilent 2100 (Agilent Technologies, USA) to confirm the insert size. Then, clusters were sequenced on an Illumina HiSeq 2500 instrument (Illumina, USA). Library construction and sequencing were performed at Shanghai Yuanshen Biomedical Technology Co., Ltd.

Data Analysis

Differentially expressed lncRNAs and mRNAs were identified through fold-change filtering ($|\log2FC|>1$ and $p<0.05$) using the "edgeR" package in R. The differentially expressed RNA profiles were normalized by log2 transformation.

LncRNA-mRNA Correlation Network

The Pearson correlation coefficient was calculated, and the R value (cutoff >0.95) was used for each pair of lncRNA-mRNA interactions. The lncRNA-mRNA correlation network was constructed by Cytoscape software.

GO and KEGG Functional Enrichment Analysis

Functional enrichment analysis of lncRNA-target mRNAs was performed using Metascape (https://metascape.org/). All statistically enriched terms (Gene Ontology and Kyoto Encyclopedia of Genes and Genomes) were identified based on accumulative hypergeometric *p* values.

Kaplan–Meier Survival Analysis of lncRNAs

To investigate the predictive value of the expression levels of lncRNAs and mRNAs for the survival of cervical cancer patients, Kaplan–Meier survival analysis was performed using GEPIA (http://gepia2.cancer-pku.cn/). The statistical significance was set at $p < 0.05$. Then, the data obtained from the analysis were verified by StarBase (http://starbase.sysu.edu.cn/).

Cox Proportional Regression Model Based on Differentially Expressed RNAs

To analyze the independent effects of individual miRNAs on the overall survival of patients with colon cancer, we performed univariate and multivariate Cox proportional regression analysis with an online tool (SangerBox tools, http://sangerbox.com/Tool). We constructed a Cox proportional hazards regression model and calculated the risk value of each patient through the formula (risk score = b × exp (RNA1) + b × exp (RNA2) +… + b × exp (RNAn), where b represents the multivariate Cox regression coefficient and exp () represents the expression level of prognostic RNAs. Next, we calculated the survival rates of the high-risk and low-risk groups and plotted the 1-year, 3-year, and 5-year survival receiver operating characteristic (ROC) curves to test the feasibility of the prediction ability of the model.

RNA Extraction and Quantitative PCR

Total RNA was extracted from cervical cancer samples using an RNeasy mini kit (Qiagen, Germany) and analyzed by 1% agarose gel electrophoresis (Bio-Rad, USA) to ensure that no degradation had occurred. A Qubit 8000 (Life Technologies, USA) was used to measure the RNA concentration. Then, the RNA was reverse transcribed into cDNA. qRT-PCR was performed using SYBR Premix Ex Taq (Takara, China) on an ABI7500 system (Applied Biosystems, CA). The following cycling parameters were used: initial denaturation at 95°C for 30 s, followed by 35 cycles of 95°C for 5 s, 58°C for 30 s and 95°C for 60 s, and 60°C for 30 s. The primer sequences for PCR were as follows: PCBP1-AS1, forward: 5′-CCAACCTGATACATTGCCT-3′ and reverse 5′-TGGAAGAAATTCCCTGCTG-3′, GAPDH: forward 5′-CTCCTCCACCTTTGACGCTG-3′ and reverse 5′-TCCT CTTGTGCTCTTGCTGG-3′. Primers were synthesized by Sangon Biotech (China). GAPDH was used as a control. The mean value of triplicate experiments was used to calculate relative lncRNA expression using the formula $\Delta Ct = Ct^{mean}$ lncRNAs – Ct^{mean} GAPDH. Expression fold changes were calculated using the $2^{-\Delta\Delta Ct}$ method.

Immune Infiltrates Analysis

The TIMER (https://cistrome.shinyapps.io/timer/) correlation module was used to evaluate potential relationships between PCBP1-AS1 expression and immune infiltrates.

Gene Set Enrichment Analysis

GSEA was performed using normalized RNA-Seq data obtained from TCGA-cervical cancer. The number of permutations was set to 100. Using GSEA, we further analyzed GO terms and KEGG pathways to investigate possible biological functions of PCBP1-AS1 (p-value <0.05).

Cell Culture

HeLa cells were obtained from the American Type Culture Collection (ATCC). The cell lines were cultured as suggested by ATCC. The cells were cultured in Dulbecco's modified Eagle's medium (Invitrogen, USA) supplemented with 10% fetal bovine serum (Invitrogen, USA), 100 U/ml penicillin (Sigma, USA), and 100 μg/ml streptomycin (Sigma, USA) under a humidified atmosphere of 5% CO_2 at 37°C.

Cell Transfection

PCBP1-AS1 small interfering RNA (si-PCBP1-AS1) and the corresponding control (si-NC) were purchased from RiboBio (Guangzhou, China). The PCBP1-AS1 overexpression plasmid (pCDH-GFP-PCBP1-AS1) and corresponding control plasmid (NC) were also purchased from RiboBio (Guangzhou, China). All oligomers and plasmids were transfected into HeLa cells using Lipofectamine 3000 reagents (Invitrogen, USA) based on the manufacturer's protocol. Briefly, when HeLa cell densities were approximately 60% in 12-well plates (Corning, USA), 50 nM siRNA oligos or 2 μg overexpression plasmids were introduced into cells using Lipofectamine 3000 reagents (Invitrogen, USA). Untreated cells were set as blank groups and transfected with empty vectors, and NC-siRNA was used as a negative control. At 24 h post transfection, the efficiency of knockdown and overexpression was determined by qRT-PCR and fluorescence microscopy. Subsequent experiments were performed at 48 h after transfection.

Proliferation Assays

The proliferation of HeLa cells was measured by cell proliferation using Cell Counting Kit-8 (Sigma, USA) in 96-well plates. Then, 3,000 cells/well were incubated for 12, 24, and 48 h. All cells were then incubated with CCK-8 reagent (10 μl per well) for 3 h, and a microplate reader (Thermo, USA) was utilized to detect the absorbance of each well at 450 nm. Each experiment was carried out three times.

Wound-Healing Assay

HeLa cells were seeded in plates (96 wells, Corning, USA) at 5×10^5 cells/well with culture medium at 37°C with 5% CO_2. Then, the confluent cell monolayer was scratched with a sterile 200 μl pipette tip, and Opti-MEM™-reduced serum medium (Gibco,

USA) was added. Microscope photos were taken after 0, 12, 24, and 48 h to record the scratched areas. ImageJ software was used to evaluate the percentage of closure.

Transwell Assay

At 48 h after transfection, HeLa cells were collected to prepare a single-cell suspension. The HeLa cell suspension (3×10^3 cells/well) was added to the Transwell upper chamber (Corning, USA), and DMEM (20% FBS) was added to 24-well plates in the lower chamber. The upper chamber was coated with Matrigel. After 24 h, 4% paraformaldehyde (Sigma, USA) was applied to fix the cells, and the cells were stained with 1% crystal violet (Sigma, USA). Cells were observed and counted under an optical microscope (Olympus, Japan).

Western Blot Analysis

Total protein was extracted using RIPA lysis buffer (Beyotime, China), and separated using 10% SDS-PAGE (Beyotime, China). Next, proteins were transferred onto PVDF membranes (Beyotime, China). Then, the target protein membrane was blocked with 5% nonfat milk for 24 h. Subsequently, the membranes were incubated with specific primary antibodies against CD4 (Beyotime, China, 1:1,000), CD19 (Beyotime, China, 1:1,000), CD56 (Beyotime, China, 1:1,000), and GAPDH (Beyotime, China, 1:1,000) for 24 h at 4°C. GAPDH was used as a control. Afterwards, the membranes were incubated with secondary antibodies (Beyotime, China, 1:1,000) for 4 h at room temperature. TMB color liquid (Beyotime, China) was used to detect protein bands. ImageJ software was used to analyze the gray value of the bands. Protein levels were calculated using the ratio of target protein/GAPDH.

Statistical Analysis

GraphPad Prism 8.0 software (California, USA) was utilized to perform statistical analysis. The discrepancies between two groups were compared by t-test. The differences were deemed statistically significant at $p < 0.05$.

RESULT

Differential Expression Patterns of Genes Between Cervical Cancer Tissues and Adjacent Tissues

To understand the lncRNAs and mRNAs involved in cervical cancer pathogenesis, we performed lncRNA microarray detection in three pairs of cervical cancer tissues and matched adjacent tissues. On average, 82.58 million reads were obtained for each sample (**Supplementary Table 2**). Among the 10,675 detected expressed lncRNAs, we identified 130 differentially expressed lncRNAs between cervical cancer tissues and adjacent tissues ($|\log2FC|>1$ and $p < 0.05$), of which 48 were upregulated and 82 were downregulated in the cancer tissue, as shown in **Figure 1A**; the red points represent statistically significant upregulated differentially expressed lncRNAs, and the blue points represent downregulated differentially expressed lncRNAs. With this same criterion, we identified 656 significantly differentially expressed mRNAs, of which 293 were upregulated and 363 were downregulated (**Figure 1A**). The top 10 lncRNAs (up- and down) and mRNAs (up- and down) are shown in **Tables 1** and **2**, respectively. Among the annotated differentially expressed lncRNAs, antisense lncRNAs accounted for 37.50% of upregulated lncRNAs, whereas the majority of downregulated lncRNAs were lincRNAs and antisense lncRNAs (accounting for 40.24% each) (**Figure 1D**). The distribution of differentially expressed lncRNAs across chromosomes was also analyzed. Among the downregulated differentially expressed lncRNAs, chromosome 1 had the most differentially expressed lncRNAs (n = 9), followed by chromosome 5 (n = 8) (**Figure 1C**). In the same analysis, chromosome 2 had the most DElncRNAs (n = 7) among the upregulated differentially expressed lncRNAs (**Figure 1C**). The heat map results in **Figure 1B** show that lncRNA and mRNA expression was distinct between cervical cancer tissues and adjacent tissues. In summary, the results from the lncRNA microarray analysis indicated that aberrantly expressed genes, including mRNAs and lncRNAs, may play important roles in the development and progression of cervical cancer.

lncRNA-RNA Interaction Network and GO Analysis

lncRNAs can regulate the transcription, translation, and splicing of downstream target mRNAs. To understand the correlation between the expression of differentially expressed lncRNAs and differentially expressed mRNAs, a lncRNA-RNA interaction network was constructed. The lncRNA-mRNA coexpression pairs in the network were selected with a threshold of correlation ≥0.95, resulting in a network consisting of 514 nodes and 7,102 significant coexpression relationships, including 127 differentially expressed lncRNAs and 387 differentially expressed mRNAs (**Figure 2A**), suggesting that these differentially expressed lncRNAs might regulate downstream target mRNAs mainly through induction mechanisms. As shown in **Figure 2A**, the coexpression regulatory network was divided into two parts (cis and trans), which are the two regulatory mechanisms by which lncRNAs regulate downstream genes. Then, using the jActive module, we identified a highly active subnetwork module (ActivePath Score = 7.20; **Figure 2B**) from the network, including 66 nodes and 192 edges with 19 lncRNAs and 47 mRNAs. GO and KEGG enrichment analyses were performed to analyze the functions of the differentially expressed mRNAs in each subnetwork module of the network. The results showed that these differentially expressed mRNAs were significantly enriched in epidermal development, regulation of hormone levels, cell cycle, epidermal cell differentiation, phosphorylation, and cell resistance. The main pathways were the HNF3A pathway and regulation of the intracellular estrogen receptor signaling pathway (**Figure 3**), which indicated that these differentially expressed mRNAs may be related to EMT. EMT causes dissociated epithelial cells to acquire migration and invasive capacities and confers cancer cells with the ability to migrate to distant tissues.

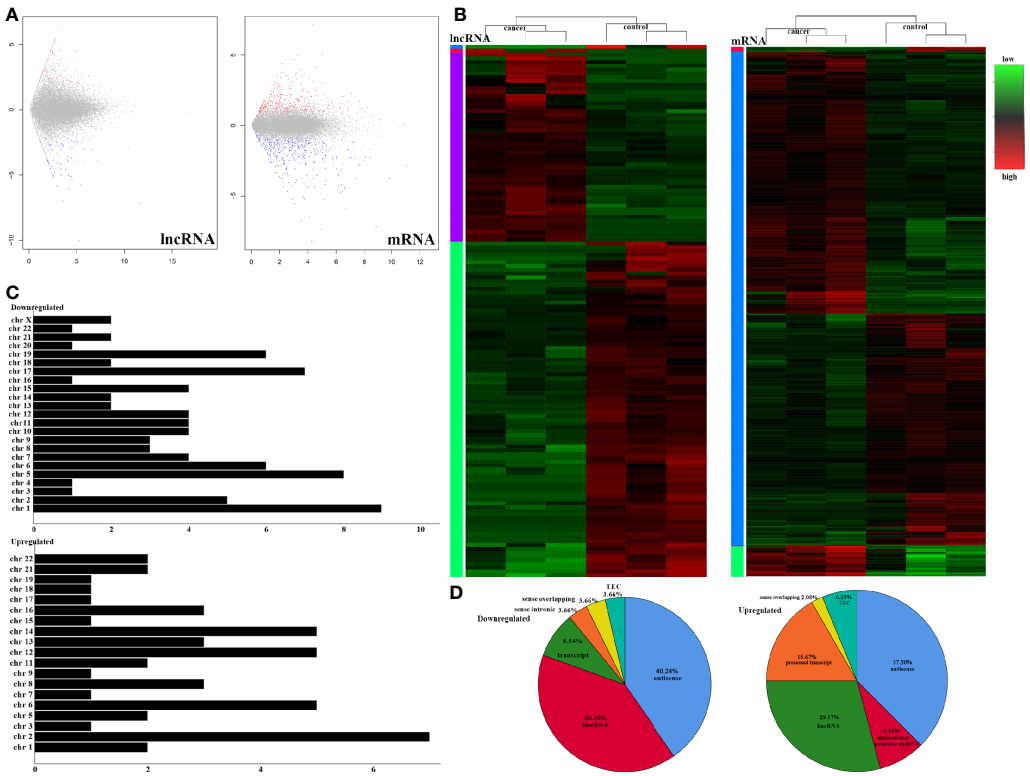


FIGURE 1 | LncRNA sequencing of CESC. **(A)** Volcano plots of differential expression profiles of lncRNAs and mRNAs. The red dots represent upregulated genes. The blue dots represent downregulated genes. The gray dots represent genes that do not differ significantly. **(B)** Heat map of differential expression profiles of lncRNAs and mRNAs. **(C)** Chromosomal distribution of all differentially expressed lncRNAs. **(D)** Fraction distribution of all category-annotated DEG long non-coding RNAs (lncRNAs).

TABLE 1 | Top10 (up- and downregulated) of DElncRNAs in normal tissues and CESC tissues.

ID	Symbol	lncRNA type	logFC	padj	chr	P Value
Upregulation						
ENSG00000253339	AC111149.2	lincRNA	12.68869	0.040587	8	0.000396
ENSG00000257588	AC025154.2	antisense	12.06594	0.000559	12	1.27E-06
ENSG00000235954	TTC28-AS1	processed_transcript	11.94252	0.046726	22	0.00051
ENSG00000236778	INTS6-AS1	antisense	11.15821	0.000319	13	5.51E-07
ENSG00000258592	AL391152.1	lincRNA	11.14377	0.034594	14	0.000323
ENSG00000248092	NNT-AS1	antisense	11.08669	0.001126	5	3.09E-06
ENSG00000232940	HCG25	antisense	10.29039	0.01226	6	7.89E-05
ENSG00000179818	PCBP1-AS1	processed_transcript	10.18153	0.003325	2	1.23E-05
ENSG00000231074	HCG18	antisense	10.17907	2.23E-05	6	2.62E-08
ENSG00000232306	AC012485.2	lincRNA	10.12062	0.004252	2	1.88E-05
Downregulation						
ENSG00000231062	AC103563.2	antisense	−13.6321	1.03E-07	2	2.42E-11
ENSG00000251562	MALAT1	lincRNA	−13.0168	2.33E-13	11	3.65E-17
ENSG00000285756	BX890604.2	lincRNA	−12.4793	0.000115	X	1.62E-07
ENSG00000203688	LINC02487	lincRNA	−12.4505	1.02E-05	6	9.64E-09
ENSG00000228789	HCG22	lincRNA	−11.7357	3.20E-05	6	4.26E-08
ENSG00000215458	AATBC	antisense	−11.6164	0.043124	21	0.000454
ENSG00000250167	AC034206.1	antisense	−11.3856	0.000559	5	1.23E-06
ENSG00000237499	AL357060.1	antisense	−11.3592	0.003325	6	1.25E-05
ENSG00000266729	DSG1-AS1	antisense	−11.1001	0.003806	18	1.55E-05
ENSG00000279717	AC005336.3	TEC	−10.8012	8.72E-06	19	7.29E-09

TABLE 2 | Top 10 (up- and downregulated) of DEmRNAs in normal tissues and CESC tissues.

ID	Symbol	Gene type	logFC	padj	Chr	PValue
Upregulation						
ENSG00000124208	TMEM189-UBE2V1	pc	10.95920703	0.028272279	20	0.000949057
ENSG00000162896	PIGR	pc	8.320731276	0.002015361	1	2.45E-05
ENSG00000157765	SLC34A2	pc	7.449877017	2.22E-10	4	1.32E-13
ENSG00000131152	AC010531.1	pc	7.366463382	0.000672218	16	5.55E-06
ENSG00000187908	DMBT1	pc	7.339636292	4.89E-05	10	2.34E-07
ENSG00000169064	ZBBX	pc	7.131696411	0.0001358	3	8.59E-07
ENSG00000083782	EPYC	pc	7.062597097	0.041539105	12	0.001665568
ENSG00000173702	MUC13	pc	6.797754334	0.001353022	3	1.42E-05
ENSG00000117983	MUC5B	pc	6.19862045	0.014474512	11	0.000361908
ENSG00000047457	CP	pc	6.149951455	3.30E-24	3	7.54E-28
Downregulation						
ENSG00000124766	SOX4	pc	1.065906521	0.005626318	6	9.64E-05
ENSG00000205593	DENND6B	pc	1.062334084	0.026703654	22	0.000882663
ENSG00000163902	RPN1	pc	1.061712008	0.042212725	3	0.001717434
ENSG00000075420	FNDC3B	pc	1.044223744	0.040616111	3	0.001619221
ENSG00000100629	CEP128	pc	1.039944887	0.045860025	14	0.00195825
ENSG00000166762	CATSPER2	pc	1.033277218	0.046291255	15	0.002005818
ENSG00000147400	CETN2	pc	1.01823592	0.025777184	X	0.000842572
ENSG00000162065	TBC1D24	pc	1.013031342	0.000746529	16	6.39E-06
ENSG00000125148	MT2A	pc	1.012283062	7.62E-05	16	3.93E-07
ENSG00000118707	TGIF2	pc	1.008808568	0.037527843	20	0.001411017

PCBP1-AS1 Is Associated With Poor Prognosis and Clinical Parameters of Cervical Cancer Patients

To identify the differentially expressed lncRNAs and mRNAs with potential prognostic value, the expression levels of 19 differentially expressed lncRNAs and 47 differentially expressed mRNAs in the network of the subnetwork module were analyzed using a univariate Cox proportional hazards regression model. Only one lncRNA (PCBP1-AS1) and four mRNAs (FAM222A, FHAD1, WDR62, and SBK1) were identified as prognostic factors ($p < 0.05$; **Figure 4B**). Kaplan–Meier curve analysis showed that PCBP1-AS1 was negatively correlated with OS ($p < 0.05$), and all mRNAs were positively correlated with OS ($p < 0.05$) (**Figure 4A**). Meanwhile, we evaluated the relationship among PCBP1-AS1 and four mRNA expression levels and various clinicopathological parameters of cervical cancer patients. Expression data and clinical characteristics were obtained from TCGA-cervical cancer database. The results showed that the expression of PCBP1-AS1, FAM222A, FHAD1, WDR62, and SBK1 was significantly correlated with tumor clinical stage, pathologic TNM, and lymphatic invasion ($p < 0.05$) (**Figure 4C**). From the multivariate Cox regression analysis, PCBP1-AS1 ($p = 0.046$; HR = 0.407, 95% CI, 0.156–1.06) and SBK1 ($p = 0.047$; HR = 0.804, 95% CI, 0.641–1.007) were independent prognostic factors (**Table 3**, **Supplementary Figure 1**).

Expression Validation of PCBP1-AS1 in Cervical Cancer Tissues

This project focuses on the regulation of lncRNAs; thus, PCBP1-AS1 was selected as a candidate for further determination of its role in cervical cancer pathogenesis. LncLocator prediction results revealed that PCBP1-AS1 was localized to the cytosol (**Figure 5A**). qRT-PCR was performed to detect PCBP1-AS1 expression levels in 15 cervical cancer tissues and 15 controls. The clinical characteristics of the cervical cancer samples are summarized in **Supplementary Table 1**. As shown in **Figure 5B**, PCBP1-AS1 had significantly higher expression levels in cervical cancer tissues than in their normal counterparts ($p < 0.001$). This result was consistent with the microarray analysis and TCGA data (**Figure 5C**). In addition, the expression of PCBP1-AS1 was positively correlated with FAM222A ($p = 0.035$), FHAD1 ($p = 0.027$), and SBK1 ($p < 0.001$) (**Figure 5D**). Moreover, univariate analysis revealed that PCBP1-AS1 expression, tumor stage, pathologic T stage, and lymph vascular invasion were significantly correlated with the OS of cervical cancer patients (**Table 4-a**). Our multivariate analysis revealed that PCBP1-AS1 expression might be an independent factor for the prognosis of cervical cancer (**Table 4-b**, **Figure 5E**). Meanwhile, the ROC curve AUC of PCBP1-AS1 expression for predicting survival was 0.603 (**Figure 5F**), which indicated that PCBP1-AS1 possessed the potential prognostic ability of cervical cancer. Furthermore, as shown in **Figure 5G**, we uncovered a correlation between PCBP1-AS1 expression and clinicopathologic characteristics. Increased PCBP1-AS1 expression levels in cervical cancer were significantly correlated with tumor stage, pathologic TNM ($p < 0.05$), and lymph invasion ($p = 0.0402$). These results indicated that cervical cancer patients with high levels of PCBP1-AS1 expression are more likely to promote the initiation and growth of cervical cancer than patients with low levels of PCBP1-AS1 expression due to the effect of tumor stage, pathologic TNM, and lymph invasion.

Relationship Between PCBP1-AS1 Expression and Tumor-Infiltrating Immune Cells

To understand the influence of PCBP1-AS1 in tumor-infiltrating lymphocytes, we analyzed the possible correlations between

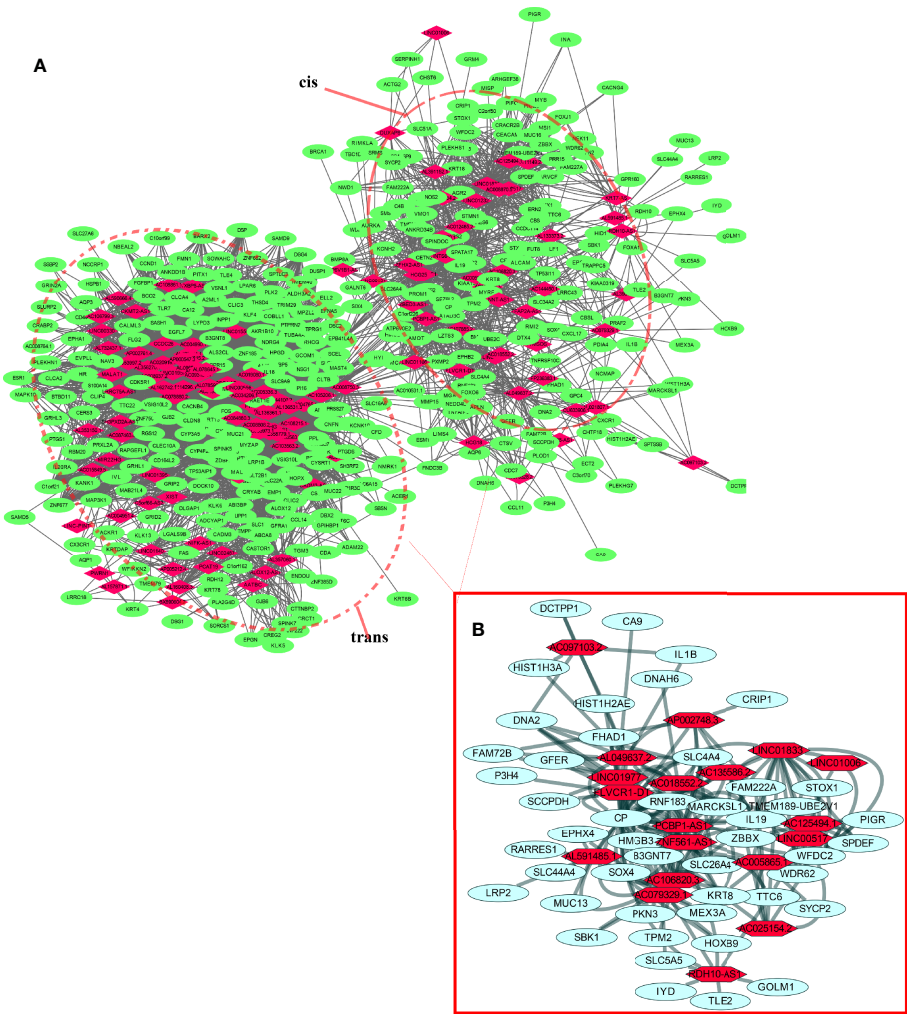


FIGURE 2 | LncRNA-mRNA interaction network. **(A)** Correlation analysis was carried out with differentially expressed lncRNAs and mRNAs, and the cutoff value was set with a threshold of correlation >0.95. Red nodes represent lncRNAs, green nodes represent mRNAs, and red dotted boxes indicate trans and cis groups. **(B)** The red solid line box indicates the highly active subnetwork module identified by the jActive module.

PCBP1-AS1 expression and levels of immune infiltration in cervical cancer. As shown in **Figure 6**, PCBP1-AS1 expression showed a positive correlation with the levels of $CD8^+$ T cells ($p < 0.05$), $CD4^+$ T cells ($p < 0.05$), B cells ($p < 0.05$), cancer-associated fibroblasts ($p < 0.05$), myeloid dendritic cells ($p < 0.05$), eosinophils ($p < 0.05$), mast cells ($p < 0.05$), neutrophils ($p < 0.05$), and regulatory T cells ($p < 0.05$). In contrast, the presence of macrophages and monocytes was negatively correlated with the levels of PCBP1-AS1 expression. The results indicated that PCBP1-AS1 played an important role in immune infiltration in cervical cancer. Meanwhile, to study whether the cervical cancer immune microenvironment was different in cases with high PCBP1-AS1 levels compared those with to low levels, we downloaded an RNA expression profile obtained from TCGA. The cervical cancer samples were divided into two groups with the median value of PCBP1-AS1 expression as a cutoff. Then, we explored the expression profiles to obtain a fraction of 18 immune cell subtypes and assessed the differences in their expression levels in the two PCBP1-AS1 expression groups (**Figure 7A**). B cells, CD4+ T cells, M0 macrophages, M2 macrophages, and activated NK cells were significantly affected by PCBP1-AS1 expression. M2 macrophages and activated NK cells were increased ($p < 0.05$) in the low expression group compared to the high expression group. In contrast, CD4+ T cells, B cells, and M0 macrophages were increased in the high expression group ($p < 0.05$). We collected five pairs of samples (five cervical cancer tissues and five adjacent tissues) and extracted total protein from them. Differences in CD4+ T cells (CD4), B cells (CD19), and NK cells (CD56) in normal samples and cervical cancer samples were detected by western

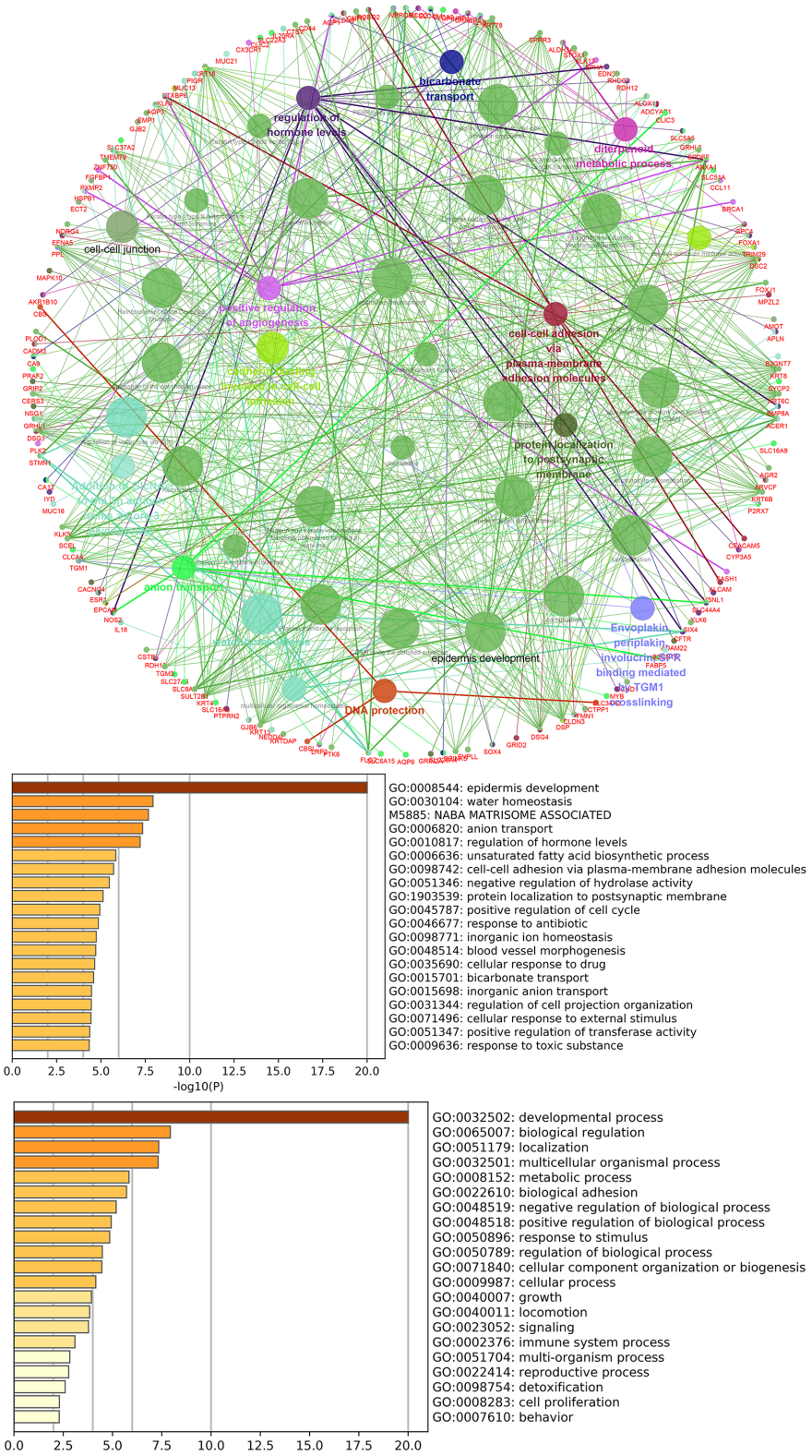


FIGURE 3 | GO and pathway analysis for differentially expressed mRNAs in the highly active subnetwork module.

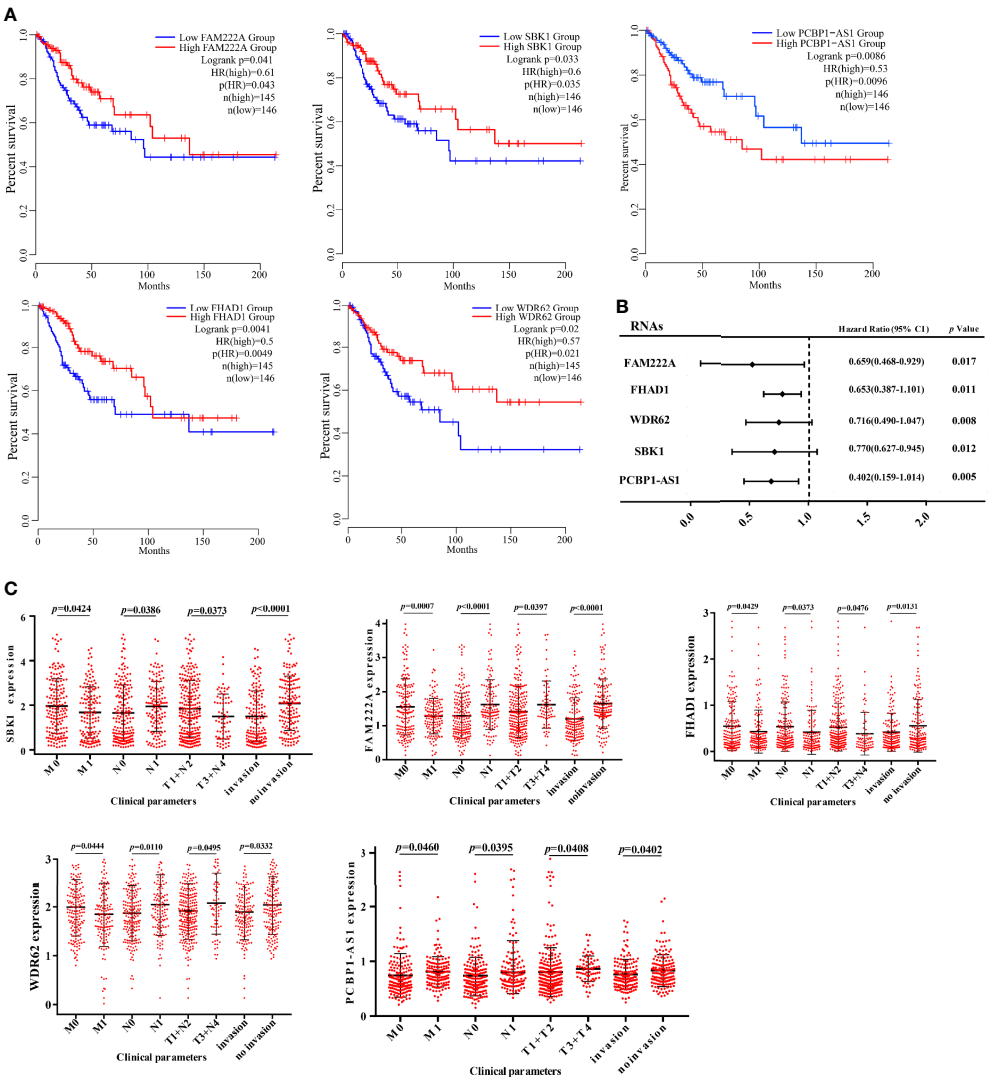


FIGURE 4 | **(A)** Kaplan–Meier analysis results of PCBP1-AS1, FAM222A, FHAD1, WDR62, and SBK1 in CESC. **(B)** Univariate Cox proportional hazards regression analysis of PCBP1-AS1, FAM222A, FHAD1, WDR62, and SBK1 in CESC. **(C)** Expression of PCBP1-AS1, FAM222A, FHAD1, WDR62, and SBK1 correlated significantly with clinicopathological parameters.

TABLE 3 | Multivariate cox regression analysis of RNA signature associated with survival in cervical cancer patients.

	Coefficient	HR	SE	P-value
FAM222A	−0.033	0.79	0.197	0.23
FHAD1	−0.029	0.799	0.277	0.418
WDR62	−0.002	0.728	0.2	0.113
SBK1	−0.007	0.804	0.115	0.047
PCBP1-AS1	−0.069	0.407	0.489	0.046

blotting. The results showed that CD56 protein was present at low levels in the tumor samples. In contrast, CD4 and CD19 protein levels were higher in tumor samples than in normal samples, which indirectly confirmed the results of the above analysis of immune infiltration (**Figures 7B, C**). In addition, we analyzed the correlations between 18 types of immune cells (**Figure 7D**), which revealed that the different infiltrating immune cell subpopulations of cervical cancer were moderately correlated.

Gene Set Enrichment Analysis of PCBP1-AS1

To further analyze the function of PCBP1-AS1, GSEA was performed, and the most differentially (FDR q-val < 0.250, NOM p-val < 0.050) enriched signaling pathways and functions were selected based on the normalized enrichment score (NES). As shown in **Figure 8B**, the GO sets of molecular functions and biological processes significantly associated with

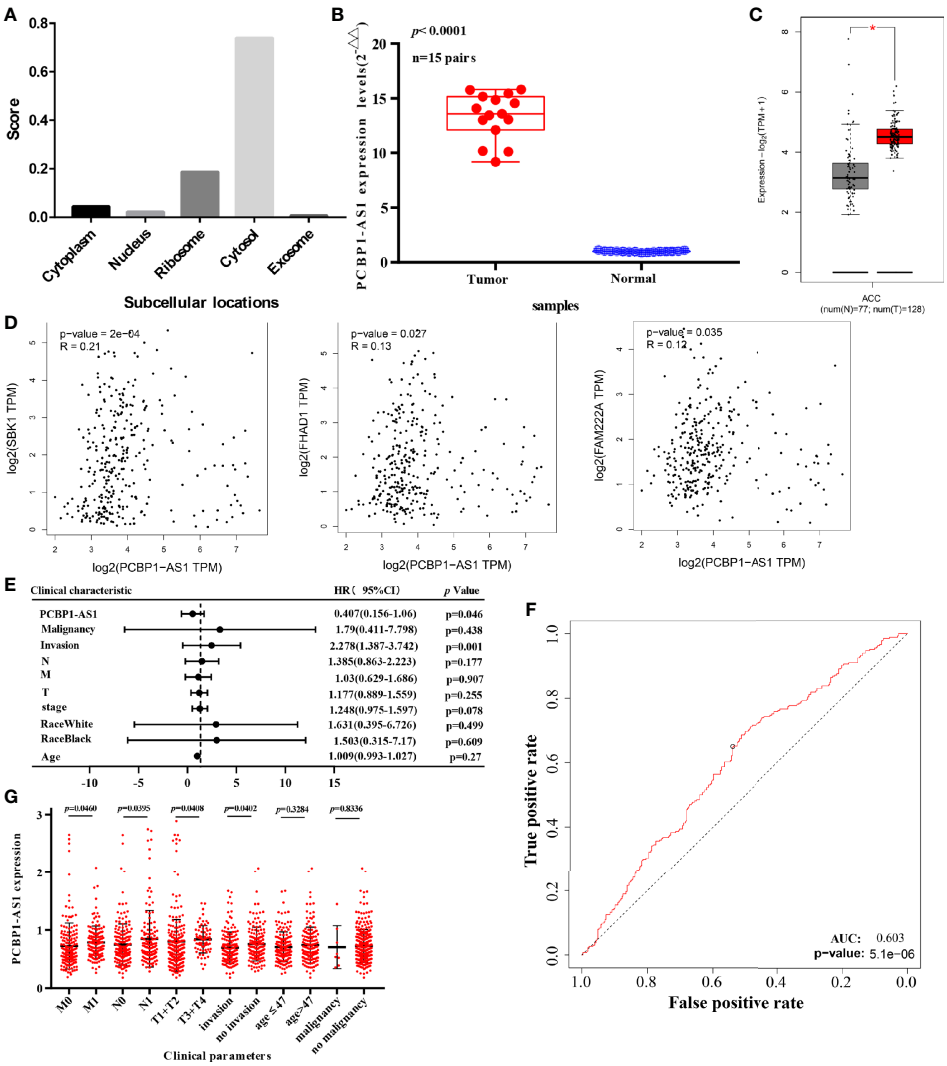


FIGURE 5 | **(A)** Subcellular localizations of PCBP1-AS1 determined by using lncLocator. **(B)** Expression of PCBP1-AS1 in CESC cancer tissues assessed by qPCR. **(C)** Expression of PCBP1-AS1 in CESC from TCGA by GEPIA.*p<0.05. **(D)** Correlation analysis among PCBP1-AS1, FAM222A, FHAD1, and SBK1. **(E)** Multivariate Cox analysis of PCBP1-AS1 expression and other clinicopathological variables. **(F)** ROC curves of PCBP1-AS1. **(G)** Expression of PCBP1-AS1 correlated significantly with clinicopathological parameters.

PCBP1-AS1 expression were cell adhesion, cell migration, cell proliferation, regulation of apoptosis, cell resistance, and chromatin regulation. KEGG pathway analysis showed that the four pathways with the strongest positive correlations with PCBP1-AS1 expression were protein export, proteasome, p53 signaling pathway, and glycolysis gluconeogenesis; the four pathways with the strongest negative correlations were phosphatidylinositol signaling, basal cell carcinoma, bladder cancer, and Notch signaling, as shown in **Figure 8A**. The above GO and KEGG pathway annotations are shown in **Table 5**. These results revealed that the expression level of PCBP1-AS1 was strongly associated with GO functions and pathways regulating cell function (cell adhesion, migration, proliferation, apoptosis, and resistance) and chromosome and protein activity.

PCBP1-AS1 Contributes to HeLa Cell Proliferation, Migration, and Invasion

To further address the biological function of PCBP1-AS1, we used siRNA and overexpression plasmids to alter the expression level of PCBP1-AS1 in HeLa cells and analyzed the effect of PCBP1-AS1 on HeLa cell proliferation, migration, and invasion. Fluorescence microscopy analysis of PCBP1-AS1 revealed successful transfection of the overexpression plasmid (**Figure 9A**), and siRNA-PCBP1-AS1 led to a significant decrease in PCBP1-AS1 expression in HeLa cells (**Figure 9B**). The effect of PCBP1-AS1 on cell proliferation was detected by the CCK-8 assay. Compared with the negative control (NC) group, overexpression of PCBP1-AS1 significantly promoted HeLa cell proliferation, whereas HeLa cell proliferation was significantly

TABLE 4 | Correlation between overall survival and multivariable characteristics *via* (a) univariate analysis (b) multivariate survival analysis.

Clinical characteristic	HR	LOWER	UPER	P
a				
Age	1.015	0.997	1.032	0.097
Race Black	1.376	0.289	6.552	0.689
Race White	1.564	0.379	6.447	0.536
Stage	1.494	1.216	1.837	<0.001
T	1.375	1.117	1.694	0.003
M	1.261	0.812	1.958	0.302
N	1.406	0.907	2.181	0.128
Invasion	2.360	1.484	3.753	<0.001
malignancy	1.679	0.408	6.906	0.473
PCBP1-AS1	0.402	0.159	1.014	0.033
Clinical characteristic	**HR**	**LOWER**	**UPER**	**P**
b				
Age	1.009	0.993	1.027	0.27
Race Black	1.503	0.315	7.17	0.609
Race White	1.631	0.395	6.726	0.499
Stage	1.248	0.975	1.597	0.078
T	1.177	0.889	1.559	0.255
M	1.03	0.629	1.686	0.907
N	1.385	0.863	2.223	0.177
Invasion	2.278	1.387	3.742	0.001
malignancy	1.79	0.411	7.798	0.438
PCBP1-AS1	0.407	0.156	1.06	0.046

impaired by PCBP1-AS1 knockdown ($p < 0.001$) (**Figure 9E**). Meanwhile, the wound-healing assay revealed that overexpression of PCBP1-AS1 significantly enhanced HeLa cell migration, and siRNA-PCBP1-AS1 showed a notably slower scratch closure rate than control cells (**Figures 9C, D**), which revealed that silencing PCBP1-AS1 inhibited HeLa cell migration ($p < 0.001$). Furthermore, the Transwell assay demonstrated that PCBP1-AS1 knockdown HeLa cells displayed significantly lower invasion potential than the control cells ($p < 0.001$) (**Figures 9F, G**). Collectively, these results suggest that the expression level of PCBP1-AS1 affected the proliferation, migration, and invasion of cervical cancer cells.

DISCUSSION

Cervical cancer is one of the most common malignancies in females, and it has the highest mortality among female reproductive system malignancies (1, 2). Despite the development of diagnostic and treatment strategies, the prognosis of cervical cancer patients is still very poor, mainly due to cancer metastasis and recurrence (5–7). Thus far, a series of studies have indicated that lncRNAs exert substantial effects on the pathogenesis of carcinomas, suggesting that lncRNAs might act as prognostic indicators in tumorigenesis and cancer development (14–16). For example, PSMB8-AS1 contributes to pancreatic cancer progression by modulating the miR-382-3p/STAT1/PD-L1 axis (21). However, limited research has been performed on the transcriptomic profiles of cervical cancer, and the functional roles of lncRNAs in cervical cancer pathogenesis remain largely unknown. Hence, comprehensively understanding the lncRNA profile of cervical cancer and analyzing the mechanism of action involving lncRNAs might provide new thinking in the pathogenesis of this disease.

In this study, high-throughput microarray analysis was performed to characterize the significantly differentially expressed lncRNAs and mRNAs between cervical cancer patients and controls, which might be involved in cervical cancer progression. In total, 130 lncRNAs and 656 mRNAs were found to be dysregulated. The functions of lncRNAs are closely associated with downstream target mRNAs, which they may regulate directly or indirectly (22, 23). Therefore, a global lncRNA-mRNA coexpression cis- and trans-regulatory network was constructed with 127 differentially expressed lncRNAs and 387 differentially expressed mRNAs, which could be successfully used for disease-related lncRNA identification. Based on the lncRNA-PCG functional network, we identified a highly active subnetwork module, including 19 lncRNAs and 47 mRNAs, by the jActive module. GO and KEGG enrichment analysis based on the 47 differentially expressed mRNAs indicated that several biological processes and pathways may play important roles in cervical cancer pathogenesis, including epidermal development, cell cycle, cell resistance, epidermal cell differentiation, and regulation of the intracellular estrogen receptor signaling pathway, which indicated that these DEmRNAs may be related to EMT. EMT causes dissociated epithelial cells to acquire migratory and invasive capacities and endows cancer cells with the ability to migrate to distant tissues. This functional annotation provides bioinformatics-based evidence regarding the potential mechanism promoting cervical cancer occurrence. Based on Kaplan-Meier analysis, PCBP1-AS1 and four mRNAs (FAM222A, FHAD1, WDR62, and SBK1) were identified as potential prognostic factors for

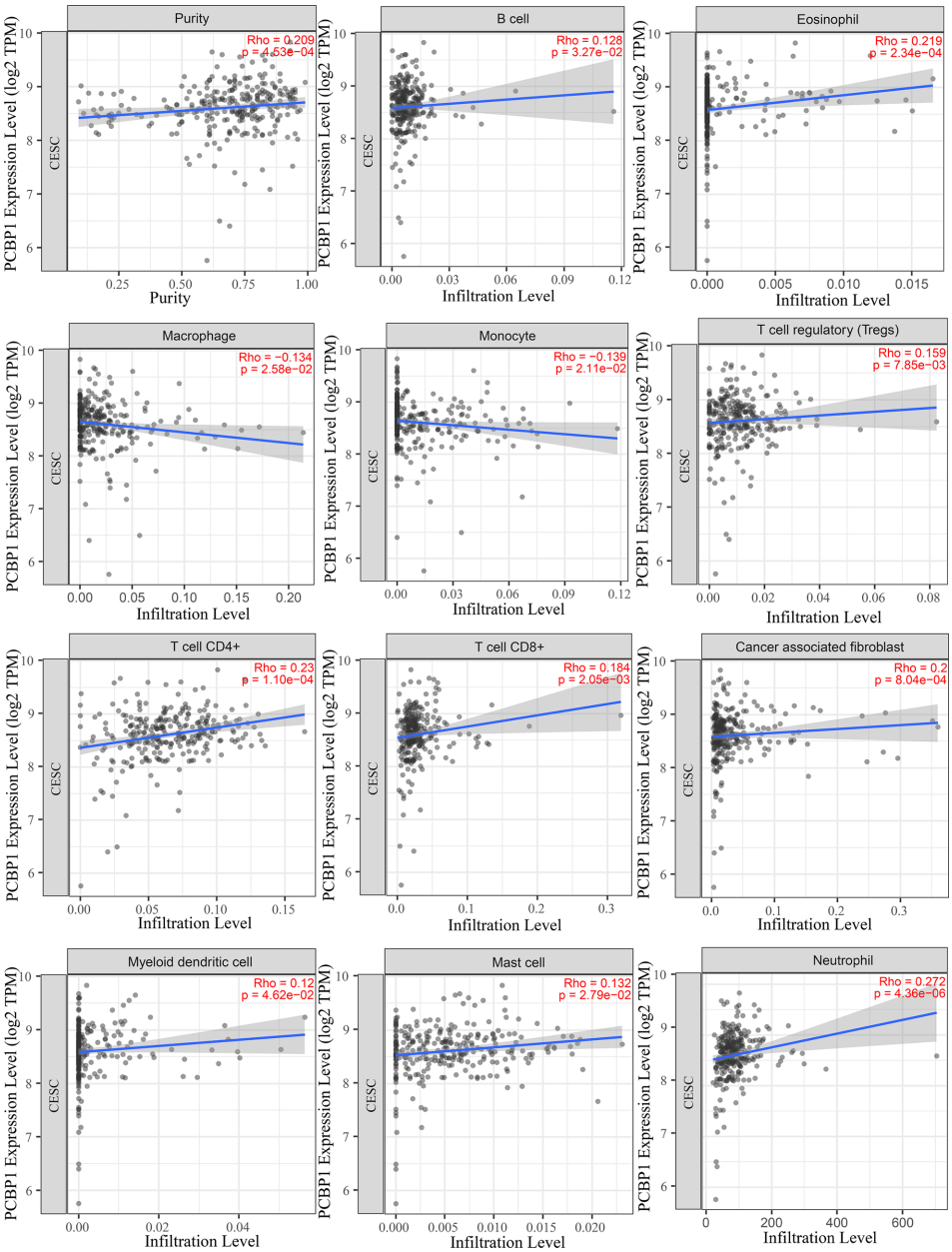


FIGURE 6 | Correlations between PCBP1-AS1 expression and immune infiltration levels, with purity adjustment.

cervical cancer patients. The expression of these DERNAs was significantly correlated with tumor clinical stage, pathologic TNM, and lymphatic invasion.

To the best of our knowledge, the five DERNAs have rarely been reported in previous studies, and their functions in cervical cancer are largely unknown. Originally, PCBP1-AS1 was identified in cervical cancer tissues through microarray expression profiling (24). However, its expression and biological function in cervical cancer tissues and cells have not been studied. Luo et al. reported that PCBP1-AS1 aggravated the progression of hepatocellular carcinoma by regulating the PCBP1/PRL-3/AKT pathway (25). Luan et al. noted that PCBP1-AS1 promoted the autophagy of glioma cells (26). FAM222A is Chromosome 12 Open Reading Frame 34. It was reported that FAM222a is related to chemotherapy resistance in gastric cancer, and its antisense RNA can regulate the migration of non-small cell lung cancer cells (27, 28). FHAD1 was reported by Zhao et al. to be a marker for the occurrence of prostate cancer (29). WDR62 was identified as a scaffold protein in the JNK signaling pathway (30). Zhou found that inactivation of

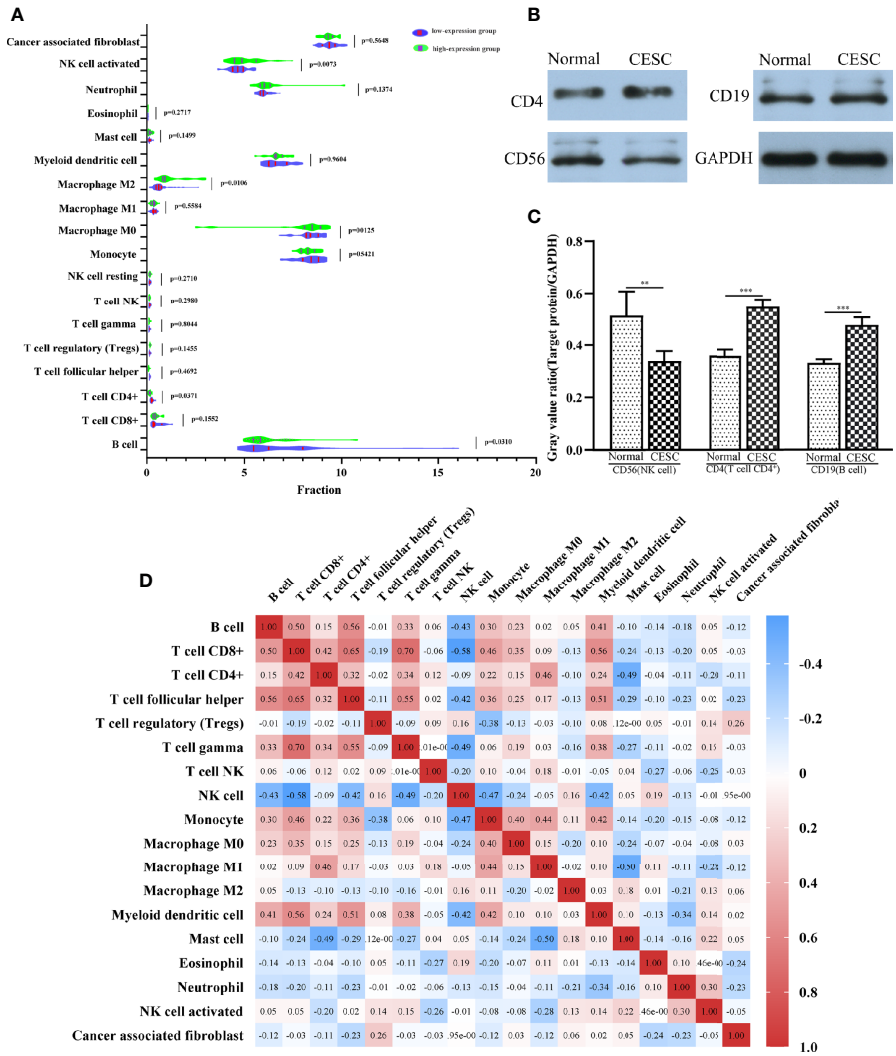


FIGURE 7 | **(A)** The varied proportions of 18 subtypes of immune cells in high and low PCBP1-AS1 expression groups in tumor samples. **(B)** Western blot results of CD4 (representing CD4+ T cells), CD19 (representing B cells), and CD56 (representing NK cells) protein expression in cervical cancer tissues and corresponding adjacent normal cervical cancer tissues. GAPDH was used as a control. **(C)** Gray value ratios of CD4/GAPDH, CD19/GAPDH, and CD56/GAPDH in cervical cancer tissues and corresponding adjacent normal cervical cancer tissues. $^{**}p < 0.01$,$^{***}p < 0.001$. **(D)** Heatmap of 18 infiltrating immune cells in tumor samples.

WDR62 could cause defects in female meiotic initiation, which led to the occurrence of female reproductive diseases (31). SBK1 is also a peptide domain. Wang found that SBK1 was dysregulated in several cancer tissues, especially in ovarian cancer, and showed that SBK1 played an important role during ovarian carcinogenesis (32). According to the above studies, these five hub genes are likely to be involved in the occurrence or development of cervical cancer.

Our research results showed that PCBP1-AS1 expression was increased in cervical cancer tissues compared with paired adjacent normal tissues and was a prognostic biomarker for cervical cancer. Additionally, we found a role for upregulated PCBP1-AS1 as an independent prognostic factor for poor OS. Cervical cancer patients with high PCBP1-AS1 expression are more likely to have a more advanced stage, TNM status, and lymph metastasis than those with low PCBP1-AS1 expression. In addition, we analyzed the connections between PCBP1-AS1 expression and immune infiltration levels in cervical cancer by TIMER. We found a relationship between PCBP1-AS1 and T cell, B cell, myeloid dendritic cell, eosinophil, mast cell, neutrophil, macrophage, and monocyte infiltration. Furthermore, the immune infiltration score analysis showed that B cells, CD4+ T cells, M0 macrophages, M2 macrophages, and NK cells were related to PCBP1-AS1 expression. The results revealed that CD4+, B cells, and M0 macrophages were increased in the high expression group, whereas the levels of M2 macrophages and activated NK cells were decreased. It is reported that NK cells are important biological barriers that

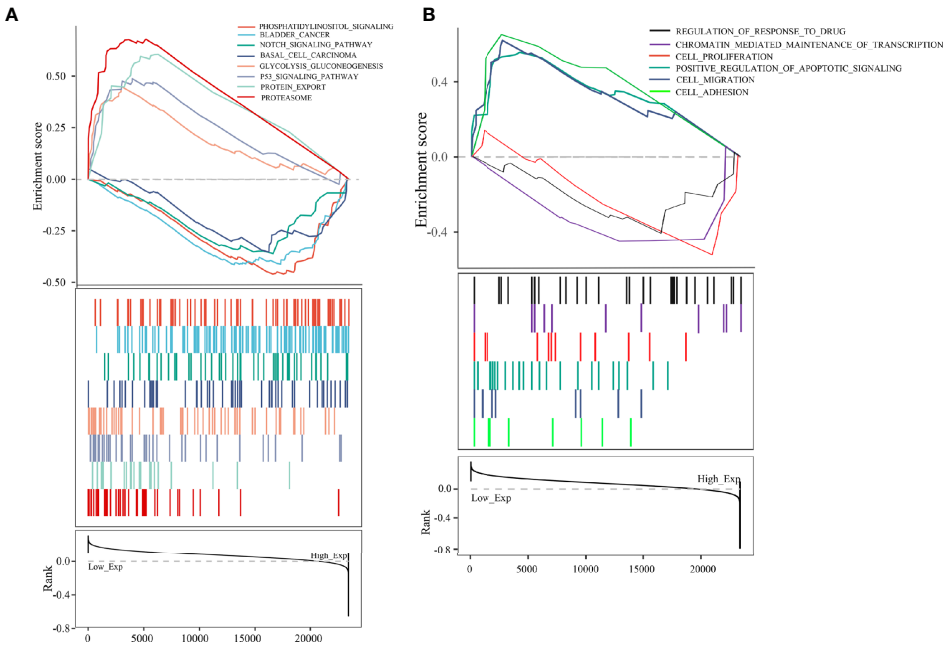


FIGURE 8 | **(A)** KEGG pathway analysis showed five positively correlated groups and five negatively correlated groups. **(B)** GO term analysis revealed three positively correlated groups and three negatively correlated groups.

TABLE 5 | Signaling pathways most significantly correlated with PCBP1-AS1 expression based on their normalized enrichment score (NES) and p-value.

	GO name	NES	NOM p-value
Positive	Cell adhesion	1.60	0.04
	Cell migration	1.59	0.027
	Negative regulation of apoptotic	1.62	0.032
Negative	Positive regulation of response to drug	–1.53	0.037
	Chromatin mediated maintenance of transcription	–1.61	0.024
	Cell proliferation	–1.58	0.032
	KEGG name	**NES**	**NOM p-value**
Positive	Protein export	1.63	0.046
	Proteasome	1.70	0.034
	p53 signaling pathway	1.52	0.032
	Glycolysis gluconeogenesis	1.57	0.035
Negative	Phosphatidylinositol signaling	–1.90	<0.001
	Basal cell carcinoma	–1.37	0.011
	Bladder cancer	–1.71	0.010
	Notch signaling	–1.43	0.044

are resident in the cervix and can identify and kill virus-infected cells rapidly through pathways that do not require preactivation (33). Other research suggests that HPV16 disables the increased NK cells in the early lesion of the cervix (34). Here, we speculate that these phenomena may be a possible mechanism by which PCBP1-AS1 regulates the functions of NK cells in cervical cancer. Furthermore, the overexpression of PCBP1-AS1 may inhibit efficient NK cell immune responses and infiltration. Overall, PCBP1-AS1 plays a crucial role in the regulation and recruitment of immune infiltrating cells in cervical cancer. However, these results need to be further validated in combination with clinical trials.

Equally important, we performed GSEA to further analyze the biological function of PCBP1-AS1. Our results showed that the main significant pathways for PCBP1-AS1 included the p53 signaling pathway and Notch signaling. Notch signaling is well known to be one of the most frequently activated signaling pathways in cancer and is involved in cell cycle regulation (35) and immune responses (36). Rong et al. reported that activated Notch signaling may lead to the development of cervical cancer by regulating Numb splicing (37). Similarly, P53 also plays a complex role in promoting the cell cycle, cell senescence, and apoptosis. Wild-type p53 can promote the cancer metabolic switch by inducing PUMA-dependent suppression of oxidative

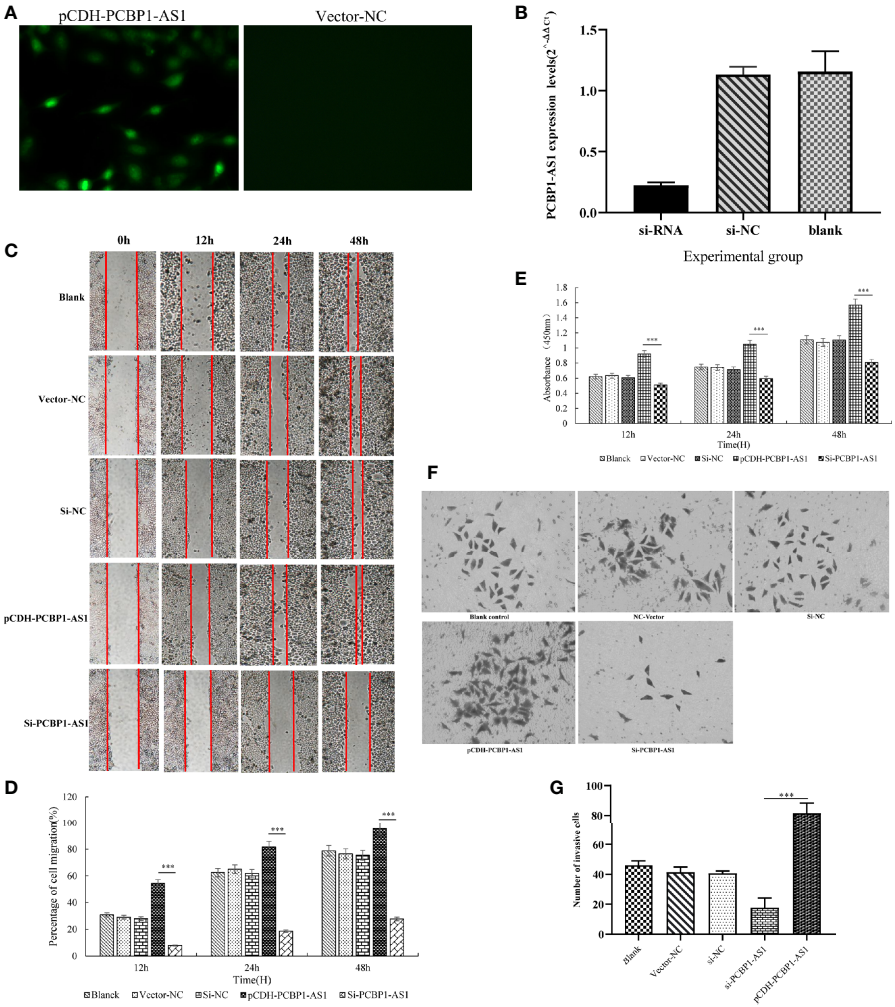


FIGURE 9 | **(A)** Overexpression of PCBP1-AS1 in HeLa cells analyzed by fluorescence microscopy. **(B)** PCBP1-AS1 knockout efficiency in HeLa cells analyzed by qPCR. **(C)** and **(D)** The effects of PCBP1-AS1 knockdown or overexpression on HeLa cell migration measured using the migration assay. ***$p < 0.001$. **(E)** The effects of PCBP1-AS1 knockdown and overexpression on the viability of HeLa cells measured using the CCK-8 assay. **(F)** and **(G)** Transwell invasion assay and average number of invasive PCBP1-AS1 knockdown and overexpression cells. ***p < 0.001.

phosphorylation (38). Furthermore, recent research found that many more genes could promote proliferation and suppress apoptosis in cervical cancer cells by inhibiting and activating the p53 signaling pathway (39–41). Our results help to deepen the understanding of the biological functions of PCBP1-AS1 in cervical cancer. Furthermore, to validate the identified biological functions, gain-of-function experiments were performed in HeLa cells. Our results showed that PCBP1-AS1 depletion could significantly inhibit cervical cancer cell proliferation. Wound healing assays and Transwell assays further demonstrated that downregulation of PCBP1-AS1 could reduce the migration and invasion ability of cervical cancer cells. Collectively, these results provide mechanistic evidence supporting the finding that PCBP1-AS1 upregulation is associated with more advanced stage, TNM status, and lymph metastasis. PCBP1-AS1 may also be a useful biomarker for cervical cancer.

In conclusion, we used transcriptome sequencing technology to profile the lncRNAs of both cervical cancer and adjacent mucosa from 15 patients. A total of 130 lncRNAs and 656 mRNAs were systematically screened, many of which played important roles in regulating cell biological functions. These sequencing data provide an important resource for future studies of key lncRNAs in cervical cancer. Of these, PCBP1-AS1 was found to be a new biomarker for the prognosis of cervical cancer patients and to regulate cell proliferation and migration. In addition, this study helps to elucidate the roles of immune cell infiltration and lncRNAs in cervical cancer. With a better understanding of the biological function of PCBP1-AS1, this molecule could act as an effective biomarker for the diagnosis and treatment of cervical cancer and may help clinicians make appropriate choices for targeted therapy for the treatment of cervical cancer in the future.

DATA AVAILABILITY STATEMENT

The raw sequencing dates presented in the study are publicly available, which has been uploaded to Gene Expression Omnibus (GEO). These data can be found here: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE167362.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Jiangning Hospital of Nanjing Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

QL and LHL were the principal investigators who designed and conceived the study and obtained financial support. LHL and QP analyzed the data and wrote the manuscript. YX, LL, and MG prepared the dataset. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

We thank the team that built the Timer, GEPIA, Starbase online analysis page and thank the free online platform of Sanger box tools, platform of Shanghai Ordovician Biotechnology Co., LTD.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.634732/full#supplementary-material

Supplementary Figure 1 | Prognostic risk score model analysis of PCBP1-AS1 and SBK1 in CESC patients. **(A)** From top to bottom: risk score distribution, patient survival status distribution, and heatmap of PCBP1-AS1 and SBK1 expression profiles ranked by risk score. **(B)** Kaplan–Meier curves for high-risk and low-risk groups. **(C)** The ROC curves for predicting survival in CESC patients by the risk score.

REFERENCES

1. Shafabakhsh R, Reiter RJ, Mirzaei H, Teymoordash SN, Asemi Z. Melatonin: A new inhibitor agent for cervical cancer treatment. *J Cell Physiol* (2019) 234:21670–82. doi: 10.1002/jcp.28865
2. Vu M, Yu J, Awolude OA, Chuang L. Cervical cancer worldwide. *Curr Probl Cancer* (2018) 42:457–65. doi: 10.1016/j.currproblcancer.2018.06.003
3. Wei M, Zhou W, Bi Y, Wang H, Liu Y, Zhang ZJ. Rising Mortality Rate of Cervical Cancer in Younger Women in Urban China. *J Gen Intern Med* (2019) 34:281–4. doi: 10.1007/s11606-018-4732-z
4. Small W, Bacon MA, Bajaj A, Chuang LT, Fisher BJ, Harkenrider MM, et al. Cervical cancer: A global health crisis. *Cancer* (2017) 123:2404–12. doi: 10.1002/cncr.30667
5. Koh WJ, Abu-Rustum NR, Bean S, Bradley K, Campos SM, Cho KR, et al. Cervical Cancer, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* (2019) 17:64–84. doi: 10.6004/jnccn.2019.0001
6. Mallmann P, Mallmann C. Neoadjuvant and Adjuvant Chemotherapy of Cervical Cancer. *Oncol Res Treat* (2016) 39:522–4. doi: 10.1159/000449023
7. Vordermark D. Radiotherapy of Cervical Cancer. *Oncol Res Treat* (2016) 39:516–20. doi: 10.1159/000448902
8. Cancer Genome Atlas Research Network. Albert Einstein College of Medicine; Analytical Biological Services; Integrated genomic and molecular characterization of cervical cancer. *Nature* (2017) 543:378–84. doi: 10.1038/nature21386
9. Minion LE, Tewari KS. Cervical cancer - State of the science: From angiogenesis blockade to checkpoint inhibition. *Gynecol Oncol* (2018) 148:609–21. doi: 10.1016/j.ygyno.2018.01.009
10. ENCODE Project Consortium, Birney E, Stamatoyannopoulos JA. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* (2007) 447:799–816. doi: 10.1038/nature05874
11. Outeiro-Pinho G, Barros-Silva D, Correia MP, Henrique R, Jerónimo C. Renal Cell Tumors: Uncovering the Biomarker Potential of ncRNAs. *Cancers (Basel)* (2020) 12(8):2214. doi: 10.3390/cancers12082214
12. Yi J, Li S, Wang C, Cao N, Qu H, Cheng C, et al. Potential applications of polyphenols on main ncRNAs regulations as novel therapeutic strategy for cancer. *BioMed Pharmacother* (2019) 113:108703. doi: 10.1016/j.biopha.2019.108703
13. Feng W, Su Z, Yin Q, Zong W, Shen X, Ju S. ncRNAs associated with drug resistance and the therapy of digestive system neoplasms. *J Cell Physiol* (2019) 234:19143–57. doi: 10.1002/jcp.28551
14. Hu Q, Ye Y, Chan LC, Li Y, Liang K, Lin A, et al. Oncogenic lncRNA downregulates cancer cell antigen presentation and intrinsic tumor suppression. *Nat Immunol* (2019) 20:835–51. doi: 10.1038/s41590-019-0400-7
15. Dong P, Xiong Y, Yue J, B Hanley S, Kobayashi N, Todo Y, et al. Exploring lncRNA-Mediated Regulatory Networks in Endometrial Cancer Cells and the Tumor Microenvironment: Advances and Challenges. *Cancers (Basel)* (2019) 11:234. doi: 10.3390/cancers11020234
16. 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.
17. Liu Y, Zhang R, Qiu F, Li K, Zhou Y, Shang D, et al. Construction of a lncRNA-PCG bipartite network and identification of cancer-related lncRNAs: a case study in prostate cancer. *Mol Biosyst* (2015) 11:384–93. doi: 10.1039/c4mb00439f
18. Wang P, Ning S, Zhang Y, Li R, Ye J, Zhao Z, et al. Identification of lncRNA-associated competing triplets reveals global patterns and prognostic markers for cancer. *Nucleic Acids Res* (2015) 43:3478–89. doi: 10.1093/nar/gkv233
19. Wang J, Li Z, Gao A, Wen Q, Sun Y. The prognostic landscape of tumor-infiltrating immune cells in cervical cancer. *BioMed Pharmacother* (2019) 120:109444. doi: 10.1016/j.biopha.2019.109444
20. Luo W, Wang M, Liu J, Cui X, Wang H. Identification of a six lncRNAs signature as novel diagnostic biomarkers for cervical cancer. *J Cell Physiol* (2020) 235:993–1000. doi: 10.1002/jcp.29015
21. Zhang H, Zhu C, He Z, Chen S, Li L, Sun C. LncRNA PSMB8-AS1 contributes to pancreatic cancer progression via modulating miR-382-3p/STAT1/PD-L1 axis. *J Exp Clin Cancer Res* (2020) 39:179. doi: 10.1186/s13046-020-01687-8
22. Zhang Y, Xu Y, Feng L, Li F, Sun Z, Wu T, et al. Comprehensive characterization of lncRNA-mRNA related ceRNA network across 12 major cancers. *Oncotarget* (2016) 7:64148–67. doi: 10.18632/oncotarget.11637
23. Liu XD, Xie DF, Wang YL, Guan H, Huang RX, Zhou PK. Integrated analysis of lncRNA-mRNA co-expression networks in the α-particle induced carcinogenesis of human branchial epithelial cells. *Int J Radiat Biol* (2019) 95:144–55. doi: 10.1080/09553002.2019.1539880
24. Campos-Parra AD, Padua-Bracho A, Pedroza-Torres A, Figueroa-González G, Fernández-Retana J, Millan-Catalan O, et al. Comprehensive transcriptome analysis identifies pathways with therapeutic potential in locally advanced cervical cancer. *Gynecol Oncol* (2016) 143:406–13. doi: 10.1016/j.ygyno.2016.08.327
25. Luo T, Gao Y, Zhangyuan G, Xu X, Xue C, Jin L, et al. lncRNA PCBP1-AS1 Aggravates the Progression of Hepatocellular Carcinoma via Regulating

PCBP1/PRL-3/AKT Pathway. *Cancer Manag Res* (2020) 12:5395–408. doi: 10.2147/CMAR.S249657
26. Luan F, Chen W, Chen M, Yan J, Chen H, Yu H, et al. An autophagy-related long non-coding RNA signature for glioma. *FEBS Open Bio* (2019) 9:653–67. doi: 10.1002/2211-5463.12601
27. Zhang X, Han J, Du L, Li X, Hao J, Wang L, et al. Unique metastasis-associated lncRNA signature optimizes prediction of tumor relapse in lung adenocarcinoma. *Thorac Cancer* (2020) 11(3):728–37. doi: 10.1111/1759-7714.13325
28. Choi SJ, Jung SW, Huh S, Chung YS, Cho H, Kang H. Alteration of DNA Methylation in Gastric Cancer with Chemotherapy. *J Microbiol Biotechnol* (2017) 27(8):1367–78. doi: 10.4014/jmb.1704.04035
29. Zhao S, Geybels MS, Leonardson A, Rubicz R, Kolb S, Yan Q, et al. Epigenome-Wide Tumor DNA Methylation Profiling Identifies Novel Prognostic Biomarkers of Metastatic-Lethal Progression in Men Diagnosed with Clinically Localized Prostate Cancer. *Clin Cancer Res* (2017) 23(1):311–9. doi: 10.1158/1078-0432.CCR-16-0549
30. Wasserman T, Katsenelson K, Daniliuc S, Hasin T, Choder M, Aronheim A. A novel c-Jun N-terminal kinase (JNK)-binding protein WDR62 is recruited to stress granules and mediates a nonclassical JNK activation. *Mol Biol Cell* (2010) 21(1):117–30e09–06-0512. doi: 10.1091/mbc
31. Zhou Y, Qin Y, Qin Y, Xu B, Guo T, Ke H, et al. Wdr62 is involved in female meiotic initiation via activating JNK signaling and associated with POI in humans. *PloS Genet* (2018) 14(8):e1007463. doi: 10.1371/journal.pgen.1007463
32. Wang P, Guo J, Wang F, Shi T, Ma D. Human SBK1 is dysregulated in multiple cancers and promotes survival of ovary cancer SK-OV-3 cells. *Mol Biol Rep* (2011) 38(5):3551–9. doi: 10.1007/s11033-010-0465-8
33. Renoux VM, Bisig B, Langers I, Dortu E, Clémenceau B, Thiry M, et al. Human papillomavirus entry into NK cells requires CD16 expression and triggers cytotoxic activity and cytokine secretion. *Eur J Immunol* (2011) 41:3240–52. doi: 10.1002/eji.201141693
34. Zhang J, Jin S, Li X, Liu L, Xi L, Wang F, et al. Human Papillomavirus Type 16 Disables the Increased Natural Killer Cells in Early Lesions of the Cervix. *J Immunol Res* (2019) 2019:9182979. doi: 10.1155/2019/9182979
35. Qin J, Wang R, Zhao C, Wen J, Dong H, Wang S, et al. Notch signaling regulates osteosarcoma proliferation and migration through Erk phosphorylation. *Tissue Cell* (2019) 59:51–61. doi: 10.1016/j.tice.2019.07.002
36. Vieceli Dalla Sega F, Fortini F, Aquila G, Campo G, Vaccarezza M, Rizzo P. Notch Signaling Regulates Immune Responses in Atherosclerosis. *Front Immunol* (2019) 10:1130. doi: 10.3389/fimmu.2019.01130
37. Rong C, Feng Y, Ye Z. Notch is a critical regulator in cervical cancer by regulating Numb splicing. *Oncol Lett* (2017) 13:2465–70. doi: 10.3892/ol.2017.5683
38. Kim J, Yu L, Chen W, Xu Y, Wu M, Todorova D, et al. Wild-Type p53 Promotes Cancer Metabolic Switch by Inducing PUMA-Dependent Suppression of Oxidative Phosphorylation. *Cancer Cell* (2019) 35:191–203.e8. doi: 10.1016/j.ccell.2018.12.012
39. Gao X, Li Q, Chen G, He H, Ma Y. MAGEA3 promotes proliferation and suppresses apoptosis in cervical cancer cells by inhibiting the KAP1/p53 signaling pathway. *Am J Transl Res* (2020) 12:3596–612.
40. Liu Y, Li L, Liu Y, Geng P, Li G, Yang Y, et al. RECK inhibits cervical cancer cell migration and invasion by promoting p53 signaling pathway. *J Cell Biochem* (2018) 119:3058–66. doi: 10.1002/jcb.26441
41. Cui L, Nai M, Zhang K, Li L, Li R. lncRNA WT1-AS inhibits the aggressiveness of cervical cancer cell via regulating p53 expression via sponging miR-330-5p. *Cancer Manag Res* (2019) 11:651–67. doi: 10.2147/CMAR.S176525

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

CASE REPORT
published: 25 March 2021
doi: 10.3389/fonc.2021.658389

Case Report: Coinheritance of Germline Mutations in *APC* and *BRCA1* in Colorectal Cancer

Wei Huang[1], *Jin Bian*[2], *Xiaoping Qian*[3], *Lin Shao*[4], *Haiyan Li*[4], *Lu Zhang*[4] *and Lin Wang*[5*]

[1] Department of Oncology, Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China, [2] Department of Oncology, Nanjing Jinling Hospital, Nanjing, China, [3] The Comprehensive Cancer Centre, Nanjing Drum Tower Hospital, Medical School of Nanjing University, Nanjing, China, [4] Department of Medicine, Burning Rock Biotech, Guangzhou, China, [5] Department of Oncology, Nanjing Tongren Hospital, Nanjing, China

Deleterious mutations in *APC* gene cause the autosomal dominant familial adenomatous polyposis (FAP) which is typically characterized by the occurrence of hundreds to thousands of colorectal adenomas that eventually lead to colorectal cancers (CRCs). *BRCA1/2* are the two major susceptibility genes for breast and ovarian cancers. Here, we reported a coinheritance of mutations in *APC* and *BRCA1* genes in a 20-year-old CRC patient with typical clinical features for FAP. Multiple relatives in the family of the patient were affected by colorectal and other cancers. Next-generation sequencing analysis using a panel consisting of 53 hereditary cancer related genes revealed a maternally inherited *APC* (exon15cn_del) mutation and a paternally inherited *BRAC1* (p.Ile1824AspfsX3) mutation. This is the first coexistence of *APC* and *BRCA1* mutations in a CRC patient with the mutation inheritance pattern comprehensively characterized in the family. The patient underwent a colonoscopy and a subtotal colectomy and was subsequently diagnosed with colonic adenocarcinomas accompanied with hundreds of tubulovillous adenomas. The case reveals the scenario where two disease-causing mutations of different hereditary tumor syndromes coexist, and illustrates the importance of evaluating detailed family history and performing a multiple-gene panel test in patients with hereditary cancer.

Keywords: familial adenomatous polyposis (FAP), *adenomatous polyposis coli* (*APC*) gene, *breast cancer susceptibility* (*BRCA*) gene, double germline mutations, next generation sequencing (NGS)

OPEN ACCESS

Edited by:
João Agostinho Machado-Neto, University of São Paulo, Brazil

Reviewed by:
Juan Luiz Coelho Da Silva, University of São Paul, Brazil
Larissa Costa De Almeida, Universidade de São Paulo, Brazil

***Correspondence:**
Lin Wang
wanglin81yy@163.com

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

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

Citation:
Huang W, Bian J, Qian X, Shao L, Li H, Zhang L and Wang L (2021) Case Report: Coinheritance of Germline Mutations in APC and BRCA1 in Colorectal Cancer. Front. Oncol. 11:658389. doi: 10.3389/fonc.2021.658389

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer worldwide (6.1%) and the second leading cause of cancer-related mortality (9.2%) (1). Approximately 10% of CRC patients are diagnosed at an age younger than 50 years, which is a hallmark of inherited cancer predisposition. Germline mutations in the mismatch repair genes *MLH1*, *MSH2*, *MSH6* and *PMS2*, or *EPCAM* lead to Lynch syndrome, the most common known cause of hereditary CRC and comprising 4% to 13.5% of early-onset CRC patients (2, 3).

Adenomatous polyposis coli (*APC*), located on chromosome 5q21-q22, is one of the tumor-suppressor genes frequently inactivated in the early progression of colorectal carcinogenesis (4, 5). Its primary transcript (NM_000038.6) has 16 exons with 1-15 coding a protein of 2843 amino acids (6). Deleterious germline mutations in *APC* cause familial adenomatous polyposis (FAP) which account for about 0.5% of all CRCs (4, 7). Individuals harboring a germline *APC* mutation can develop multiple adenomas caused by inactivation of the remaining allele in the colorectum *via* gain of additional somatic *APC* mutations or loss of heterozygosity (LOH) at this locus. FAP is typically characterized by the occurrence of hundreds to thousands of colorectal adenomas within 20 years which invariably lead to CRC if not detected early and removed.

Breast cancer susceptibility genes (*BRCA*) consisting of *BRCA1* and *BRCA2*, are also important tumor-suppressor genes (8, 9). *BRCA1* gene, located at chromosome 17q21, consists of 23 exons encoding a protein of 1863 amino acids. *BRCA2* is located at chromosome 13q12 and included 27 exons encoding a large protein product of 3418 amino acids. Mutations in *BRCA1/2* genes have been discovered in multiple malignancies (10). The cumulative breast cancer risk and ovarian cancer risk for mutation carriers are 5 times and 10-20 times higher than that for non-carriers (11), respectively. However, investigations on whether *BRCA1/2* mutations increase lifetime risk of developing colorectal cancer have yielded conflicting results (12–15).

In the present study, we describe a 20-year-old male with familial CRC harboring concurrent germline mutations in *APC* and *BRCA1* which were inherited maternally and paternally respectively.

CASE REPORT

The reported patient was a 20-year-old male (**Figure 1**, III2) with a family history of cancers. His mother (II6) was first diagnosed with a grade II adenocarcinoma of colon accompanied with multiple polyps at the age of 48 and received a systemic chemotherapy. Subsequently, the detailed family history of his mother was assessed. His maternal grandfather (I3) died of gastric cancer when he was 60 years old. His maternal grandmother (I4) was diagnosed with CRC at the age of 59 and died at the age of 76. Two of his maternal aunts (II7&II9) died from CRC at the age of 39 and 50, respectively. His maternal uncle (II11) died of lung cancer at the age of 59. One of his female maternal cousins (III4) was diagnosed with CRC and subsequently received surgery at the age of 30 and was still alive.

Due to the familial cancer history, in June 2018, the white blood cell sample was collected from the mother of the patient (proband, II6) and subjected to a next generation sequencing (NGS)-based genetic test using a panel consisting of 53 hereditary cancer related genes (Ugene, Burning Rock Biotech, China) with a median sequencing depth of 400x. The copy number variation analysis based on the sequencing depth revealed a copy number 1 for the exon 15 in *APC* gene, indicating a heterozygous loss of *APC* exon 15 (exon15cn_del, **Figure 2A**). Genetic test was also performed on two of the patient's maternal cousins (III7&III6) and revealed the same *APC* mutation in the male cousin (III7) but no mutation present in female cousin (III6).

NGS with the same 53-gene panel performed on the reported patient (III2) and his sister (III1) revealed existence of the germline *APC* exon15cn_del mutation in both (**Figure 2B**).

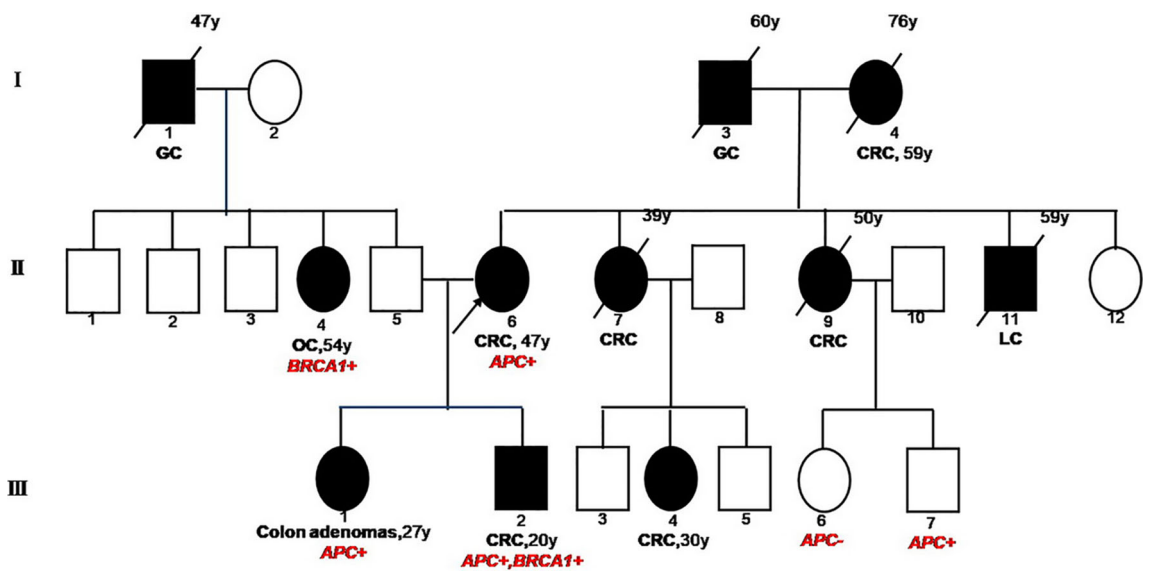


FIGURE 1 | Familial Pedigree of case. The proband is indicated by an arrowhead. Squares represent males, circles represent females. Solid symbols represent affected individuals. Symbols with slash indicate deceased individuals. Age at cancer diagnosis is reported following the corresponding disease and the age of death is reported on the top right corner of symbol. GC, gastric cancer; CRC, colorectal cancer; OC, ovarian cancer; LC, lung cancer.

However, in addition to *APC* mutation, the male patient also harbored an open reading frame shift mutation in *BRAC1* (c. 5470_5477delATTGGGCA, p.Ile1824AspfsX3, **Figure 2C**). The paternal family history of the patient showed that his paternal grandfather (I1) died from gastric cancer at the age of 47. His paternal aunt (II4) was diagnosed with ovarian cancer at the age of 54 years and is still alive. We also identified a same *BRAC1* (p.Ile1824AspfsX3) germline mutation in her.

Taken together, the family clinical history and identified deleterious mutations were highly suggestive of FAP. The 20-year-old patient underwent a colonoscopy on Feb 21, 2019. More than ten polyps in size of 0.4-4cm with erosions on some of them were discovered. Three big polyps on the hepatic flexure of colon were biopsied which indicated high-grade intraepithelial neoplasia. The patient subsequently received a subtotal colectomy on March 25, 2019 and was diagnosed with stage T3N1bM0 colonic adenocarcinomas accompanied with more than one hundred of tubulovillous adenomas with the larger ones measuring 0.3-1cm. Immunohistochemical tests were performed with the surgical sample and revealed a status of HER2 (0), proficient DNA mismatch repair (pMMR) and BRAF (–). The patient subsequently received a modified oxaliplatin (L-OHP) with leucovorin (LV) and bolus/continuous infusion of 5-fluorouracil (5-FU) (mFOLFOX6) regimen for 12 cycles.

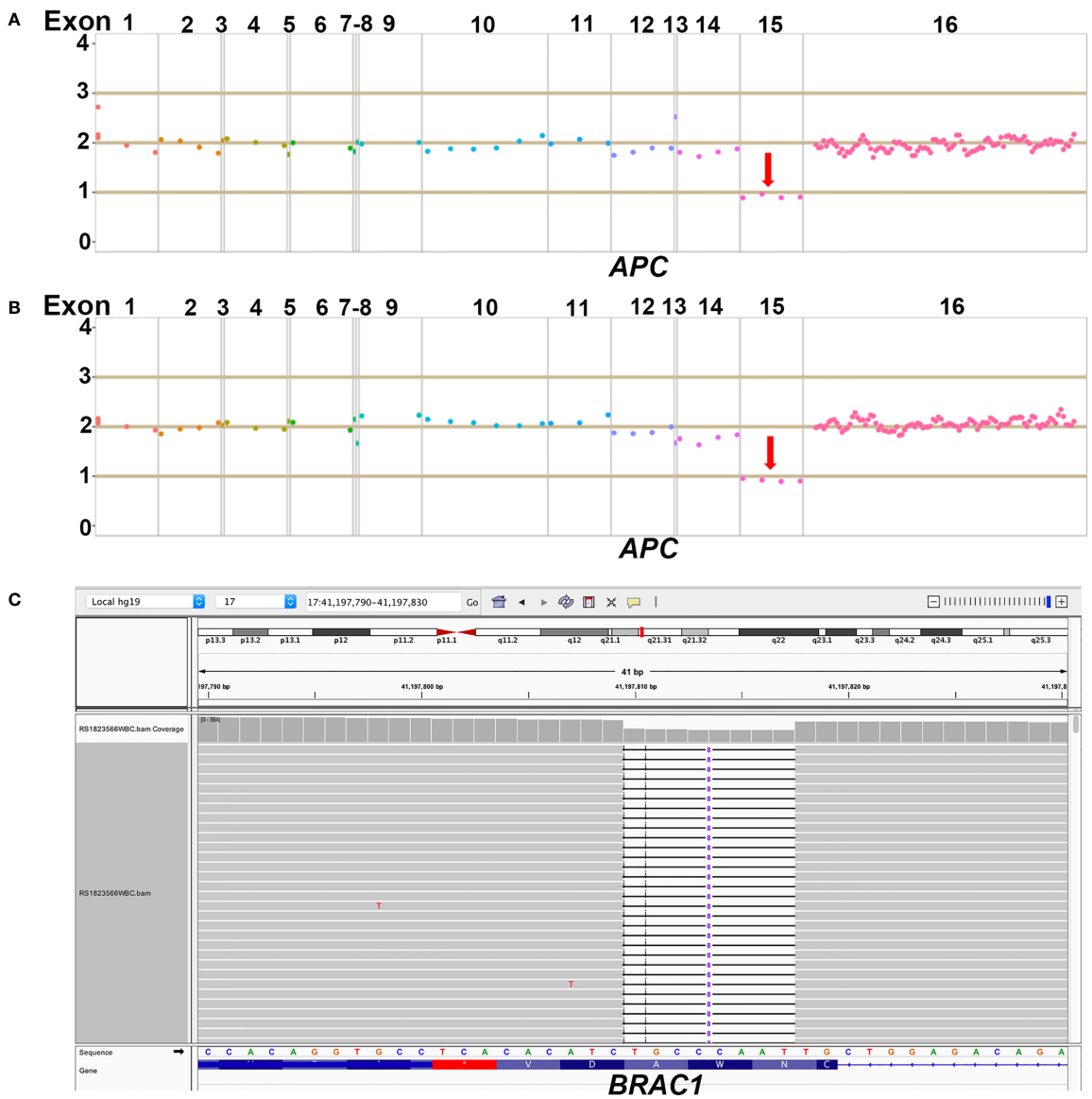


FIGURE 2 | Demonstration of NGS results of *APC* and *BRCA1* germline mutations. **(A)** The heterozygous loss of exon 15 in *APC* gene was detected in the proband; **(B)** The same heterozygous loss of exon 15 in *APC* gene was also detected in the 20-year old male patient; **(C)** The heterozygous p.Ile1824AspfsX3 in *BRCA1* gene was detected in the 20-year old male patient.

Repeated image tests and a colonoscopy in Dec, 2019 revealed no evidence of recurrence or metastasis but the presence of multiple polyps measuring 0.3-0.8cm on the rectum, which were subsequently diagnosed with adenomatous polyps. The patient remained alive as the submission of the manuscript. The sister of the patient (III1) was also diagnosed with multiple adenomatous polyps of colon in May 2019 when she was 27 years and underwent subtotal colectomy.

DISCUSSION

Although this is the second report of CRC with concurrent *APC* and *BRCA* germline mutations, it is the first case which the inheritance of both mutations was well characterized by comprehensively sequencing the family members. Dolkar et al. first reported a 44-year-old Caucasian male with concurrent *APC* and *BRCA* germline mutations who had colonic adenocarcinoma accompanied with 15 additional colon polyps (16). The patient's father had pancreatic cancer and his mother as well as his maternal cousin had colon cancer. Sequencing identified a pathogenic substitution mutation at nucleotide position 1213 in exon 9 of *APC* gene resulting in a premature stop codon (p.R405X) plus a deleterious c.8297delC variant in the *BRCA2* gene. The report only provided genetic tests for the proband therefore origins of mutations were not well recognized. In the present case, the male patient carried germline mutations in *APC* and *BRAC1* genes. By performing genetic analysis on multiple affected family members, we were able to delineate the inheritance pattern of mutations in this family. The mother of the patient (proband), sibling and maternal cousins all carried the *APC* mutation and his paternal aunt carried the *BRAC1* mutation which demonstrated a maternal origin of the former mutation and a paternal origin of the latter.

Pathogenic variants in *APC* gene are predominantly located in the exon 15 and always cause a premature truncation of the APC protein through nonsense substitutions or frameshifts (7). One of the mutations identified in this case was a heterozygous loss of exon 15 in *APC* gene which has been reported previously (17) and defined as a pathogenic variant according to the guidance of American College of Medical Genetics and Genomics (ACMG) (18).

BRCA proteins play essential roles in repair of DNA double-strand breaks *via* a homologous recombination mechanism (10). Deficiencies in BRCA proteins cause chromosomal instability which is associated with tumorigenesis. The open reading frame shift mutation in *BRAC1* (p.Ile1824AspfsX3) reported in the case resulted in a truncated BRCA1 protein of 1825 amino acids. The mutation has been detected in 3 out of 133 Chinese women with familial breast/ovarian cancer and was characterized as pathogenic (19).

It is known that *BRCA1/2* are the two major susceptibility genes for breast and ovarian cancers. However it is still controversial whether the existence of germline mutation in *BRCA1/2* increases the risk of CRC. A study genotyped 2,398 CRC patients and 4,570 controls showed the presence of *BRCA1* mutation in 0.42% of cases and in 0.48% of controls ($P = 0.8$). Although the *BRCA1* mutation frequency was found slightly higher (0.93%) in patients with family CRC history, the study did not support the correlation of *BRCA1* mutations with increased risk of CRC (14). Another prospective study in 7,015 women with a *BRCA* mutation revealed an increased risk of CRC in carriers with *BRCA1* mutations younger than 50 years but not in carriers with *BRCA2* mutations or elder females (15).

Mutations in *APC* cause autosomal dominant FAP which often leads to CRC eventually. However, it is inconclusive whether *APC* and *BRCA* interact intrinsically which might predispose individuals with germline mutations in both genes to an increased risk of cancers. In the present case, the patient with double mutations developed colonic adenocarcinomas with hundreds of adenomas at the age of 20. He had a much earlier onset than other relatives in the family who only carried *APC* mutation. Previous studies investigating the coexistence of *APC* polymorphism I1307K with *BRCA* germline mutations demonstrated that *APC* I1307K increased the penetrance of *BRCA* mutations for breast cancer but not for ovarian cancer (20, 21). A study in mice also implied that although *APC* mutation might function early in the neoplastic process, coinheritance of a *BRCA2* alteration did not modify the *APC* mutation-driven phenotypes and therefore did not enhance tumorigenesis (22).

In present report, the well-specified family cancer history and the FAP-typical clinical characteristics observed in the proband made it relatively easy to uncover the disease-causing mutations. However in clinical practice, features of hereditary tumor syndrome are not always observed in patients and family histories are often not well-recognized. Therefore, the detection rate of pathogenic variants is always unsatisfactory. In a study aiming to screen the mutation rate in cancer susceptibility genes in 1,058 unselected CRC patients revealed 9.9% of patients carried mutations in cancer susceptibility genes and 7.0% carried mutations in non-Lynch syndrome (LS) genes. Notably, 15 of 23 carriers of high-penetrance non-LS mutations lacked classic clinical histories suggesting genetic factors that underlie CRC frequently occur beyond well-recognized familial CRC syndromes (23). Therefore, multigene panel testing in unselected patients with CRC will identify substantially more disease-causing mutations and bring more opportunities for genetically driven cancer prevention (3, 23).

The reported case harbors both *APC* and *BRAC1* pathogenic mutations which also confer an increased risk for duodenal cancer, pancreatic cancer, thyroid cancer, breast cancer and prostate cancer (7, 24). In order to detect tumor at early stage and receive proper management, the patient is suggested to consider routine esophagogastroduodenoscopy, thyroid and breast ultrasound imaging, pancreatic and prostate CT scan annually.

In conclusion, we reported a familial CRC case with coinheritance of mutations in both *APC* and *BRCA1* and well-characterized inheritance pattern in the family. The patient benefited from colonoscopy and subsequent management on the basis of genetic testing results. The case illustrates the importance of evaluating detailed family history and

performing a multiple-gene panel test in cancer patient allowing for the identification of more disease-causing mutations and bringing more opportunities for genetically driven cancer prevention. In addition, one should also be aware of the scenario where double disease-causing mutations of different hereditary tumor syndromes coexist, in which the predisposition to specific cancer needs further investigation.

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

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

LW: conception and design. WH: manuscript writing and manuscript review. WH and JB: clinical management of the patient. LS and HL: gene sequencing. XQ, LZ, and LW: manuscript revision. All authors contributed to the article and approved the submitted version.

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. Yurgelun MB, Allen B, Kaldate RR, Bowles KR, Judkins T, Kaushik P, et al. Identification of a Variety of Mutations in Cancer Predisposition Genes in Patients With Suspected Lynch Syndrome. *Gastroenterology* (2015) 149 (3):604–13.e20. doi: 10.1053/j.gastro.2015.05.006
3. Pearlman R, Frankel WL, Swanson B, Zhao W, Yilmaz A, Miller K, et al. Prevalence and Spectrum of Germline Cancer Susceptibility Gene Mutations Among Patients With Early-Onset Colorectal Cancer. *JAMA Oncol* (2017) 3 (4):464–71. doi: 10.1001/jamaoncol.2016.5194
4. Leoz ML, Carballal S, Moreira L, Ocana T, Balaguer F. The genetic basis of familial adenomatous polyposis and its implications for clinical practice and risk management. *Appl Clin Genet* (2015) 8:95–107. doi: 10.2147/TACG.S51484
5. Aitchison A, Hakkaart C, Day RC, Morrin HR, Frizelle FA, Keenan JI. APC Mutations Are Not Confined to Hotspot Regions in Early-Onset Colorectal Cancer. *Cancers (Basel)* (2020) 12(12):3829. doi: 10.3390/cancers12123829
6. Kinzler KW, Nilbert MC, Su LK, Vogelstein B, Bryan TM, Levy DB, et al. Identification of FAP locus genes from chromosome 5q21. *Science* (1991) 253 (5020):661–5. doi: 10.1126/science.1651562
7. Jasperson KW, Patel SG, Ahnen DJ. APC-Associated Polyposis Conditions. In: MP Adam, HH Ardinger, RA Pagon, SE Wallace, LJH Bean, K Stephens, editors. *GeneReviews((R))*. Seattle (WA): University of Washington, Seattle (1993).
8. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* (1994) 266(5182):66–71. doi: 10.1126/science.7545954
9. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature* (1995) 378(6559):789–92. doi: 10.1038/378789a0
10. Venkitaraman AR. Functions of BRCA1 and BRCA2 in the biological response to DNA damage. *J Cell Sci* (2001) 114(Pt 20):3591–8.
11. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. *J Clin Oncol* (2007) 25(11):1329–33. doi: 10.1200/JCO.2006.09.1066
12. Cullinane CM, Creavin B, O'Connell EP, Kelly L, O'Sullivan MJ, Corrigan MA, et al. Risk of colorectal cancer associated with BRCA1 and/or BRCA2 mutation carriers: systematic review and meta-analysis. *Br J Surg* (2020) 107 (8):951–9. doi: 10.1002/bjs.11603
13. Oh M, McBride A, Yun S, Bhattacharjee S, Slack M, Martin JR, et al. BRCA1 and BRCA2 Gene Mutations and Colorectal Cancer Risk: Systematic Review and Meta-analysis. *J Natl Cancer Inst* (2018) 110(11):1178–89. doi: 10.1093/jnci/djy148
14. Suchy J, Cybulski C, Gorski B, Huzarski T, Byrski T, Debniak T, et al. BRCA1 mutations and colorectal cancer in Poland. *Fam Cancer* (2010) 9(4):541–4. doi: 10.1007/s10689-010-9378-x
15. Phelan CM, Iqbal J, Lynch HT, Lubinski J, Gronwald J, Moller P, et al. Incidence of colorectal cancer in BRCA1 and BRCA2 mutation carriers: results from a follow-up study. *Br J Cancer* (2014) 110(2):530–4. doi: 10.1038/bjc.2013.741
16. Dolkar T, Zikria J, Bussell S. Coexisting germline mutations in APC and BRCA2 in a patient with colon cancer. *CAP Today* (2015) 29(6):62–6.
17. Aretz S, Stienen D, Uhlhaas S, Pagenstecher C, Mangold E, Caspari R, et al. Large submicroscopic genomic APC deletions are a common cause of typical familial adenomatous polyposis. *J Med Genet* (2005) 42(2):185–92. doi: 10.1136/jmg.2004.022822
18. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* (2015) 17(5):405–24. doi: 10.1038/gim.2015.30
19. Cao WM, Gao Y, Yang HJ, Xie SN, Ding XW, Pan ZW, et al. Novel germline mutations and unclassified variants of BRCA1 and BRCA2 genes in Chinese women with familial breast/ovarian cancer. *BMC Cancer* (2016) 16:64. doi: 10.1186/s12885-016-2107-6
20. Maresco DL, Arnold PH, Sonoda Y, Federici MG, Bogomolniy F, Rhei E, et al. The APC I1307K allele and BRCA-associated ovarian cancer risk. *Am J Hum Genet* (1999) 64(4):1228–30. doi: 10.1086/302345
21. Leshno A, Shapira S, Liberman E, Kraus S, Sror M, Harlap-Gat A, et al. The APC I1307K allele conveys a significant increased risk for cancer. *Int J Cancer* (2016) 138(6):1361–7. doi: 10.1002/ijc.29876
22. Bennett LM, McAllister KA, Ward T, Malphurs J, Collins NK, Seely JC, et al. Mammary tumor induction and premature ovarian failure in ApcMin mice are not enhanced by Brca2 deficiency. *Toxicol Pathol* (2001) 29(1):117–25. doi: 10.1080/019262301301418928
23. Yurgelun MB, Kulke MH, Fuchs CS, Allen BA, Uno H, Hornick JL, et al. Cancer Susceptibility Gene Mutations in Individuals With Colorectal Cancer. *J Clin Oncol* (2017) 35(10):1086–95. doi: 10.1200/JCO.2016.71.0012
24. Petrucelli N, Daly MB, Pal T. BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer. In: MP Adam, HH Ardinger, RA Pagon, SE Wallace, LJH Bean, K Stephens, editors. *GeneReviews((R))*. Seattle (WA): University of Washington, Seattle (1993).

Conflict of Interest: Author LS, HL and LZ were employed by company Burning Rock Biotech.

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.

frontiers in Oncology

ORIGINAL RESEARCH
published: 06 April 2021
doi: 10.3389/fonc.2021.657665

Apigenin Inhibits the Growth of Hepatocellular Carcinoma Cells by Affecting the Expression of microRNA Transcriptome

Shou-Mei Wang[1†], Pei-Wei Yang[1†], Xiao-Jun Feng[1†], Yi-Wei Zhu[2], Feng-Jun Qiu[2], Xu-Dong Hu[2] and Shu-Hui Zhang[1*]*

[1] Department of Pathology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Chinese Medicine, Shanghai, China, [2] Department of Biology, School of Basic Medical Sciences, Shanghai University of Traditional Chinese Medicine, Shanghai, China

OPEN ACCESS

Edited by:
Chengqian Yin,
Shenzhen Bay Laboratory, China

Reviewed by:
Bo Zhu,
Dana-Farber Cancer Institute, United States
Tongzheng Liu,
Jinan University, China

***Correspondence:**
Xu-Dong Hu
huxudongsh@126.com
Shu-Hui Zhang
shzhang@126.com

[†]These authors share first authorship

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *23 January 2021*
Accepted: *12 March 2021*
Published: *06 April 2021*

Citation:
Wang S-M, Yang P-W, Feng X-J, Zhu Y-W, Qiu F-J, Hu X-D and Zhang S-H (2021) Apigenin Inhibits the Growth of Hepatocellular Carcinoma Cells by Affecting the Expression of microRNA Transcriptome. Front. Oncol. 11:657665. doi: 10.3389/fonc.2021.657665

Background: Apigenin, as a natural flavonoid, has low intrinsic toxicity and has potential pharmacological effects against hepatocellular carcinoma (HCC). However, the molecular mechanisms involving microRNAs (miRNAs) and their target genes regulated by apigenin in the treatment of HCC have not been addressed.

Objective: In this study, the molecular mechanisms of apigenin involved in the prevention and treatment of HCC were explored *in vivo* and *in vitro* using miRNA transcriptomic sequencing to determine the basis for the clinical applications of apigenin in the treatment of HCC.

Methods: The effects of apigenin on the proliferation, cell cycle progression, apoptosis, and invasion of human hepatoma cell line Huh7 and Hep3B were studied *in vitro*, and the effects on the tumorigenicity of Huh7 cells were assessed *in vivo*. Then, a differential expression analysis of miRNAs regulated by apigenin in Huh7 cells was performed using next-generation RNA sequencing and further validated by qRT-PCR. The potential genes targeted by the differentially expressed miRNAs were identified using a curated miRTarBase miRNA database and their molecular functions were predicted using Gene Ontology and KEGG signaling pathway analysis.

Results: Compared with the control treatment group, apigenin significantly inhibited Huh7 cell proliferation, cell cycle, colony formation, and cell invasion in a concentration-dependent manner. Moreover, apigenin reduced tumor growth, promoted tumor cell necrosis, reduced the expression of Ki67, and increased the expression of Bax and Bcl-2 in the xenograft tumors of Huh7 cells. Bioinformatics analysis of the miRNA transcriptome showed that hsa-miR-24, hsa-miR-6769b-3p, hsa-miR-6836-3p, hsa-miR-199a-3p, hsa-miR-663a, hsa-miR-4739, hsa-miR-6892-3p, hsa-miR-7107-5p, hsa-miR-1273g-3p, hsa-miR-1343, and hsa-miR-6089 were the most significantly up-regulated miRNAs, and their key gene targets were MAPK1, PIK3CD, HRAS, CCND1, CDKN1A, E2F2, etc. The core regulatory pathways of the up-regulated miRNAs were associated with the

hepatocellular carcinoma pathway. The down-regulated miRNAs were hsa-miR-181a-5p and hsa-miR-148a-3p, and the key target genes were MAPK1, HRAS, STAT3, FOS, BCL2, SMAD2, PPP3CA, IFNG, MET, and VAV2, with the core regulatory pathways identified as proteoglycans in cancer pathway.

Conclusion: Apigenin can inhibit the growth of HCC cells, which may be mediated by up-regulation or down-regulation of miRNA molecules and their related target genes.

Keywords: apigenin, hepatocellular carcinoma, microRNA, transcriptome sequencing, pathway

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the main types of primary liver cancer. According to the latest statistics, more than 300,000 people die of HCC every year in China. It is the fifth largest cancer in the world and the second leading cause of cancer-related death (1, 2). Early clinical diagnosis and cure rate for HCC are low, and the rates of postoperative recurrence, metastasis, and mortality are extremely high. Even in developed countries, the relative survival rate of HCC in five years is only 7% (3). Due to the occult onset and high malignant degree of liver cancer, most patients with liver cancer are diagnosed at advanced stages, with treatment options mainly relying on chemotherapy and targeted therapy. Systemic chemotherapy with Western medicine, such as Adriamycin, Epirubicin, Fluorouracil, Cisplatin and Mitomycin, etc. has high general efficacy but also high toxicity, and the targeted therapeutic drugs are expensive. Therefore, a new and more affordable method is urgently needed to eradicate HCC.

Apigenin, a natural flavonoid, exists in *Scutellaria barbata*, *Lobelia chinensis*, *Oldenlandia diffusa*, Centipeda, Rhizoma Polygontum Cuspidatum, Veratrum Nigrum, Semen Plantaginis, Caulis Trachelospermi, and other Chinese herbal medicine. It has low intrinsic toxicity and has potential anti-oxidant, anti-inflammatory, anti-viral, and anti-cancer properties (4). Studies have demonstrated that apigenin can inhibit the proliferation of HCC cells, induce cell differentiation and apoptosis, inhibit cancer cell invasion and distant metastasis, inhibit angiogenesis, regulate immunity, enhance the sensitivity and reduce toxicity of chemotherapies (5–11). However, the molecular mechanism of apigenin in regulating the growth, invasion, and metastasis of HCC cells is still superficial and needs to be further investigated.

MiRNAs are highly conserved single-stranded non-coding RNAs that are involved in the regulation of various cellular activities, such as cell proliferation, apoptosis, differentiation, inflammation, migration, and invasion, and play important roles in tumorigenesis (12). It has been reported that apigenin has a good anti-HCC pharmacological effect, but the target miRNA molecules and related genes regulated by apigenin in the prevention and treatment of HCC are indeterminate. Therefore, in this study, the inhibitory effect of apigenin on human hepatoma cell line Huh7 and Hep3B was determined *in vitro* and *in vivo*, and the differential expression profile of miRNAs in Huh7 cells treated with apigenin was analyzed and screened using next-generation miRNA sequencing technology. Finally, we investigated the miRNAs and potential target genes of apigenin associated with the growth inhibition of HCC cells using ClueGo plug-in in Cytoscape-3.7.1 software, and miRTarBase database, and explored the molecular mechanism of apigenin in HCC.

MATERIALS AND METHODS

Reagent and Cell Line

Apigenin and human HCC cell line (Huh7 and Hep3B) were used in this study. The cell origins and specifications were described previously (13) and are provided in Supplementary Materials.

Cell Proliferation and Colony Formation Assays

Cell proliferation and colony formation assays were executed as previously described (13) and are provided in Supplementary Materials. Each experiment was repeated three times in duplicates.

Cell Cycle and Apoptosis Analysis

Huh7 and Hep3B cells were collected for cell cycle or apoptosis analysis according to the manufacturer's instructions, as described previously (14) and supplied in Supplementary Materials.

Transwell Invasion Assay

Cell culture inserts (8 μm pore size; Corning-Costar, USA) and Matrigel invasion chambers (Corning-Costar, USA) were used according to the manufacturer's instructions. The transwell invasion assay was performed as previously described (13).

Tumorigenicity in Nude Mice

The assay of tumorigenicity in nude mice as previously described (15) and in Supplementary Materials. Male BALB/c nude mice (weighing 18 - 20 g) were purchased from Shanghai Lingchang Biotechnology Co. Ltd. (Certificate No. 20180003007216) and were maintained under specific pathogen-free conditions. All experiments were performed according to the guidelines of the Committee on Protection, Welfare and Ethics of Experimental Animals in Yueyang Hospital of Integrated Traditional Chinese and Western Medicine affiliated to Shanghai University of Traditional Chinese Medicine (no. YYLAC-2019-036-4-2).

Immunohistochemistry and Tunel assay

Immunohistochemistry and Tunel assay were performed according to the manufacturer's instructions and as previously described (16) and in Supplementary Materials. The primary antibodies were supplied in **Supplementary Table 1**. Tunel apoptosis assay kit (Catalog No. C1098) was purchased from Beyotime Institute of Biotechnology (Jiangsu, China).

MicroRNA Transcriptome Expression Analysis

The method was provided in Supplementary Materials. The criteria for differentially expressed miRNAs and mRNAs were set at cutoff value of >2.0 fold change, and a threshold of RPM >10 for miRNA and FPKM >0.3 for mRNA based on the ratio between the treatment and the control. MiRNA experiments were performed by the Shanghai YunXu Bio-tech Company, Shanghai, China.

RNA Isolation and Quantitative Real-Time PCR (qRT-PCR) Analysis

Quantitative real-time PCR (qRT-PCR) was performed according to the manufacturer's instructions and as described previously (17). Primer sequences were shown in **Supplementary Table 2**. All samples were run in triplicate, and the relative miRNA and mRNA expression levels were calculated according to the $2^{-\triangle\triangle Ct}$ method (14).

Bioinformatics Analysis for the Characterization of miRNA and Related Target Genes

SPSS 24.0 software was used for K-means clustering analysis. The standardized reads of differentially expressed miRNAs between apigenin and control groups were clustered by the Heatmap.2 function in R language. The total number of comparative reads was used to standardize the comparative splicing reads (junction reads) for each sample with $\log_2$ conversion. The differentially expressed miRNAs across the experimental groups were selected based on a $\log_2$ fold-change ($\log_2$ FC) > 2.0 and P-value < 0.001. The target genes of differentially expressed miRNAs were predicted based on CLUEGO plug-in in Cytoscape software (version 3.7.1) and miRTarBase database. Gene ontology (Go) annotations, including the terms "biological process", "molecular function", and "cellular component," and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis were then carried out for the identified target genes through by DAVID tools (https://david.ncifcrf.gov/) (18).

Statistical Analysis

All experiments were performed in triplicate and the results are presented as mean ± standard deviation (SD). Statistical analysis between the groups was performed by Student's t-test using SPSS 24.0 software (SPSS, Armonk, NY, USA). A value of $P < 0.05$ was considered statistically significant.

RESULTS

Apigenin Inhibits Huh7 and Hep3B Cell Growth

To observe the effect of apigenin on the proliferation of hepatoma cells, we treated Huh7 and Hep3B cells with various concentrations (5, 10, and 20 μM) of apigenin for 3 days. We found that apigenin inhibited HCC cell proliferation in a dose-dependent manner ($P < 0.05$) (**Figures 1A, B**). To investigate whether apigenin affected the ability of Huh7 and Hep3B cells to survive and form colonies, the same number of viable cells treated with apigenin or DMSO were seeded at a low cell density on the petri dish. After 14 days of culturing, the colonies were visualized and counted microscopically. As shown in **Figure 1C**, compared to the control group, apigenin significantly reduced the number and size of colonies formed in a concentration-dependent manner ($P < 0.01$).

Apigenin Affects Cell Cycle Arrest and Apoptosis

To determine whether apigenin inhibits the growth of hepatoma cells by affecting cell cycle progression (G0/G1, S, and G2/M) and cellular apoptosis, we treated Huh7 and Hep3B cells with different concentrations of apigenin (5 μM, 10 μM, and 20 μM) for 48 h and investigated the cell cycle distribution by flow cytometry. The results showed that the number of G2/M phase cells increased and G0/G1 phase cells decreased significantly in Huh7 cells while the number of G0/G1 phase cells increased and G2/M phase cells decreased significantly in Hep3B cells ($P < 0.05$) (**Figure 1D**), suggesting that the inhibition of cell growth by apigenin is associated with cell cycle arrest. PI and annexin V staining was used to detect the effect of apigenin on cell apoptosis. The results displayed that the rate of apoptosis of Huh7 and Hep3B cells increased significantly after treatment with apigenin at different concentrations for 48 h ($P < 0.01$) (**Figure 1E**). This indicates that apigenin induces apoptosis in a dose-dependent manner. Then, apoptosis was detected by Tunel assay and the results displayed that the rate of apoptosis of tumor tissues from mice treated with apigenin increased significantly after treatment with apigenin ($P < 0.01$) (**Figure 2D**).

Apigenin Restrains the Invasion of Huh7 and Hep3B Cells

Cell invasion assay provides a quantitative approach to study the invasion and metastasis of tumor cells induced by various cytokines. In this study, a transwell invasion experiment was performed to observe the effect of apigenin on the invasion ability of Huh7 and Hep3B cells. When observed under the microscope, the cells stained with crystal violet dye appeared rounded in shape. The invasion cells were more numerous in the control group but were sparse in the apigenin treatment group. Statistical analysis showed that compared with the control group, the invasion number of Huh7 and Hep3B cells treated with apigenin (5 μM, 10 μM, and 20 μM) was significantly reduced ($P < 0.01$) (**Figure 1F** and **Supplementary Table 3**). It is,

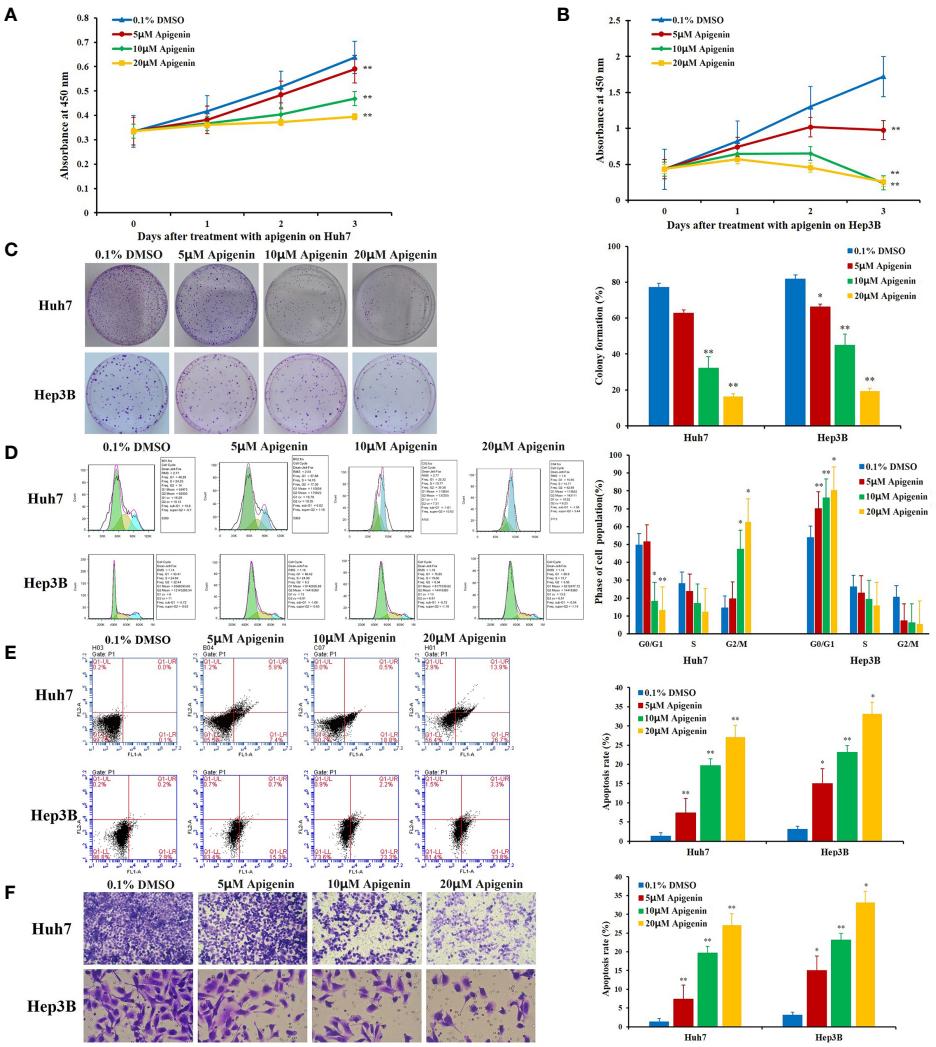


FIGURE 1 | Effects of apigenin on the growth, colony formation, cell cycle progression, and apoptosis rate of Huh7 and Hep3B cells *in intro*. **(A, B)** Changes in cell viability as determined by the CCK-8 assay. **(C)** Colony formation. **(D)** Cell cycle distribution. **(E)** Apoptosis analysis by Annexin-V/PI staining. **(F)** Transwell invasion assay. *$P < 0.05$, **$P < 0.01$, indicating significant differences in comparison to the control treatment (0.1% DMSO) group. Data are shown as mean ± SD from three independent experiments.

therefore, suggested that apigenin is effective in inhibiting the invasion of Huh7 and Hep3B cells and is negatively correlated with the concentration of apigenin.

Apigenin Suppresses Tumorigenicity of Huh7 Cells in Nude Mice

Next, we determined the effect of apigenin on tumorigenicity of Huh7 cells *in vivo*. In accordance with the findings *in vitro*, intraperitoneal injection of apigenin (25 mg/kg/day) suppressed tumor growth, showing a significant reduction in tumor volume and weight in comparison to the control treatment (0.1% DMSO) group ($P < 0.05$; **Figures 2A–C**). Microscopically, the tumor cells were arranged in diffuse compact trabeculae, with variable degrees of anaplasia and increased mitotic activity in the control group. In contrast, obvious necrosis was detected in the apigenin group. Immunohistochemical analysis of paraffin-embedded sections demonstrated increased Bax and decreased Bcl2 and Ki67 staining after apigenin treatment (**Figure 2D**).

Apigenin Influenced miRNA Differential Expression as Assessed by Transcriptome Sequencing Analysis

Under high-throughput RNA-sequencing (RNA-seq), the differentially expressed miRNAs in apigenin-treated Huh7 hepatocellular carcinoma cells were identified, of which 130 up-regulated and 9 down-regulated in apigenin treated cells (**Supplementary Table 4**). Further, we based on $\log_2$ fold-change ($\log_2$ FC) ≥ 2.0 and P value ≤ 0.001 selection criteria, a total of 32 miRNAs were obtained, of which 30 were up-regulated and 2 were down-regulated in apigenin treated cells

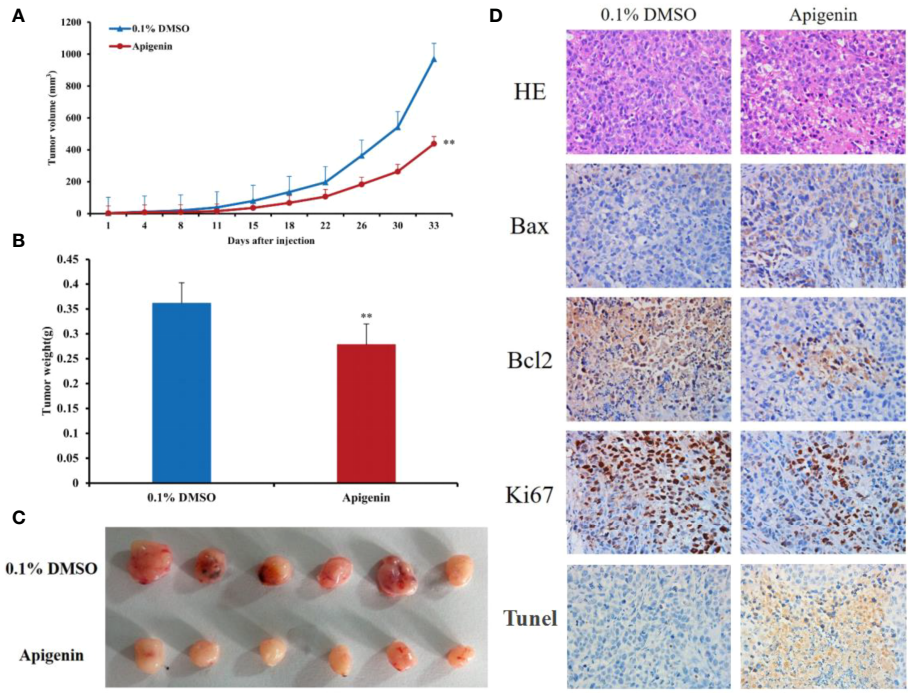


FIGURE 2 | Effects of apigenin on the tumorigenicity of Huh7 xenograft in nude mice. **(A)** Tumor growth curves of Huh7 cells treated with apigenin (25 mg/kg/day) or vehicle control (0.1% DMSO). **(B)** Tumor masses of each treatment group. **(C)** Photographs of resected tumors (n = 6) from each treatment group. **(D)** Representative tumor sections stained with H&E, anti-Bax, anti-Bcl2, anti-Ki67 antibodies and Tunel assay kit. Original magnification: ×400. $^{**}P < 0.01$ in comparison to the control treatment (0.1% DMSO) group.

(**Supplementary Tables 5**, **6**). The Heatmap.2 function of theR software package was used to cluster the differentially expressed miRNAs based on their expression profiles, and the results are shown as a hierarchical clustering heat map. As shown in **Supplementary Figure 1**, 139 differentially expressed miRNAs between the apigenin and control treatment groups could be effectively distinguished. We further made the heat map of 32 top apigenin-modulated miRNAs (**Figure 3**). Next, the miRNA target mRNA networks were constructed for the 30 upregulated and 2 down-regulated miRNAs using Cytoscape-3.7.1 software with the ClueGo plug-in and miRTarBase database (**Supplementary Figures S2, S3**).

Confirmation of the Differentially Expressed miRNAs and Target Genes by qRT-PCR

To verify the RNA-seq data, qRT-PCR analysis was conducted for 12 selected differentially expressed miRNAs. The qRT-PCR and RNA-seq expression results were consistent, although there were no distinct variations statistically in the expression of some genes. The results of the differential expression analysis of miRNAs displayed that the expression of hsa-miR-7847-3p, hsa-miR-663a, hsa-miR-1273g-3p, hsa-miR-619-5p, hsa-miR-34a-5p, hsa-miR-5787, and hsa-let-7i-5p levels were increased in apigenin-treated cells as compared with the control cells. Conversely, hsa-miR-1260b, hsa-miR-760, hsa-miR-215-3p, hsa-miR-181a-5p, and hsa-miR-148a-3p expression was markedly down-regulated in the cells treated with apigenin (**Figure 4A**). In the differential gene expression level, only CCND1 gene were markedly up-regulated in apigenin-treated cells as compared with the control cells, the remaining 6 genes including MAPK1, PIK3R5, CCND1, ADCY1, ADCY3, GNAQ and EGF were down-regulated in the cells treated with apigenin as compared with the control group. However, there was no significant difference in MAPK1, ADCY3, and GNAQ gene (**Figure 4B**).

Effect of Apigenin on Differentially Expressed miRNAs and Related Target Genes by KEGG Pathway Analysis

We further analyzed the enrichment of the KEGG pathways through the KEGG database. Then, we were able to infer the main pathways regulated by the differentially expressed miRNAs and their key target genes through KEGG pathway analysis.

KEGG pathway enrichment showed that the effect of apigenin on the up-regulated miRNAs in Huh7 cells was associated with multiple signaling pathways and target genes (**Figure 5A**). These miRNAs were involved in a variety of cancer-related signaling pathways, such as hepatocellular carcinoma, pathways in cancer, cellular senescence, bladder cancer, proteoglycans in cancer,

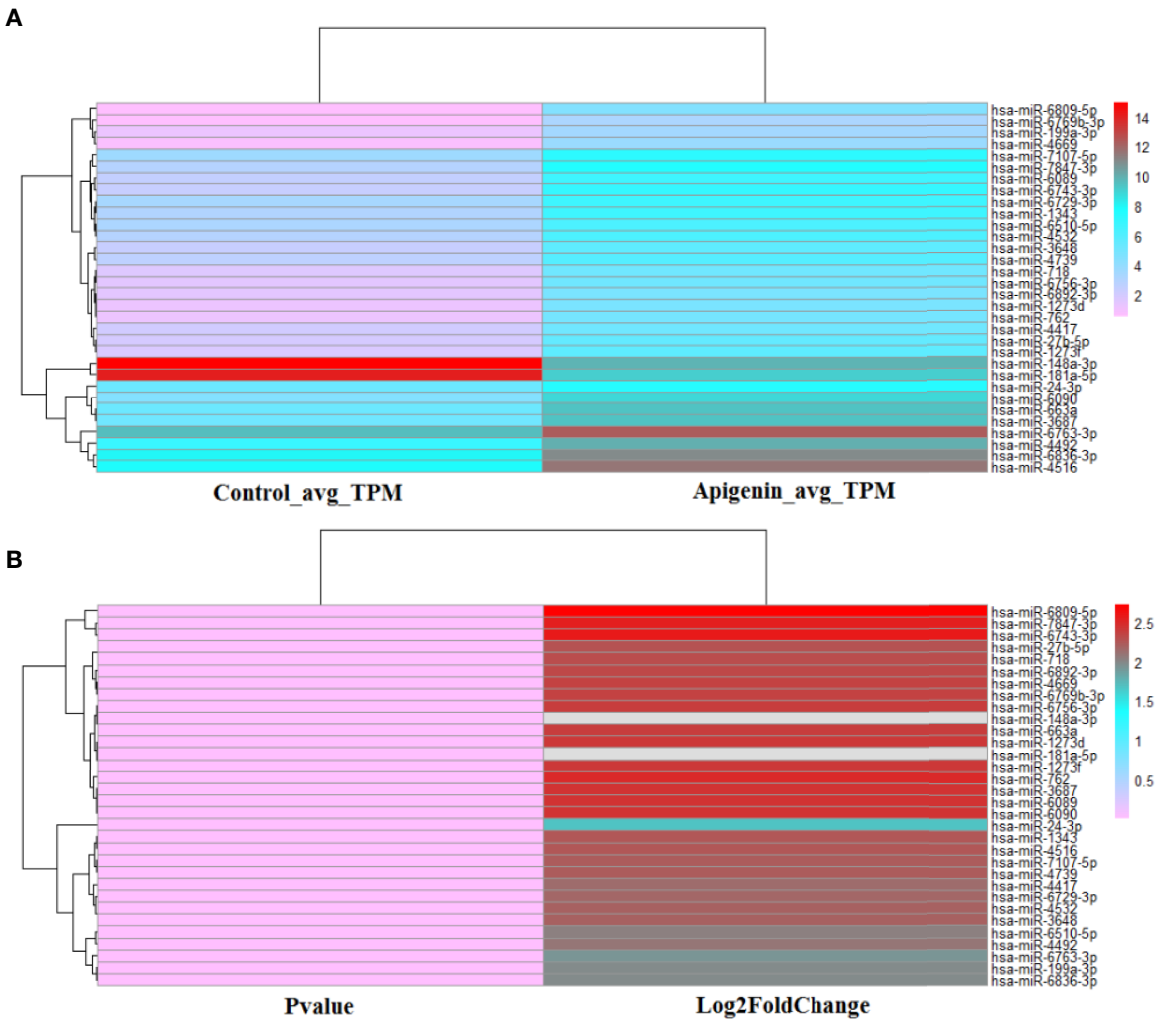


FIGURE 3 | The heat map of 32 top apigenin-modulated miRNAs in Huh 7 cells. **(A)** The hierarchical clustering analysis was performed to analyze the average TPM expression of miRNAs in Huh 7 cells treated with 10 μM apigenin vs. 0.1% DMSO. **(B)** The hierarchical clustering analysis was performed to analyze the P value and Log2Fold of the 32 top apigenin-modulated miRNAs (P value ≤ 0.001). Red indicates high relative expression and pink indicates low relative expression.

gastric cancer, melanoma, microRNAs in cancer, breast cancer, glioma, chronic myeloid leukemia, pancreatic cancer, FoxO signaling pathway, human T-cell leukemia virus 1 infection, human cytomegalovirus infection, colorectal cancer, endometrial cancer, small cell lung cancer, prostate cancer, PI3K-Akt signaling pathway, AGE-RAGE signaling pathway in diabetic complications, hepatitis B, non-small cell lung cancer, Kaposi sarcoma-associated herpesvirus infection, acute myeloid leukemia, Epstein-Barr virus infection, viral carcinogenesis, hepatitis C, human papillomavirus infection, and the erbB signaling pathway (**Supplementary Table 7**). The hepatocellular carcinoma pathway accounted for 75.93%, FoxO signaling pathway for 20.37% (**Figure 5B**). Furthermore, MAPK1, E2F2, CDK4, CDKN1A, CCND1 were found key target genes of hepatocellular carcinoma pathway. MAPK1, CDKN1A, and CCND1 are key target genes of FoxO signaling pathway(**Figure 5E**). Meanwhile, apigenin up-regulated miRNAs and related target genes in Huh7 cells were arranged in descending order of Degree value. According to a Degree value > mean 9.7, the key target genes included MAPK1, PIK3CD, HRAS, CCND1, CDKN1A, E2F2, MYC, CDK4, MTOR, CDK2, STAT3, PTEN, CDKN1, BIGF1, MAPK14, TGFB1, CDKN2A, VEGFA, WNT4, CCNA2, MET, VAV3, APC2, and APC (**Supplementary Table 8**).

Similarly, the effect of apigenin on the down-regulated miRNAs in Huh7 cells was also related to a variety of signaling pathways and target genes (**Figure 5C**). These signaling pathways consisted of microRNAs in cancer, proteoglycans in cancer, cellular senescence, FoxO signaling pathway, PD-L1 expression and PD-1 checkpoint pathway in cancer, focal adhesion, Th17 cell differentiation, human T-cell leukemia virus 1 infection, natural killer cell mediated cytotoxicity,

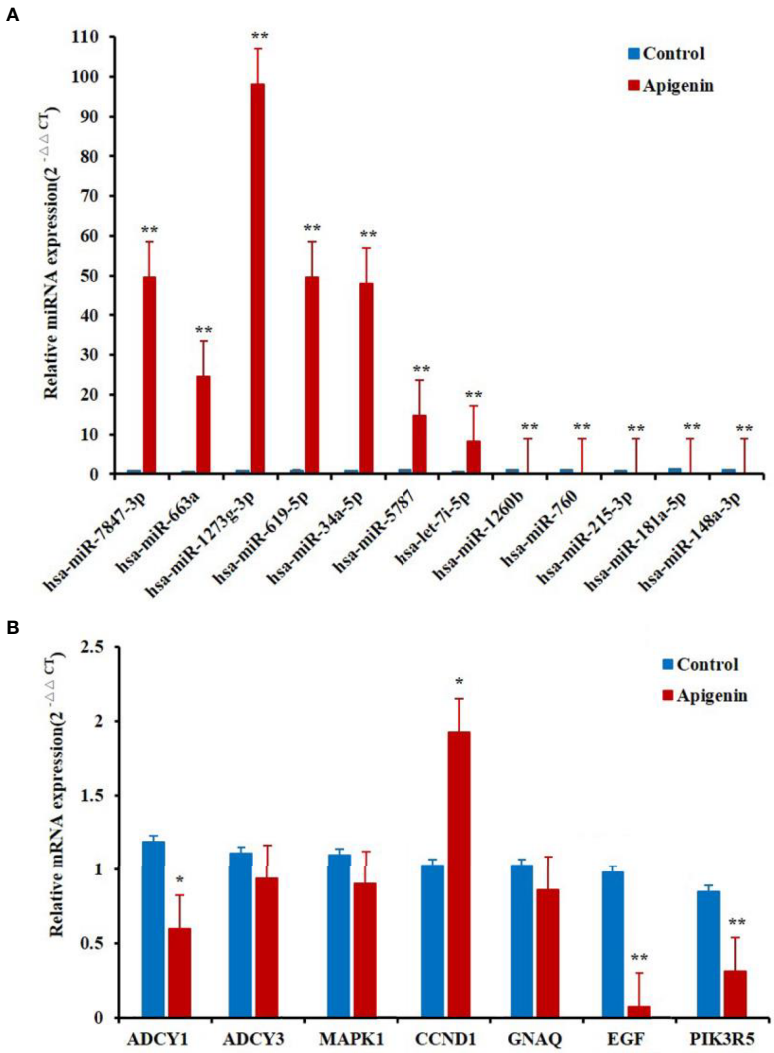


FIGURE 4 | Differentially expressed miRNAs and related target genes between apigenin-treated cells and the control group by RT-PCR. **(A)** The verification of the RNA-seq data by qPCR analysis of 12 miRNAs. **(B)** The notarization of the RNA-seq data by qPCR analysis of seven genes. *$P < 0.05$ and **$P < 0.01$, compared with the control (0.1% DMSO) group. miRNAs and genes expression by RT-PCR.

colorectal cancer, apoptosis, gastric cancer, AGE-RAGE signaling pathway in diabetic complications, T cell receptor signaling pathway, renal cell carcinoma, neurotrophin signaling pathway, Th1 and Th2 cell differentiation, Chagas disease (American trypanosomiasis), long-term potentiation, acute myeloid leukemia, yersinia infection, growth hormone synthesis, secretion and action, sphingolipid signaling pathway, Prolactin signaling pathway, melanoma, central carbon metabolism in cancer, Fc epsilon RI signaling pathway, non-small cell lung cancer, and VEGF signaling pathway (**Supplementary Table 9**). Proteoglycans in cancer accounted for 67.27%, microRNAs in cancer for 12.73% (**Figure 5D**). According to a Degree value > mean 7.2, the key target genes were MAPK1, HRAS, STAT3, FOS, BCL2, SMAD2, PPP3CA, IFNG, MET, and VAV2 (**Supplementary Table 8**).

Effect of Apigenin on the Expression miRNAs and Related Target Genes by Go functional Analysis

To understand the biological significance of the identified miRNAs and their target genes, functional analysis was performed using the Go database. Go enrichment analysis mainly included three parts: biological process (BP), molecular function (MF), and cell component (CC). We were able to annotate and infer the functions of the differentially expressed miRNAs by analyzing the Go functions, with a false discovery rate (FDR) ≤ 0.05 indicating biological significance. Compared with the control group, 88 Go-BP terms of miRNAs target genes up-regulated in the apigenin treatment group were mainly involved in Cell cycle regulation accounted for 75.0% of all biological functions, regulation of transferase activity for

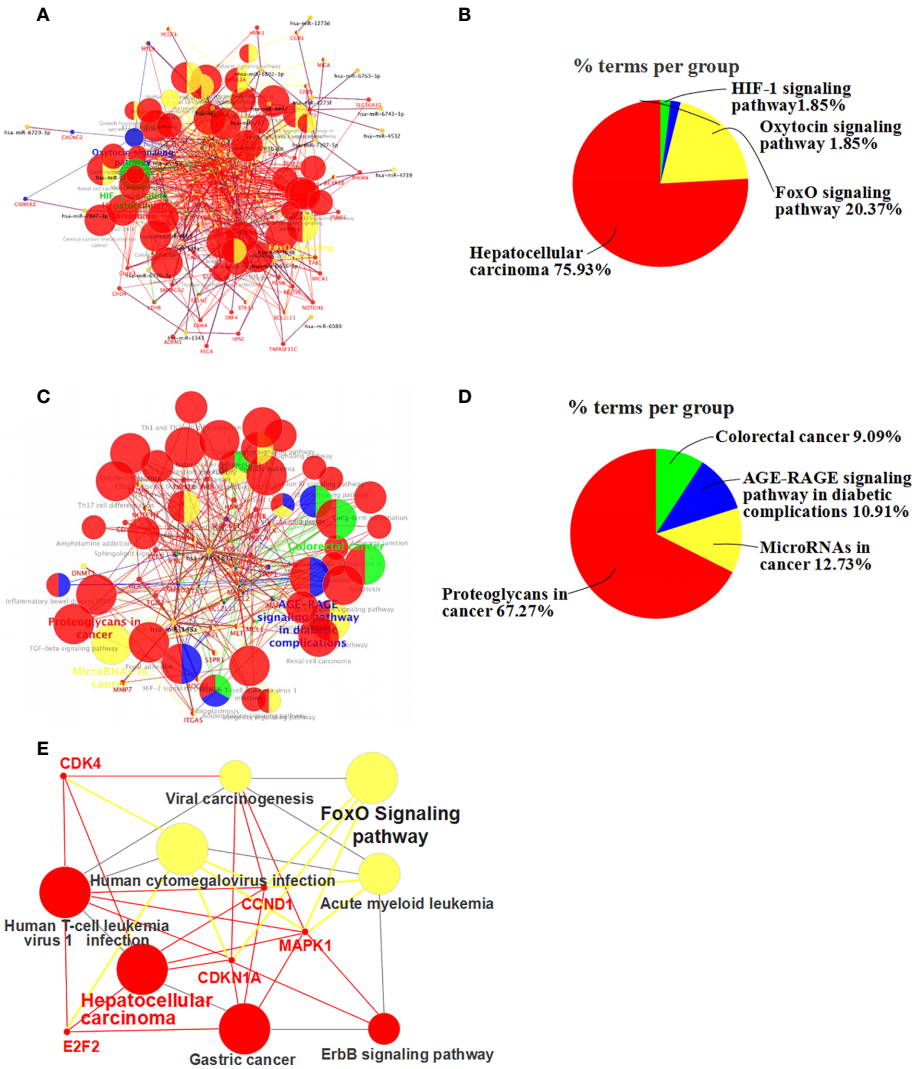


FIGURE 5 | KEGG pathway enrichment analysis of differentially expressed miRNA-related target genes. **(A)** Up-regulated miRNA and associated target gene signaling pathways (up-regulated miRNAs-pathway-genes) in apigenin-treated Huh7 cells. **(B)** The percentage of the up-regulated miRNA-related signaling pathways. **(C)** Down-regulated miRNA and target gene signaling pathways in apigenin-treated Huh7 cells (down-regulated miRNAs-pathway-genes). **(D)** The percentage of the down-regulated miRNA-related signaling pathways. **(E)** Up-regulated miRNA-FoxO signaling pathways-genes in apigenin-treated Huh7 cells.

14.42%, regulation of kinase activity for 3.85%, regulation of growth for 2.88%, apoptotic process for 1.92%, regulation of cellular response to stress for 0.96% and response to inorganic substance for 0.96% (**Figure 6A**). Compared with the control group, 18 Go-MF terms of the miRNA target genes were up-regulated in the apigenin group and included histone kinase activity accounted for 44.44%, purine nucleotide transmembrane transporter activity for 16.67%, RNA polymerase II specific DNA-binding transcription factor binding for 16.67%, protein serine/threonine kinase activity for 11.11%, and myristoyl-CoA hydrolase activity and magnesium ion binding for 5.56% (**Figure 6B**). Compared with the control group, 18 Go-CC terms of miRNA target genes up-regulated in the apigenin group were mainly involved in cyclin-dependent protein kinase holoenzyme complex accounted for 36.36%, cyclin B1-CDK1 complex for 18.18%, chromosome region for 13.64%, cell-substrate junction for 9.09%, adherens junction for 9.09%, gamma-tubulin complex, and plasma membrane raft and transport vesicle for 4.55% (**Figure 6C**).

Compared with the control group, 82 Go-BP terms of miRNA target genes down-regulated in the apigenin group were primarily associated with endothelial cell differentiation accounted for 27.68% of all biological functions, regulation of cellular senescence for 16.96%, regulation of anoikis for 16.07%, negative regulation of anoikis for 11.61%, intrinsic apoptotic signaling pathway responding to DNA damage for 6.25%, and positive regulation of insulin receptor signaling pathway for 5.36% (**Figure 6D**). Only 2 Go-MF terms of the miRNA target

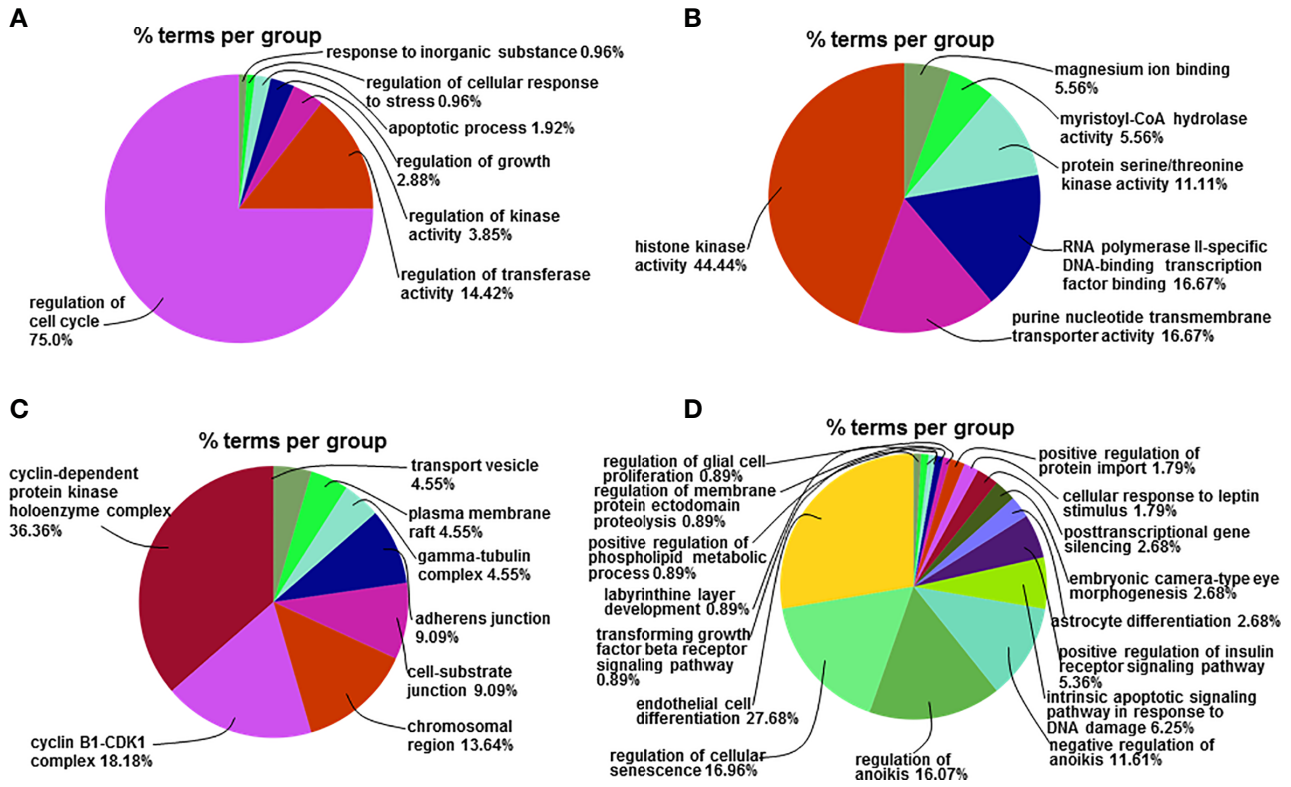


FIGURE 6 | Go functional analysis of the differentially expressed miRNA-related target genes. **(A)** The percentage of the up-regulated mRNAs-related GO-BP terms enriched in the corresponding group in total up-regulated mRNAs-related GO-BP terms. **(B)** The percentage of the up-regulated mRNAs-related GO-MF terms enriched in the corresponding group in total up-regulated mRNAs-related GO-MF terms. **(C)** The percentage of the up-regulated mRNAs-related GO-CC terms enriched in the corresponding group in total up-regulated mRNAs-related GO-CC terms. **(D)** The percentage of the down-regulated mRNAs-related GO-BP terms enriched in the corresponding group in total down-regulated mRNAs-related GO-BP terms.

genes were down-regulated in the apigenin group with phosphotyrosine residue binding accounted for 100% of the Go terms enriched. There were no Go-CC terms found for the down-regulated miRNA target genes in the apigenin group.

DISCUSSION

In this study, we demonstrated that apigenin inhibits the growth and invasion of HCC cells both *in vitro* and *in vivo*. *In vitro*, this effect showed that apigenin reduced HCC cell proliferation activity, induced cell cycle arrest and apoptosis, and suppressed the invasion of Huh7 and Hep3B cells. However, the effect of apigenin on cell cycle was completely different in Huh7 and Hep3B cells, suggesting that apigenin may have different effects on cell cycle regulation in different cell lines. Other studies have also confirmed that apigenin can induce the arrest of HepG2 cells in the G2/M phase and induce apoptosis (19).After the treatment of Bel-7402/adriamycin (ADM) cells, apigenin can induce cell arrest in the S phase (20). In addition, apigenin plays a role in other cancers by inducing apoptosis and cell cycle regulation (21, 22). *In vivo*, the inhibitory effect was manifested as apigenin halting the tumorigenicity of Huh7 cells in nude mice by decreasing the tumor volume, weight and increasing apoptosis, and impacting the expression of Bax, Bcl-2, and Ki-67. Therefore, apigenin has antitumor effects, which are consistent with the previous data demonstrated in various cancers, especially HCC (17).

Under high-throughput RNA-sequencing, we showed that apigenin-treated Huh7 cells were affected at the miRNA expression level. Apigenin up-regulated miRNAs were involved mainly in the hepatocellular carcinoma-related pathway and the FoxO signaling pathway, indicating that most of the differentially expressed miRNAs regulated by apigenin are related to the occurrence and development of HCC. Overall, the cell cycle regulation made up a large percentage of the Go terms. The differential expression of hsa-miR-199a-3p, hsa-miR-663a, and hsa-miR-24 in apigenin-treated Huh7 cells was mainly related to the proliferation and invasion of hepatoma cells. MiR-199a/b-5p was reported to inhibit the activation of the ROCK1/MLC and PI3K/Akt signaling pathways by negatively regulating ROCK1 expression, leading to the inhibition of liver cancer metastasis (23). MiR-663a/b can inhibit the proliferation and invasion of HCC cells by regulating TGF-β 1 and target gene GAB2 (24, 25). MiR-24 is known to promote HCC cell growth, metastasis, and invasion by targeting P53 or metallothionein 1M (26, 27), while our results

demonstrated that apigenin up-regulated the expression of hsa-miR-24 and inhibited the growth of Huh7 hepatoma cells. In addition, up-regulation of miRNAs, such as hsa-miR-6769b-3p, hsa-miR-6836-3p, hsa-miR-4739, hsa-miR-6892-3p, hsa-miR-7107-5p, hsa-miR-1273g-3p, hsa-miR-1343, and hsa-miR-6089, by apigenin have not been reported previously.

The down-regulated miRNAs in apigenin-treated Huh7 cells were found to be involved mainly in proteoglycans in the cancer-related pathway and microRNAs in the cancer-related pathway. Of the identified Go terms, the differentiation of endothelial cells and apoptosis make up a large percentage. The down-regulation of hsa-miR-181a-5p and hsa-miR-148a-3p in apigenin-treated Huh7 cells was likely associated with the apoptosis and invasion of tumor cells. The deletion of the miR-181 family is known to inhibit the migration of tumor cells through the regulation of MAPK (28). In addition, miR-148b is found to suppress the proliferation, migration and invasion of HepG2 and SMMC 7721 cells by targeting Rho-associated protein kinase 1 (29). In this study, apigenin down-regulated the expression of hsa-miR-148a-3p and inhibited the proliferation of Huh7 hepatoma cells, which is inconsistent with the report in the literature. This discrepancy may be due to the different roles hsa-miR-148a-3p plays in different hepatoma cells. Although down-regulation of hsa-miR-148a-3p could promote cell proliferation and invasion, apigenin appears to affect multiple miRNAs at the same time, which results in the inhibition of the proliferation and invasion of hepatoma cells.

The miRNAs up-regulated in Huh7 cells during apigenin treatment were mainly linked to the suppression of MAPK1, PIK3CD, HRAS, CCND1, CDKN1A, E2F2, MYC, CDK4, MTOR, CDK2, STAT3, PTEN, CDKN1, BIGF1, MAPK14, TGFB1, CDKN2A, VEGFA, WNT4, CCNA2, MET, VAV3, APC2, and APC expression. BIGF1 is a newly discovered gene in HCC during this study. Among the down-regulated genes, CCND1 was found to be down-regulated by HOTAIR gene knockout Huh7 cells, which inhibited the proliferation and induced cell cycle arrest in Huh7 cells (30). Apigenin inhibited HCC growth by down-regulating CDK4 and up-regulating CyclinD1 *via* p38 MAPK-p21 signaling in Huh7, SMMC-7721 and HepG2 cell lines (31). Down-regulation of CDK2, cyclin A, cyclin B1 and cyclin E contributed to the anti-proliferation effect of apigenin treatment in human bladder cancer T-24 cells (32). In addition, MTOR, STAT3, HRAS and c-MYC as oncogenes, together with the transcription factor E2F2, play important roles in the proliferation, differentiation, apoptosis and invasion of HCC cells (9, 33, 34). However, in disagreement with our findings, an increase in CDKN1A expression was found to promote cell apoptosis (35). Overall, the predicted results of apigenin on miRNA target genes are consistent with the reported genes described in the literature. MYC, CDKN2A, PTEN, HARS, APC2, and APC are often referred to as miRNAs in cancer. CDKN2A and PTEN genes are involved in the p53 signaling pathway. WNT4, HGF and TGFB1 are associated with the molecular signaling pathways in cancer, proteoglycans in cancer, and the relaxin signaling pathway. CCND1, CDK4, MAPK1, CDKN1A, and E2F2 genes are linked to the development of hepatocellular carcinoma. At present, limited studies on PIK3CD, HRAS, E2F2, BIGF1, WNT4, and VAV3 genes have been reported in hepatoma cells, indicating the need for further research.

The down-regulated miRNAs in apigenin treated Huh7 cells showed anti-HCC pharmacological effects by increasing the expression of MAPK1, HRAS, STAT3, FOS, BCL2, SMAD2, PPP3CA, IFNG, MET, and VAV2. Of which, SMAD2, PPP3CA, MAPK1, and FOS genes are known to be involved in the regulation of tumor cell proliferation and invasion. Activation of TGF-β1/Smad2 signaling can promote epithelial-to-mesenchymal transition (EMT) and invasion of HCC (36). PPP3CA was down-regulated in the chip analysis of HCC patients with hepatitis C (37). The expression of MAPK1 inhibits the proliferation of Huh7 and Hep-G2 hepatoma cells (38). De-repression of c-Fos gene expression caused by miR-139 down-regulation contributes to MHCC97H cell metastasis (39). Inhibiting the expression of VEGF, VAV2, and CDC42 contributed to the suppression of angiogenesis and metastasis of HCC (40). PPP3CA is related to cell senescence, VEGF signaling pathway, etc. IFNG and VAV2 are associated with HIF-1 signaling pathway, TGF-β signaling pathway, etc. These genes have been rarely reported in liver cancer, which suggests the need for further study. In addition, we found that key target genes such as MAPK1 were included in the up-regulated and down-regulated miRNA networks, but our experimental results showed that there was no significant difference in MAPK1 after apigenin treatment, which may be caused by the mutual canceling of the up-regulated and down-regulated effects.

In summary, our study demonstrated that apigenin has an inhibitory effect in hepatoma cells, which is associated with anti-proliferation, induction of cell cycle arrest and apoptosis, and the suppression of HCC cell invasion. In addition, as a promising drug, apigenin has a variety of properties including obvious antioxidant and anti-inflammatory activity (41, 42). It is important to note that the effects of apigenin vary with different doses, and multiple studies have shown that 20-100 μM apigenin inhibits cell proliferation, invasion, and apoptosis induction (43, 44). It has also been suggested that high doses of apigenin may cause oxidative stress-induced liver damage (41). Therefore, in the application of the grasp of the dose is very important. This study is the first to utilize the whole genome expression profiles to identify apigenin-regulated miRNAs in HCC cells. These differentially expressed miRNAs may be involved in the specific molecular mechanism of the inhibitory activity of apigenin in hepatoma cells.

DATA AVAILABILITY STATEMENT

The authors acknowledge that the data presented in this study must be deposited and made publicly available in an acceptable repository, prior to publication. Frontiers cannot accept an article that does not adhere to our open data policies.

ETHICS STATEMENT

The animal study was reviewed and approved by Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Chinese Medicine, Shanghai, China.

AUTHOR CONTRIBUTIONS

X-DH and S-HZ designed the study. S-MW, P-WY, and X-JF contributed equally to this research as co-first authors. F-JQ and Y-WZ contributed to animal rearing. All authors contributed to the article and approved the submitted version.

FUNDING

This study is supported by the National Natural Science Foundation of China (grant number: 82003994, 81172311); Highland Project from Integrated Traditional Chinese and Western medicine Branch of Shanghai University of Traditional Chinese Medicine; the special research project of traditional Chinese medicine in Henan Province (No. 2019JDZX2034).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.657665/full#supplementary-material

REFERENCES

1. European Association for the Study of the Liver and Electronic address eee and European Association for the Study of the L. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol* (2018) 69 (1):182–236. doi: 10.1016/j.jhep.2018.03.019
2. Global Burden of Disease Liver Cancer C, Akinyemiju T, Abera S, Ahmed M, Alam N, Alemayohu MA, et al. The Burden of Primary Liver Cancer and Underlying Etiologies From 1990 to 2015 at the Global, Regional, and National Level: Results From the Global Burden of Disease Study 2015. *JAMA Oncol* (2017) 3(12):1683–91. doi: 10.1001/jamaoncol.2017.3055
3. Sayiner M, Golabi P, Younossi ZM. Disease Burden of Hepatocellular Carcinoma: A Global Perspective. *Dig Dis Sci* (2019) 64(4):910–7. doi: 10.1007/s10620-019-05537-2
4. Madunic J, Madunic IV, Gajski G, Popic J, Garaj-Vrhovac V. Apigenin: A dietary flavonoid with diverse anticancer properties. *Cancer Lett* (2018) 413:11–22. doi: 10.1016/j.canlet.2017.10.041
5. Bhattacharya S, Mondal L, Mukherjee B, Dutta L, Ehsan I, Debnath MC, et al. Apigenin loaded nanoparticle delayed development of hepatocellular carcinoma in rats. *Nanomedicine* (2018) 14(6):1905–17. doi: 10.1016/j.nano.2018.05.011
6. Coelho PL, Oliveira MN, da Silva AB, Pitanga BP, Silva VD, Faria GP, et al. The flavonoid apigenin from Croton betulaster Mull inhibits proliferation, induces differentiation and regulates the inflammatory profile of glioma cells. *Anticancer Drugs* (2016) 27(10):960–9. doi: 10.1097/CAD.0000000000000413
7. Kang CH, Molagoda IMN, Choi YH, Park C, Moon DO, Kim GY. Apigenin promotes TRAIL-mediated apoptosis regardless of ROS generation. *Food Chem Toxicol* (2018) 111:623–30. doi: 10.1016/j.fct.2017.12.018
8. Seydi E, Rasekh HR, Salimi A, Mohsenifar Z, Pourahmad J. Selective Toxicity of Apigenin on Cancerous Hepatocytes by Directly Targeting their Mitochondria. *Anticancer Agents Med Chem* (2016) 16(12):1576–86. doi: 10.2174/1871520616666160425110839
9. Cao HH, Chu JH, Kwan HY, Su T, Yu H, Cheng CY, et al. Inhibition of the STAT3 signaling pathway contributes to apigenin-mediated anti-metastatic effect in melanoma. *Sci Rep* (2016) 6:21731. doi: 10.1038/srep21731
10. Hu XY, Liang JY, Guo XJ, Liu L, Guo YB. 5-Fluorouracil combined with apigenin enhances anticancer activity through mitochondrial membrane potential (DeltaPsim)-mediated apoptosis in hepatocellular carcinoma. *Clin Exp Pharmacol Physiol* (2015) 42(2):146–53. doi: 10.1111/1440-1681.12333
11. Gao AM, Zhang XY, Hu JN, Ke ZP. Apigenin sensitizes hepatocellular carcinoma cells to doxorubic through regulating miR-520b/ATG7 axis. *Chem Biol Interact* (2018) 280:45–50. doi: 10.1016/j.cbi.2017.11.020
12. Giordano S, Columbano A. MicroRNAs: new tools for diagnosis, prognosis, and therapy in hepatocellular carcinoma? *Hepatology* (2013) 57(2):840–7. doi: 10.1002/hep.26095
13. Song WH, Feng XJ, Gong SJ, Chen JM, Wang SM, Xing DJ, et al. microRNA-622 acts as a tumor suppressor in hepatocellular carcinoma. *Cancer Biol Ther* (2015) 16(12):1754–63. doi: 10.1080/15384047.2015.1095402
14. Feng XJ, Pan Q, Wang SM, Pan YC, Wang Q, Zhang HH, et al. MAP4K4 promotes epithelial-mesenchymal transition and metastasis in hepatocellular carcinoma. *Tumour Biol* (2016) 37(8):11457–67. doi: 10.1007/s13277-016-5022-1
15. Yang P-W, Lu Z-Y, Pan Q, Chen TT, Feng XJ, Wang SM, et al. MicroRNA-6809-5p mediates luteolin-induced anticancer effects against hepatoma by targeting flotillin 1. *Phytomedicine* (2019) 57:18–29. doi: 10.1016/j.phymed.2018.10.027
16. Liu AW, Cai J, Zhao XL, Jiang TH, He TF, Fu HQ, et al. ShRNA-targeted MAP4K4 inhibits hepatocellular carcinoma growth. *Clin Cancer Res* (2011) 17 (4):710–20. doi: 10.1158/1078-0432.CCR-10-0331
17. Cai J, Zhao XL, Liu AW, Nian H, Zhang SH. Apigenin inhibits hepatoma cell growth through alteration of gene expression patterns. *Phytomedicine* (2011) 18(5):366–73. doi: 10.1016/j.phymed.2010.08.006
18. Chen J, Liu C, Cen J, Liang T, Xue J, Zeng H, et al. KEGG-expressed genes and pathways in triple negative breast cancer: Protocol for a systematic review and data mining. *Medicine (Baltimore)* (2020) 99(18):e19986. doi: 10.1097/MD.0000000000019986
19. Khan TH, Sultana S. Apigenin induces apoptosis in HepG2 cells: Possible role of TNF-α and IFN-γ. *Toxicology* (2006) 217:206–12. doi: 10.1016/j.tox.2005.09.019
20. Gao AM, Ke ZP, Wang JN, Yang JY, Chen SY, Chen H. Apigenin sensitizes doxorubicin-resistant hepatocellular carcinoma BEL-7402/ADM cells to doxorubicin via inhibiting PI3K/Akt/Nrf2 pathway. *Carcinogenesis* (2013) 34(8):1806–14. doi: 10.1093/carcin/bgt108
21. Kaur P, Shukla S, Gupta S. Plant flavonoid apigenin inactivates Akt to trigger apoptosis in human prostate cancer: an in vitro and in vivo study. *Carcinogenesis* (2008) 29(11):2210–7. doi: 10.1093/carcin/bgn201
22. Meng S, Zhu Y, Li JF, Wang X, Liang Z, Li SQ, et al. Apigenin inhibits renal cell carcinoma cell proliferation. *Oncotarget* (2017) 8(12):19834–42. doi: 10.18632/oncotarget.15771
23. Zhan Y, Zheng N, Teng F, Bao L, Liu F, Zhang M, et al. MiR-199a/b-5p inhibits hepatocellular carcinoma progression by post-transcriptionally suppressing ROCK1. *Oncotarget* (2017) 8(40):67169–80. doi: 10.18632/oncotarget.18052
24. Zhang C, Chen B, Jiao A, Li F, Sun N, Zhang G, et al. miR-663a inhibits tumor growth and invasion by regulating TGF-beta1 in hepatocellular carcinoma. *BMC Cancer* (2018) 18(1):1179. doi: 10.1186/s12885-018-5016-z
25. Guo L, Li B, Miao M, Yang J, Ji J. MicroRNA663b targets GAB2 to restrict cell proliferation and invasion in hepatocellular carcinoma. *Mol Med Rep* (2019) 19(4):2913–20. doi: 10.3892/mmr.2019.9934
26. Chen L, Luo L, Chen W, Xu HX, Chen F, Chen LZ, et al. MicroRNA-24 increases hepatocellular carcinoma cell metastasis and invasion by targeting p53: miR-24 targeted p53. *BioMed Pharmacother* (2016) 84:1113–8. doi: 10.1016/j.biopha.2016.10.051
27. Dong X, Ding W, Ye J, Yan D, Xue F, Xu L, et al. MiR-24-3p enhances cell growth in hepatocellular carcinoma by targeting metallothionein 1M. *Cell Biochem Funct* (2016) 34(7):491–6. doi: 10.1002/cbf.3213

28. Song MK, Park YK, Ryu JC. Polycyclic aromatic hydrocarbon (PAH)-mediated upregulation of hepatic microRNA-181 family promotes cancer cell migration by targeting MAPK phosphatase-5, regulating the activation of p38 MAPK. *Toxicol Appl Pharmacol* (2013) 273(1):130–9. doi: 10.1016/j.taap.2013.08.016
29. Chen X, Bo L, Lu W, Zhou G, Chen Q. MicroRNA-148b targets Rho-associated protein kinase 1 to inhibit cell proliferation, migration and invasion in hepatocellular carcinoma. *Mol Med Rep* (2016) 13(1):477–82. doi: 10.3892/mmr.2015.4500
30. Zhou JJ, Cheng D, He XY, Meng Z, Li WZ, Chen RF. Knockdown of Hotair suppresses proliferation and cell cycle progression in hepatocellular carcinoma cell by downregulating CCND1 expression. *Mol Med Rep* (2017) 16(4):4980–6. doi: 10.3892/mmr.2017.7162
31. Li Y, Cheng X, Chen C, Huijuan W, Zhao H, Liu W, et al. Apigenin, a flavonoid constituent derived from P. villosa, inhibits hepatocellular carcinoma cell growth by CyclinD1/CDK4 regulation via p38 MAPK-p21 signaling. *Pathol Res Pract* (2020) 216(1):152701. doi: 10.1016/j.prp.2019.152701
32. Shi MD, Shiao CK, Lee YC, Shih YW. Apigenin, a dietary flavonoid, inhibits proliferation of human bladder cancer T-24 cells via blocking cell cycle progression and inducing apoptosis. *Cancer Cell Int* (2015) 15:33. doi: 10.1186/s12935-015-0186-0
33. Yang J, Pi C, Wang G. Inhibition of PI3K/Akt/mTOR pathway by apigenin induces apoptosis and autophagy in hepatocellular carcinoma cells. *BioMed Pharmacother* (2018) 103:699–707. doi: 10.1016/j.biopha.2018.04.072
34. Fang ZQ, Li MC, Zhang YQ, Liu XG. MiR-490-5p inhibits the metastasis of hepatocellular carcinoma by down-regulating E2F2 and ECT2. *J Cell Biochem* (2018) 119(10):8317–24. doi: 10.1002/jcb.26876
35. Couture R, Mora N, Al Bittar S, Najih M, Touaibia M, Martin LJ. Luteolin modulates gene expression related to steroidogenesis, apoptosis, and stress response in rat LC540 tumor Leydig cells. *Cell Biol Toxicol* (2020) 36(1):31–49. doi: 10.1007/s10565-019-09481-9
36. Yang Y, Liu Q, Li Z, Zhang R, Jia C, Yang Z, et al. GP73 promotes epithelial-mesenchymal transition and invasion partly by activating TGF-β1/Smad2 signaling in hepatocellular carcinoma. *Carcinogenesis* (2018) 39(7):900–10. doi: 10.1093/carcin/bgy010
37. Zekri AR, Hassan ZK, Bahnassy AA, Sherif GM, ELdahshan D, Abouelhoda M, et al. Molecular prognostic profile of Egyptian HCC cases infected with hepatitis C virus. *Asian Pac J Cancer Prev* (2012) 13(11):5433–8. doi: 10.7314/APJCP.2012.13.11.5433
38. Fu X, Zhang J, He X, Yan X, Wei J, Huang M, et al. Circular RNA MAN2B2 promotes cell proliferation of hepatocellular carcinoma cells via the miRNA-217/MAPK1 axis. *J Cancer* (2020) 11(11):3318–26. doi: 10.7150/jca.36500
39. Fan Q, He M, Deng X, Wu WK, Zhao L, Tang J, et al. Derepression of c-Fos caused by microRNA-139 down-regulation contributes to the metastasis of human hepatocellular carcinoma. *Cell Biochem Funct* (2013) 31(4):319–24. doi: 10.1002/cbf.2902
40. Wang R, Zhao N, Li S, Fang JH, Chen MX, Yang J, et al. MicroRNA-195 suppresses angiogenesis and metastasis of hepatocellular carcinoma by inhibiting the expression of VEGF, VAV2, and CDC42. *Hepatology* (2013) 58(2):642–53. doi: 10.1002/hep.26373
41. Singh P, Mishra SK, Noel S, Sharma S, Rath SK. Acute exposure of apigenin induces hepatotoxicity in Swiss mice. *PloS One* 2012 7(2):e31964. doi: 10.1371/journal.pone.0031964
42. Middleton E, Drzewiecki G. Flavonoid inhibition of human basophil histamine release stimulated by various agents. *Biochem Pharmacol* (1984) 33(21):3333–8. doi: 10.1016/0006-2952(84)90102-3
43. Liu X, Li L, Lv L, Chen D, Shen L, Xie Z. Apigenin inhibits the proliferation and invasion of osteosarcoma cells by suppressing the Wnt/β-catenin signaling pathway. *Oncol Rep* (2015) 34(2):1035–41. doi: 10.3892/or.2015.4022
44. Masuelli L, Benvenuto M, Mattera R, Di Stefano E, Zago E, Taffera G, et al. In Vitro and In Vivo Anti-tumoral Effects of the Flavonoid Apigenin in Malignant Mesothelioma. *Front Pharmacol* (2017) 8:373. doi: 10.3389/fphar.2017.00373

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: 06 April 2021
doi: 10.3389/fonc.2021.621422

Next-Generation Sequencing Reveals High Uncommon EGFR Mutations and Tumour Mutation Burden in a Subgroup of Lung Cancer Patients

Gang Guo[1†], Gaofeng Li[1†], Yinqiang Liu[2†], Heng Li[1], Qi Guo[1], Jun Liu[3], Xiumei Yang[1], Tao Shou[4] and Yunfei Shi[2*]*

[1] Department of Thoracic Surgery, Yunnan Cancer Hospital, Kunming, China, [2] Department of Thoracic Surgery, First Affiliated Hospital of Kunming Medical University, Kunming, China, [3] Department of Thoracic Surgery, First People's Hospital of Yunnan Province, Kunming, China, [4] Department of Medical Oncology, First People's Hospital of Yunnan Province, Kunming, China

OPEN ACCESS

Edited by:
Qiyuan Li,
Xiamen University, China

Reviewed by:
Haiying Cheng,
Albert Einstein College of Medicine, United States
Jie Meng,
Central South University, China

***Correspondence:**
Tao Shou
yn_shoutao@hotmail.com
Yunfei Shi
km-syf@163.com

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

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 26 October 2020
Accepted: 15 February 2021
Published: 06 April 2021

Citation:
Guo G, Li G, Liu Y, Li H, Guo Q, Liu J, Yang X, Shou T and Shi Y (2021) Next-Generation Sequencing Reveals High Uncommon EGFR Mutations and Tumour Mutation Burden in a Subgroup of Lung Cancer Patients. Front. Oncol. 11:621422. doi: 10.3389/fonc.2021.621422

Xuanwei County in Southwest China shows the highest incidence and mortality rate of lung cancer in China. Although studies have reported distinct clinical characteristics of patients from Xuanwei, the molecular features of these patients with non-small cell lung cancer (NSCLC) remain unclear. Here, we comprehensively characterised such cases using next-generation sequencing (NGS). Formalin-fixed, paraffin-embedded tumour samples from 146 patients from Xuanwei with NSCLC were collected for an NGS-based target panel assay; their features were compared with those of reference Chinese and The Cancer Genome Atlas (TCGA) cohorts. Uncommon *EGFR* mutations, defined as mutations other than L858R, exon 19del, exon 20ins, and T790M, were the predominant type of *EGFR* mutations in the Xuanwei cohort. Patients harbouring uncommon *EGFR* mutations were more likely to have a family history of cancer ($p = 0.048$). A higher frequency of *KRAS* mutations and lower frequency of rearrangement alterations were observed in the Xuanwei cohort ($p < 0.001$). Patients from Xuanwei showed a significantly higher tumour mutation burden than the reference Chinese and TCGA cohorts ($p < 0.001$). Our data indicates that patients from Xuanwei with NSCLC harbouring G719X/S768I co-mutations may benefit from treatment with EGFR-tyrosine kinase inhibitors. Our comprehensive molecular profiling revealed unique genomic features of patients from Xuanwei with NSCLC, highlighting the potential for improvement in targeted therapy and immunotherapy.

Keywords: NSCLC, tumour mutation burden, uncommon EGFR mutations, Xuanwei county, NGS

INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths in China (1). Many patients (57%) are diagnosed with metastatic disease, leading to a 5-year relative survival rate of 5% (2). Approximately 85% of lung cancers are non-small cell lung cancers (NSCLCs), of which lung adenocarcinoma and lung squamous cell carcinoma are the most common subtypes (3). Compared with other regions in China, Xuanwei County in Yunnan Province has the highest mortality rate of lung cancer. The age-standardised mortality rates of lung cancer patients from Xuanwei were six and three times

higher than those of patients from rural areas of China, among females and males, respectively (4). Hospitals in Yunnan Province organise free CT examinations to ensure early detection of lung cancer, and many patients are diagnosed at stage I. Xuanwei is rich in smoky (bituminous) coal, which may be associated with the high mortality rate of lung cancer in this area. Retrospective studies have shown that a lifelong use of smoky coal is associated with a 36- and 99-fold increase in mortality in men and women, respectively, compared with smokeless coal use (5, 6). Notably, lung cancer in Xuanwei has some remarkable characteristics, such as higher incidence in non-smoking females, diagnosis at a younger age, rapid tumour progression, multiple lung lesions, poor overall prognosis, and family aggregation (7). Overall, lung cancer patients in Xuanwei may present a distinct subgroup globally, leading researchers to consider whether epidemiological and clinicopathological peculiarities can be interpreted based on genomic features. Recent studies suggested that the NSCLC cohort in Xuanwei harboured a significantly higher co-mutation rate in *EGFR* exons 18 and 20 (8). NSCLCs are often found to have a high tumour mutation burden (TMB), which has been associated with apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) signatures (9). However, the detailed characteristics of these *EGFR* mutations, comprehensive molecular profiling, and TMB characteristics of patients with NSCLC in Xuanwei are unclear.

We performed comprehensive genomic testing in an NSCLC cohort from Xuanwei. The genomic features of this cohort were compared with those of a reference Chinese NSCLC cohort (1,802 patients, excluding patients from Xuanwei) and data from The Cancer Genome Atlas (TCGA) mainly comprising a Western population from Europe and the US (10).

METHODS

Patient Enrolment

In total, 1948 Chinese patients diagnosed with NSCLC at the Yunnan Cancer Hospital, First People's Hospital of Yunnan Province, or First Affiliated Hospital of Kunming Medical University were recruited. Formalin-fixed paraffin-embedded tumour samples were collected between December 2017 and January 2019. Matched blood samples were collected as reference controls. Of these patients, 146 were from Xuanwei and defined as the Xuanwei cohort. The remaining 1,802 Chinese patients with NSCLC were defined as the reference Chinese cohort. This study was approved by the Institution Review Board of the First Hospital of Kunming Medical University and conducted according to the Declaration of Helsinki. Informed consent was obtained from all enrolled patients.

Abbreviations: APOBEC, apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like; COSMIC, Catalogue of Somatic Mutations in Cancer; Mb, megabase; NGS, next-generation sequencing; NSCLC, non-small cell lung cancer; RECIST, Response Criteria in Solid Tumours; TCGA, The Cancer Genome Atlas; TKI, tyrosine kinase inhibitor; TMB, tumour mutation burden; TMB-H, high TMB; TMB-L, low TMB; CT, computed tomography.

Next-Generation Sequencing (NGS)

All tumour tissues and matched blood samples underwent targeted NGS-based genomic testing (OrigiMed, Shanghai, China) in a College of American Pathologists-accredited and Clinical Laboratory Improvement Amendments-certified laboratory (11). Approximately 50 ng of cancer tissue DNA was extracted from 40 mm formalin-fixed paraffin-embedded tumour samples and blood samples using the DNA Extraction Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. All coding exons and selected introns of targeted genes were captured for hybridisation capture panel and then sequenced on an Illumina NextSeq-500 Platform (Illumina Incorporated, San Diego, CA). For formalin-fixed paraffin-embedded samples, sequencing depth was 900× mean coverage (minimum 700×); for matched blood samples, sequencing depth was 300×. Genomic alterations, including single nucleotide variants, short and long insertions/deletions, copy number variations, and gene rearrangements, were subjected to advanced analysis. TMB score was calculated from a 450-gene panel data (**Supplementary Table 1**) for each sample by counting the number of somatic mutations, including coding single nucleotide variants and insertions/deletions, per megabase (Mb) of the sequence examined. Known somatic mutations in the Catalogue of Somatic Mutations in Cancer (COSMIC; https://cancer.sanger.ac.uk/cosmic/signatures) and known germline polymorphisms in the U.S. National Centre for Biotechnology Information's Single Nucleotide Polymorphism Database were not counted (12). Particularly, 35 and 111 tumour samples were subjected to 37 and 450 cancer-related gene panel testing (**Supplementary Tables 1**, **2**), respectively. TMB analysis was available for 111 cases. A high TMB (TMB-H) was defined as $\geq$10 muts/Mb, and a low TMB (TMB-L) was defined as $<$10 muts/Mb. Mutational signature analysis was conducted using the deconstructSigs package v1.8.0. All the detected somatic mutations, including synonymous mutations in the cohort, were imported for signature analysis. Finally, the weights of 30 known cancer mutation signatures in COSMIC were generated (13, 14). Uncommon *EGFR* mutations were defined as mutations other than L858R, exon 19del, exon 20ins, and T790M (15).

Response Evaluation

All nine patients received oral EGFR-tyrosine kinase inhibitor (TKI) treatment. Radiological follow-up was performed first, after 1 month, and then, once every 2 months, via computed tomography of the thorax and upper abdomen. Response was assessed according to the Response Criteria in Solid Tumours (RECIST) 1.1 (16). Progression-free survival was defined as the interval from the date of initiation of EGFR-TKI therapy to the date of disease progression or death from any cause, whichever occurred first.

Statistical Analyses

Statistical analyses were performed using the R Statistical Software package (R Foundation for Statistical Computing, Vienna, Austria). To analyse differences in continuous variables and TMB, the Wilcoxon test was performed when comparing each two groups, and the Kruskal-Wallis test was performed

TABLE 1 | Clinicopathological baseline characteristics of patients from Xuanwei and reference Chinese patients with NSCLC.

Characteristics	Xuanwei cohort (*N* = 146)	Reference Chinese cohort (*N* = 1,802)	*p*-value
Gender [*N* (%)]			0.7914
Male	82 (56.2%)	985 (54.7%)	
Female	64 (43.8%)	817 (45.3%)	
Age (median years, range)	55 (36–78)	66 (22–92)	<0.001
Stage [*N* (%)]			<0.001
I	86 (58.8%)	632 (35.1%)	
II	14 (9.6%)	192 (10.7%)	
III	22 (15.1%)	356 (19.8%)	
IV	22 (15.1%)	620 (34.3%)	
Unknown	2 (1.4%)	2 (0.1%)	
Smoking history [*N* (%)]			0.1899
Yes	58 (39.7%)	550 (30.5%)	
Never	86 (58.9%)	1,046 (58.1%)	
Unknown	2 (1.4%)	206 (11.4%)	
Histology [*N* (%)]			0.008
Adenocarcinoma	131 (89.8%)	1,567 (87.0%)	
Squamous cell carcinoma	11 (7.5%)	225 (12.5%)	
Others	4 (2.7%)	10 (0.5%)	
Family history [*N* (%)]			<0.001
Yes (lung cancer family history)	61 (41.8%) (59, 96.7%)	479 (26.6%) 232 (48.4%)	
No	83 (56.8%)	1183 (65.6%)	
Unknown	2 (1.4%)	140 (7.8%)	
Lesion number [*N* (%)]			
1	49 (33.6%)	–	
≥2	91 (62.3%)	–	
Unknown	6 (4.1%)	–	

when comparing all three groups. The Chi-square or Fisher's exact tests was used for association of categorical variables. The threshold for statistical significance was set at $p < 0.05$. The significance associated with each symbol is as follows: $^{***}p < 0.001$; $^{**}p < 0.01$; and $^{*}p < 0.05$.

RESULTS

Patients

Clinicopathological data of Xuanwei and reference Chinese patients with NSCLC are summarised in **Table 1**. The median age of the Xuanwei cohort was lower than that of the reference Chinese cohort (55 vs. 66 years, $p < 0.001$). Patients from Xuanwei showed less incidence of squamous cell carcinoma than reference Chinese patients (7.5 vs. 12.5%, $p = 0.008$). According to the pathology and medical history following the American Journal of Critical Care Cancer Staging Manual, patients were classified based on their main clinical stages (I–IV). The Xuanwei cohort contained more stage I–II patients than the reference Chinese cohort (68.5 vs. 45.7%, $p = 0.0018$). Moreover, Xuanwei cases had a greater cancer-related family history than reference Chinese cases (41.8 vs. 26.6%, $p < 0.001$), with most showing a family history of lung cancer (96.7%).

Driver Genes in the Xuanwei NSCLC Cohort

The nine driver genes of NSCLC found in patients from Xuanwei are summarised in **Figure 1A**. Comparison of the mutational profile of driver genes in the three cohorts showed significant differences (**Figure 1B**). The *KRAS* mutation frequency in the Xuanwei cohort was significantly higher than that in reference Chinese groups (26.3 vs. 11.2%, $p < 0.001$). The frequency of rearrangement alterations identified in the Xuanwei cohort was lower than that in the reference Chinese cohort (2.7 vs. 11.9%, $p < 0.001$). Rearrangement in *ROS1* and *NTRK1/2/3* were not found in the Xuanwei cohort (**Figures 1B,C**). The most common *KRAS* mutation in the Xuanwei cohort was *KRAS* G12C (53.8%), followed by *KRAS* G12V (23.1%) (**Figure 1D**).

EGFR Mutation Profile of the Xuanwei Cohort

A higher mutation frequency of *EGFR* was observed in the Xuanwei NSCLC cohort than in TCGA cases (46.6% vs. 10.7%, $p < 0.001$), although this frequency was comparable to that in reference Chinese cases (46.6 vs. 50.3%, $p = 0.44$) (**Figures 1A,B**). Notably, comparison of *EGFR* mutation subtypes demonstrated that patients from Xuanwei, compared to reference Chinese and Western patients, harboured a striking mutation pattern of

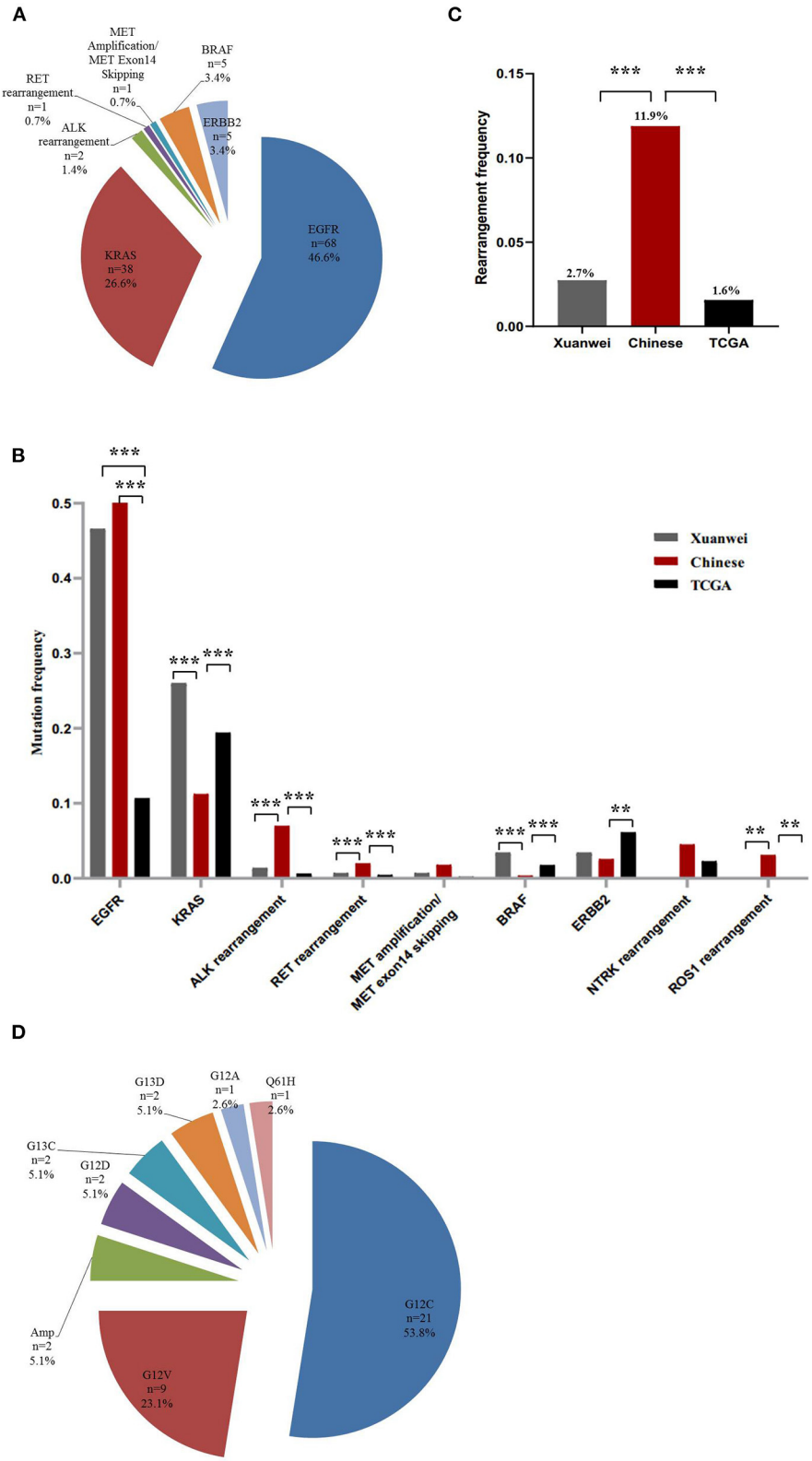


FIGURE 1 | Driver gene mutation profile in the Xuanwei NSCLC cohort. **(A)** Mutation frequency of nine genes in Xuanwei cohorts. **(B)** Composition of the alteration type in the nine genes among the three groups, namely the Xuanwei cohort (left column), reference Chinese cohort (middle column), and TCGA cohort (right column). **(C)** Composition of the rearrangement alterations among the three groups. **(D)** Distribution of *KRAS* mutation subtypes in patients from Xuanwei. ***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$.

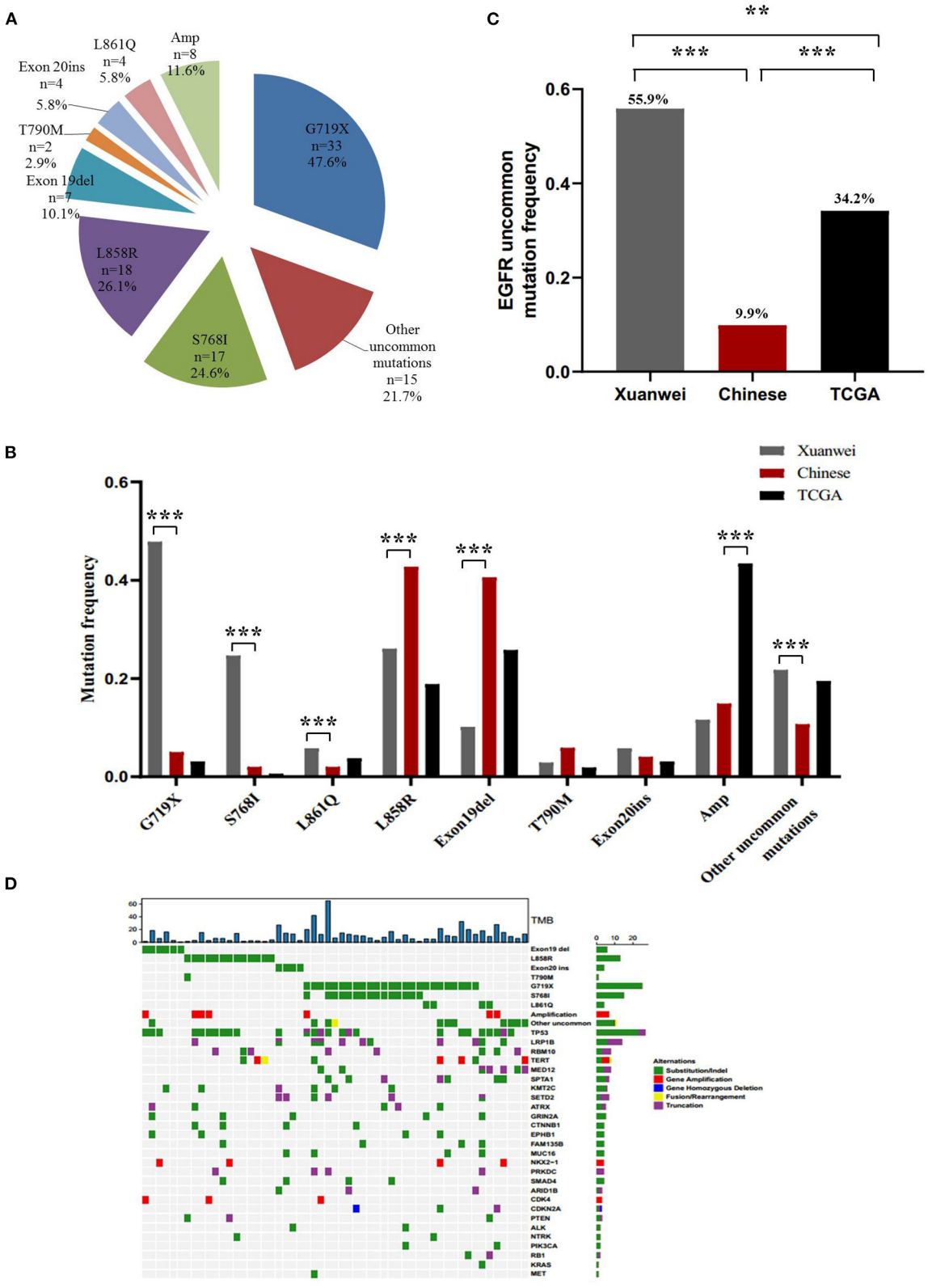


FIGURE 2 | *EGFR* mutation spectrum in the Xuanwei, reference Chinese, and TCGA NSCLC cohorts. **(A)** *EGFR* mutation subtypes of patients from Xuanwei. **(B)** Comparison of *EGFR* mutation profile among the three groups. **(C)** Ratio of uncommon *EGFR* mutations in the three groups. **(D)** Mutation profiles of patients from Xuanwei with *EGFR* genomic alterations. Mutant frequencies in the cohort are shown on the right. TMB for each patient is shown at the top. ***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$.

TABLE 2 | Comparison of characteristics of patients from Xuanwei with NSCLC harbouring common and uncommon *EGFR* mutations.

Characteristics	Overall	Common	Uncommon	*p*-value
	(*N* = 68)	(*N* = 30)	(*N* = 38)	
Gender				
Male	29 (42.6%)	12 (40%)	17 (44.7%)	0.81
Female	39 (57.4%)	18 (60%)	21 (55.3%)	
Age (years)				
Median age (range)	54 (36–78)	53 (36–72)	55 (38–78)	0.08
Stage				
I	44 (64.7%)	19 (63.3%)	25 (65.8%)	0.95
II	4 (5.8%)	2 (6.7%)	2 (5.5%)	
III	8 (11.8%)	4 (13.3%)	4 (10.5%)	
IV	11 (16.2%)	4 (13.3%)	7 (18.2%)	
Unknown	1 (1.5%)	1 (3.4%)	0	
Histology				
Adenocarcinoma	67 (98.5%)	29 (96.7%)	38 (100%)	
others	1 (1.5%)	1 (3.3%)	0	
Family history				0.048
Yes	28 (41.2%)	8 (26.7%)	20 (52.6%)	
No	39 (57.3%)	21 (70%)	18 (47.4%)	
Unknown	1 (1.5%)	1 (3.3%)	0	
Smoking history				
Yes	19 (27.9%)	9 (30%)	10 (26.3%)	0.43
Never	48 (70.5%)	20 (66.6%)	28 (73.7%)	
Unknown	1 (1.6%)	1 (3.4%)	0	
Lesions number				
1	21 (30.9%)	10 (33.3%)	11 (29%)	0.79
≥ 2	45 (66.2%)	19 (66.3%)	26 (68.4%)	
Unknown	2 (2.9%)	1 (0.4%)	1 (2.6%)	

higher *EGFR* G719X (47.6% vs. 5.0 vs. 3.1%, $p < 0.001$) and S768I (24.6% vs. 2% vs. 0.6%, $p < 0.001$) mutations, whereas classical *EGFR*-sensitive mutations, such as L858R (26.1% vs. 42.8% vs. 18.9%, $p < 0.001$) and exon 19del (10.1% vs. 40.6% vs. 25.8%, $p < 0.001$), showed significantly lower frequencies (**Figures 2A,B**; **Supplementary Figures 1A,B**).

Uncommon *EGFR* mutations were the predominant *EGFR* mutation type in the Xuanwei cohort compared to the reference Chinese cohort (55.9 vs. 9.9%, $p < 0.001$) (**Figure 2C**). The most commonly co-mutated genes are shown in **Figure 2D**. Tumours harboured a higher ratio of *EGFR* G719X and S768I co-mutations, which were mutually exclusive of L858R and exon 19del (**Figure 2D**). The clinicopathological characteristics of the Xuanwei NSCLC cohort with either uncommon or common *EGFR* mutations are summarised in **Table 2**. Patients with uncommon *EGFR* mutations were more likely to have a family history of cancer than those with common *EGFR* mutations ($p = 0.048$).

Comprehensive Profiling and Mutational Signatures in the Xuanwei Cohort

The most commonly mutated genes in the Xuanwei NSCLC cohort were *TP53* (51%), *EGFR* (49%), *KRAS* (28%), *LRP1B* (26%), and *SPTA1* (23%) (**Figure 3A**). The mutation statuses of NSCLC-related pathways were analysed. Genes involved in these signalling pathways that were included in the 450-gene panel are listed in **Supplementary Table 3**. Gene mutations in the Wnt/MAPK/ERBB signalling pathways were the most common in patients from Xuanwei with NSCLC (**Figure 3B**). The Xuanwei cohort showed a higher median TMB than the reference Chinese cohort and TCGA cases (13.1, 4.6, and 6.9 muts/Mb, respectively; $p < 0.001$) (**Figure 3C**). TMB-H was detected in 58.6% of the Xuanwei cohort and 23.5% of the reference Chinese cohort ($p < 0.001$) (**Figure 3D**). TMB and mutational signatures reflect the process of mutation accumulation in cancer. To gain further insights into the mutational process of patients from Xuanwei, we characterised the mutation signatures via analysis of TMB-H and TMB-L tumours using the somatic mutation data. The profile of somatic mutations is shown in **Figure 3E**. We identified 3,596 single nucleotide variants and 58 insertions/deletions from 111 paired sequences of NSCLC in Xuanwei cohorts. The six subtypes of base substitutions (C > A, C > G, C > T, T > A, T > C, and T > G) were unevenly represented in single nucleotide variants. C > A was the most common substitution (1,685, 49.8%), followed by C > T (752, 22.2%) (**Figure 3E**).

We analysed the Spearman correlation coefficient between TMB and mutational signatures. Four predominant signatures (APOBEC, Smoking, Signature 13, and Signature 24) were observed in tumours with TMB-H, whereas Signature 1, Signature 9, and Signature 23 showed a correlation with TMB-L,

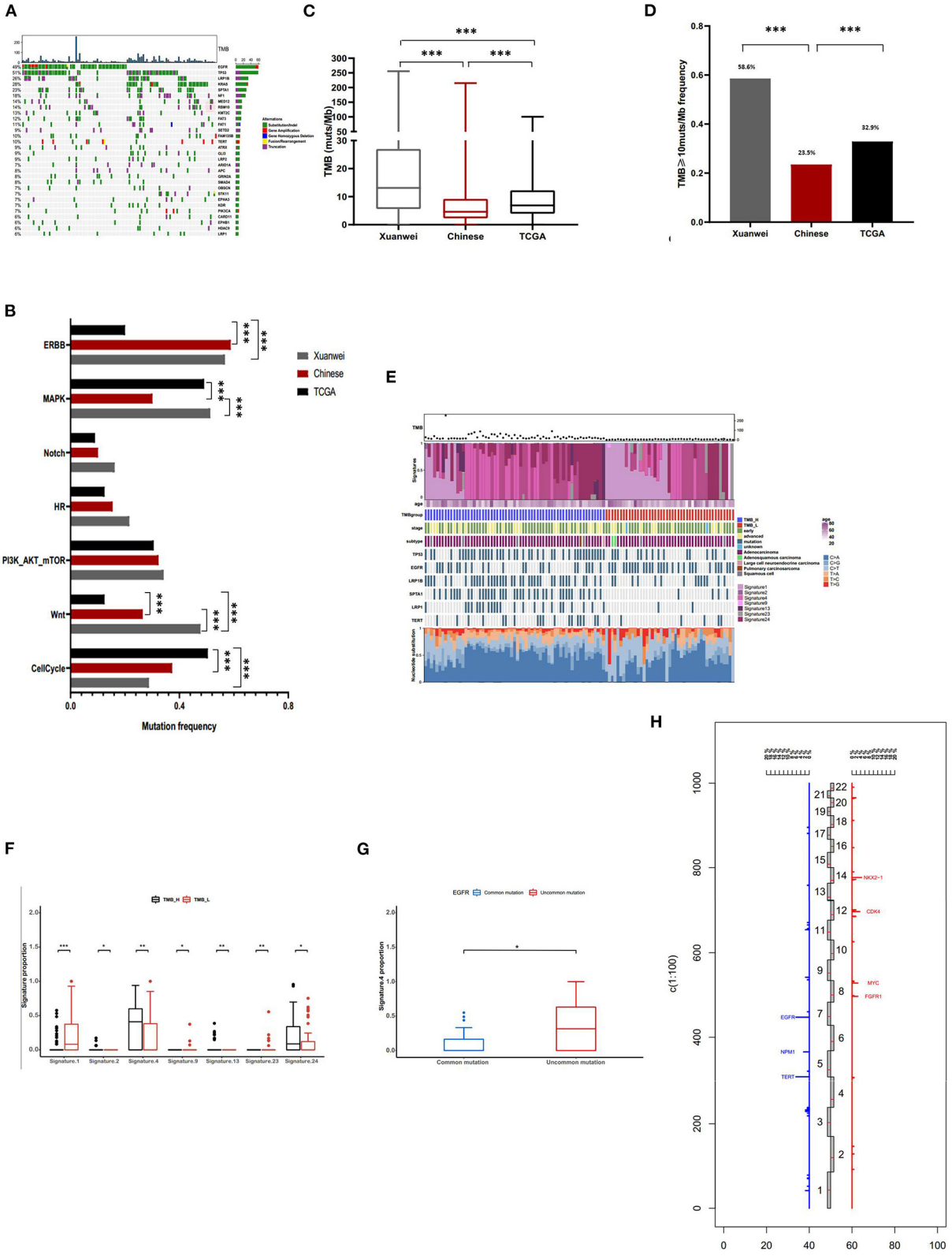


FIGURE 3 | Comprehensive profiling of and somatic mutation signatures in patients from Xuanwei with NSCLC. **(A)** Comprehensive profiling of 111 patients from the Xuanwei cohort. **(B)** Comparison of the detection rate of gene mutations in pathways in the three cohorts. **(C)** Comparison of the proportion of the TMB value and **(D)** patients with TMB-H. **(E)** Somatic mutation signatures of the Xuanwei cohort. **(F,G)** Correlation between TMB value/*EGFR* mutation type and mutation signature. **(H)** Frequency of copy number variations of all chromosomal changes. Numbers 1–22 in the middle represent the human chromosome number. ***$p < 0.001$, **$p < 0.01$, and *$p < 0.05$.

TABLE 3 | Mutation characteristics and outcome of EGFR-TKI treatment.

Patient	Age (years)	Sex	Stage	*EGFR* mutation	Best response	EGFR-TKI	Therapy	Progression-free survival (months)	Resistance mechanism
1	70	Male	IV	*EGFR* p.G719S/p.S768I/ p.N1107D	SD	Afatinib	First-line	16	Disappearance of EGFR
2	70	Female	IV	*EGFR* p.G719S/ p.S768I/Amplification	SD	Afatinib	First-line	7	Unknown
3	68	Female	IV	*EGFR* p.G719S/ p.S768I	PR	Afatinib	First-line	8	Disappearance of EGFR
4	54	Female	IIIB	*EGFR* p.G719S/ p.S768I	CR, ongoing treatment	Afatinib	First-line	12	
5	70	Female	IV	*EGFR* p.G719C/p.S768I	PR	Gefitinib	First-line	11	*EGFR* p.G719C/p.S768I
6	48	Male	IV	*EGFR* p.G719C/p.S768I/ Amplification	PR, ongoing treatment	Osimertinib	First-line	8	
7	49	Male	IV	*EGFR* p.G719C/p.K714E /p.V717L	SD, ongoing treatment	Gefitinib	First-line	2	
8	59	Male	IV	*EGFR* p.G719C/p.S768I	PR	Gefitinib	First-line	12	*MET* amplification, *PTEN* L325V
9	52	Male	IV	*EGFR* p.G719C/p.S768I	PD	Osimertinib	Third-line	1	Unknown

PD, progressive disease; PR, partial response; SD, stable disease.

suggesting the extensive accumulation of mutations due to smoking and exposure to aflatoxin. Although Signature 24 has been found in cancer samples from patients with known exposure to aflatoxin (supported by COSMIC data) (**Figure 3F**), further exploration of this association is necessary to confirm this hypothesis. Uncommon *EGFR* mutations showed high frequency about smoking-associated signature (**Figure 3G**).

Copy number variation analysis showed that the most frequently amplified genes were *EGFR* (6.3%), *TERT* (6.3%), *NKX2-1* (4.5%), *CDK4* (3.6%), *FGFR1* (2.7%), *MYC* (2.7%), *NMP1* (2.7%), and *SDHA* (2.7%) (**Figure 3H**). The association between gene mutations and TMB was analysed (**Supplementary Figure 2A**); among the common actionable mutations, *EGFR* mutation was significantly inversely correlated with TMB (6.9 vs. 17 muts/Mb, $p < 0.001$). Tumours with uncommon *EGFR* mutations showed a significantly higher TMB than those with common *EGFR* mutations (11.2 vs. 3.1 muts/Mb, $p < 0.01$) (**Supplementary Figures 2A,B**). *KRAS* mutation cases showed a higher median TMB than *KRAS* wild-type mutation cases (21.1 vs. 9.6 muts/Mb, $p < 0.01$). Tumours with *KRAS* G12C showed a lower TMB than those with non-*KRAS* G12C mutations (12 vs. 20.5 muts/Mb, $p < 0.05$) (**Supplementary Figures 2A,B**). *TP53*, *LRP1B*, *SPTA1*, *NF1*, and *KMT2C* mutations frequently occurred in patients with NSCLC and were significantly positively correlated with TMB ($p < 0.001$) (**Supplementary Figure 2C**).

Antitumor Activity of EGFR-TKI in Patients From Xuanwei With Uncommon *EGFR* Mutations

Nine patients from Xuanwei with advanced lung adenocarcinoma bearing uncommon *EGFR* mutations, with a median age of 59 years, started EGFR-TKI treatment. Detailed mutation characteristics and the outcome of EGFR-TKI treatment are shown in **Table 3**. Nine patients included in the efficacy assessment had the sensitive uncommon G719X (9/9, 100%) and S768I (8/9, 89%) mutations. Before EGFR-TKI treatment, five patients showed metastasis in the lungs, two were diagnosed with brain metastasis, and one was diagnosed with bone metastasis. Eight patients received EGFR-TKI as first-line therapy, and one patient received osimertinib as third-line therapy. As per RECIST 1.1 guidelines, five patients achieved partial response or complete response, three showed stable disease, and the one that received osimertinib as third-line therapy showed progressive disease. Treatment-related adverse events included rash (2/9), pruritus (2/9), headache (2/9), stomatitis (2/9), and constipation (1/9). No grade 3 or more adverse events were identified. At data cut-off (February 1, 2020), three patients presented sustained disease control and remained on EGFR-TKI treatment, while the other six showed progressive disease. Four patients provided serial plasma to characterise the acquired resistance mechanisms: *EGFR* mutation loss was observed in two patients at the time of EGFR-TKI treatment. *MET* amplification emerged post-EGFR-TKI treatment in one patient; no resistance mechanisms were detected in the remaining patients via NGS. As blood analysis is not 100% sensitive for the detection of acquired mutations, other genetic alterations may be associated with resistance.

DISCUSSION

High mortality and incidence rates of lung cancer have been documented in Xuanwei County. Our study revealed that the Xuanwei cohort had a greater cancer-related family history compared with the reference Chinese cohort, and most patients from Xuanwei showed a family history of lung cancer (96.7%). However, we did not detect any germline mutations in patients with family history. Patients with uncommon *EGFR* mutations were more likely to have a family history of cancer than those with common *EGFR* mutations ($p = 0.05$). A multicentre case-control study suggested that environmental tobacco smoke exposure could influence the *EGFR* mutation profile of lung cancer in never smokers and that exposure during adulthood might reduce the probability of *EGFR* mutation (17). Uncommon *EGFR* mutations and TMB-H were the predominant genomic features of patients with NSCLC from Xuanwei. Mutational signatures analysis demonstrated that uncommon *EGFR* mutations and TMB-H were correlated with smoking signature and we speculate that the environment (smoky coal) in Xuanwei may influence molecular features of patients from Xuanwei with NSCLC, which requires further analysis.

Patients with NSCLC harbouring *EGFR* mutations may benefit from EGFR-TKI therapy (18–20). The "uncommon" *EGFR* mutations account for 10–18% of all *EGFR* mutations, and NGS testing can broaden the spectrum of aberrations within the "uncommon group" in patients with NSCLC (21). Patients with *EGFR* mutations, including L858R, ex19del, and T790M, show a good response rate to first- or third-generation EGFR-TKIs, whereas those with uncommon *EGFR* mutations generally exhibit less benefit from targeted therapy (22). In our study, the NGS-based analysis of patients from Xuanwei with NSCLC revealed their comprehensive and unique profile of genomic alterations; uncommon *EGFR* mutations, mainly including G719X, S768I, and L861Q, were the predominant *EGFR* mutation types in the Xuanwei cohort, forming a distinctive subgroup of NSCLC globally. For patients with such non-classical *EGFR* mutations, previous studies illustrated that progression-free survival was significantly longer after afatinib treatment than that after first-generation EGFR-TKI (gefitinib/erlotinib) treatment (11.3 vs. 3.6 months, $p = 0.03$) (23). Osimertinib shows favourable activity with manageable toxicity in patients with metastatic or recurrent NSCLC harbouring uncommon *EGFR* mutations, achieving a median progression-free survival of 8.2 months (95% CI, 5.9 to 10.5 months) (24). According to our study, patients from Xuanwei with NSCLC harbouring G719X/S768I co-mutations may benefit from first-, second-, and third-generation EGFR-TKI treatment. The efficacy of these EGFR-TKIs in advanced patients from Xuanwei with uncommon *EGFR* mutations requires further analysis. Adjuvant therapy with gefitinib led to significantly longer disease-free survival in patients with completely resected stage II–IIIA NSCLC

with *EGFR* mutations (exon 19del and exon 21 L858R) than platinum-based chemotherapy (25). Whether patients from Xuanwei with uncommon *EGFR* mutations could benefit from adjuvant EGFR-TKI treatment is also worth further investigation.

Acquired resistance to EGFR-TKIs is a common event, and several mechanisms, including T790M, *MET* amplification, and *PTEN* down-regulation, have been reported for the common *EGFR* mutations exon 19del and L858R (26). However, mechanisms underlying EGFR-TKI resistance have not been investigated for uncommon *EGFR* mutations. Our study revealed that *MET* amplification and loss of *EGFR* mutation loss were related to EGFR-TKI resistance in patients harbouring *EGFR* G719X/S768I co-mutations. Nevertheless, further studies are required to confirm these mechanisms.

KRAS is a G-protein with intrinsic GTPase activity. *KRAS* mutations are associated with reduced responsiveness to EGFR-TKI therapy and poor survival (27). The *KRAS* mutation frequency in the Xuanwei cohort was significantly higher than that in the reference Chinese cohort. Furthermore, *KRAS* G12C accounted for 53.8% of *KRAS*-mutant patients in the Xuanwei cohort. AMG 510 is a novel, first-in-class, small molecule that specifically and irreversibly inhibits KRAS G12C by permanently locking it in an inactive GDP-bound state, and it demonstrated promising antitumor activity in patients with advanced NSCLC harbouring the *KRAS* G12C mutation (28). Patients from Xuanwei, who harbour the *KRAS* G12C mutation, may benefit from this inhibitor.

The TMB is an evolving biomarker for identifying eligible patients for checkpoint inhibitors in NSCLC (29, 30). Patients from Xuanwei with NSCLC showed a much higher median TMB than those in the reference Chinese and TCGA cohorts. Thus, patients from Xuanwei with NSCLC may benefit from immunotherapy. Furthermore, the TMB of patients with uncommon *EGFR* mutations was significantly higher than that of patients with common *EGFR* mutations in the Xuanwei cohort. Whether patients with uncommon *EGFR* mutations can benefit from immunotherapy requires further investigation.

This study comprehensively elucidated the molecular features of patients from Xuanwei with NSCLC via NGS. These patients appear to have a higher uncommon *EGFR* mutation ratio and a higher TMB value than reference Chinese patients and TCGA cohorts, and patients with uncommon *EGFR* mutations seem to show a good response to EGFR-TKI therapy. Studies with larger cohorts are needed to validate these observations, and further clinical research is warranted to provide insights into how comprehensive genomic profiling can guide treatment decisions for patients from Xuanwei with NSCLC.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. This data can be found here: https://db.cngb.org/cnsa/review/show/CNP0001608_20210319_9919b624/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by First Hospital of Kunming Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YS, TS, and GG contributed to conception and design of the study. HL, QG, JL, and XY provided study or patients material. GL and YL collected and/or assembled data. GL, YL, and YS performed the statistical analysis and interpreted data. YS, TS, GG, and YL wrote the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study was funded by the National Natural Science Foundation of China (81760554), Applied Basic Research in Yunnan Province (2017FE468-055), Kunming Medical Association Special Project for Applied Basic Research of Yunnan (2017FE467-142), and Yunnan Province Health and Family Planning Commission Medical Reserve Talents Plan (H-2017013).

ACKNOWLEDGMENTS

We thank the patients for providing their samples for this study and OrigiMed for conducting genomic profiling.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.621422/full#supplementary-material

REFERENCES

1. Cao M, Chen W. Epidemiology of lung cancer in China. *Thorac Cancer.* (2019) 10:3–7. doi: 10.1111/1759-7714.12916
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* (2020) 70:7–30. doi: 10.3322/caac.21590
3. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc.* (2008) 83:584–94. doi: 10.4065/83.5.584
4. Chen G, Sun X, Ren H, Wan X, Huang H, Ma X, et al. The mortality patterns of lung cancer between 1990 and 2013 in Xuanwei, China. *Lung Cancer.* (2015) 90:155–60. doi: 10.1016/j.lungcan.2015.08.006
5. Xiao Y, Shao Y, Yu X, Zhou G. The epidemic status and risk factors of lung cancer in Xuanwei City, Yunnan Province, China. *Front Med.* (2012) 6:388–94. doi: 10.1007/s11684-012-0233-3
6. Marx A, Chan JKC, Coindre J-M, Detterbeck F, Girard N, Harris NL, et al. The 2015 World Health Organization classification of tumors of

the thymus: continuity and changes. *J Thorac Oncol.* (2015) 10:1383–95. doi: 10.1097/JTO.0000000000000654
7. Li R, Liu Y, Wang T, Tang J, Xie L, Yao Z, et al. The characteristics of lung cancer in Xuanwei County: a review of differentially expressed genes and noncoding RNAs on cell proliferation and migration. *Biomed Pharmacother.* (2019) 119:109312. doi: 10.1016/j.biopha.2019.109312
8. Chen Y, Ye L, Stanford RR, Zhang D, Zhang Z, Wei W. Distinct epithelial growth factor receptor mutation profile in non-small-cell lung cancer patients from the Xuanwei area of China. *Mol Clin Oncol.* (2016) 4:749–755. doi: 10.3892/mco.2016.805
9. Chen H, Chong W, Teng C, Yao Y, Wang X, Li X. The immune response-related mutational signatures and driver genes in non-small-cell lung cancer. *Cancer Sci.* (2019) 110:2348–56. doi: 10.1111/cas.14113
10. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet.* (2016) 48:607–16. doi: 10.1038/ng.3564
11. Cao J, Chen L, Li H, Chen H, Yao J, Mu S, et al. An accurate and comprehensive clinical sequencing assay for cancer targeted and immunotherapies. *Oncologist.* (2019) 24:e1294–302. doi: 10.1634/theoncologist.2019-0236
12. Tang B, Yan X, Sheng X, Si L, Cui C, Kong Y, et al. Safety and clinical activity with an anti-PD-1 antibody JS001 in advanced melanoma or urologic cancer patients. *J Hematol Oncol.* (2019) 12:7. doi: 10.1186/s13045-018-0693-2
13. Guo J, Huang J, Zhou Y, Zhou Y, Yu L, Li H, et al. Germline and somatic variations influence the somatic mutational signatures of esophageal squamous cell carcinomas in a Chinese population. *BMC Genomics.* (2018) 19:538. doi: 10.1186/s12864-018-4906-4
14. Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* (2016) 17:31. doi: 10.1186/s13059-016-0893-4
15. Wen S, Dai L, Wang L, Wang W, Wu D, Wang K, et al. Genomic signature of driver genes identified by target next-generation sequencing in Chinese non-small cell lung cancer. *Oncologist.* (2019) 24:e1070–81. doi: 10.1634/theoncologist.2018-0572
16. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* (2009) 45:228–47. doi: 10.1016/j.ejca.2008.10.026
17. Torres-Duran M, Ruano-Ravina A, Kelsey KT, Parente-Lamelas I, Leiro-Fernández V, Abdulkader I, et al. Environmental tobacco smoke exposure and EGFR and ALK alterations in never smokers' lung cancer. Results from the LCRINS study. *Cancer Lett.* (2017) 411:130–5. doi: 10.1016/j.canlet.2017.09.042
18. Mitsudomi T, Morita S, Yatabe Y, Negoro S, Okamoto, I, et al. Gefitinib versus cisplatin plus docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): an open label, randomised phase 3 trial. *Lancet Oncol.* (2010) 11:121–8. doi: 10.1016/S1470-2045(09)70364-X
19. Yang JC-H, Wu Y-L, Schuler M, Sebastian M, Popat S, Yamamoto N, et al. Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung adenocarcinoma (LUX-Lung 3 and LUX-Lung 6): analysis of overall survival data from two randomised, phase 3 trials. *Lancet Oncol.* (2015) 16:141–51. doi: 10.1016/S1470-2045(14)71173-8
20. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol.* (2012) 13:239–46. doi: 10.1016/S1470-2045(11)70393-X
21. O'Kane GM, Bradbury PA, Feld R, Leighl NB, Liu G, Pisters K-M, et al. Uncommon EGFR mutations in advanced non-small cell lung cancer. *Lung Cancer.* (2017) 109:137–44. doi: 10.1016/j.lungcan.2017.04.016
22. Xu J, Jin B, Chu T, Dong X, Yang H, Zhang Y, et al. EGFR tyrosine kinase inhibitor (TKI) in patients with advanced non-small cell lung cancer (NSCLC) harboring uncommon EGFR mutations: a real-world study in China. *Lung Cancer.* (2016) 96:87–92. doi: 10.1016/j.lungcan.2016.01.018
23. Shen Y-C, Tseng G-C, Tu C-Y, Chen W-C, Liao W-C, Chen W-C, et al. Comparing the effects of afatinib with gefitinib or Erlotinib in patients with advanced-stage lung adenocarcinoma harboring non-classical epidermal growth factor receptor mutations. *Lung Cancer.* (2017) 110:56–62. doi: 10.1016/j.lungcan.2017.06.007
24. Cho JH, Lim SH, An HJ, Kim KH, Park KU, Kang EJ, et al. Osimertinib for patients with non-small-cell lung cancer harboring uncommon EGFR mutations: a multicenter, open-label, phase ii trial (KCSG-LU15-09). *J Clin Oncol.* (2020) 38:488–95. doi: 10.1200/JCO.19.00931
25. Zhong W-Z, Wang Q, Mao W-M, Xu S-T, Wu L, Shen, et al. Gefitinib versus vinorelbine plus cisplatin as adjuvant treatment for stage II-IIIA (N1-N2) EGFR-mutant NSCLC (ADJUVANT/CTONG1104): a randomised, open-label, phase 3 study. *Lancet Oncol.* (2018) 19:139–48. doi: 10.1016/S1470-2045(17)30729-5
26. Wu SG, Shih JY. Management of acquired resistance to EGFR TKI-targeted therapy in advanced non-small cell lung cancer. *Mol Cancer.* (2018) 17:38. doi: 10.1186/s12943-018-0777-1
27. Eberhard DA, Johnson BE, Amler LC, Goddard AD, Heldens SL, Herbst RS, et al. Mutations in the epidermal growth factor receptor and in KRAS are predictive and prognostic indicators in patients with non-small-cell lung cancer treated with chemotherapy alone and in combination with erlotinib. *J Clin Oncol.* (2005) 23:5900–9. doi: 10.1200/JCO.2005.02.857
28. Canon J, Rex K, Saiki AY, Mohr C, Cooke K, Bagal D, et al. The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature.* (2019) 575:217–23. doi: 10.1038/s41586-019-1694-1
29. Hellmann MD, Ciuleanu T-E, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. *N Engl J Med.* (2018) 378:2093–104. doi: 10.1056/NEJMoa1801946
30. 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

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

CASE REPORT
published: 12 April 2021
doi: 10.3389/fonc.2021.569429

Case Report: Next-Generation Sequencing Reveals Tumor Origin in a Female Patient With Brain Metastases

Qun Li[1], Xiaoyan Zhang[2], Jiao Feng[3], Dezhi Cheng[4], Lin Cai[1], Zhang'an Dai[1], Shuyu Zhao[5], Jianmin Li[5], Jingjing Huang[2], Yu Fang[2], Honglin Zhu[2], Danhua Wang[2], Sizhen Wang[6], Tonghui Ma[2] and Xianghe Lu[1]**

[1] Neurosurgery department, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [2] Department of Translational Medicine, Genetron Health (Beijing) Technology, Co. Ltd., Beijing, China, [3] Holistic Integrative Pharmacy Institutes and Comprehensive Cancer Diagnosis and Treatment Center, College of Medicine, Hangzhou Normal University, Hangzhou, China, [4] Thoracic surgery department, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [5] Pathology department, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [6] Genetron Health (Beijing) Technology, Co. Ltd., Beijing, China

OPEN ACCESS

Edited by:
Qinghua Xu,
Canhelp Genomics, China

Reviewed by:
Hamid Morjani,
Université de Reims Champagne-Ardenne, France
Vandna Kukshal,
Washington University School of Medicine in St. Louis, United States

***Correspondence:**
Xianghe Lu
dr_luxh@sina.com
Tonghui Ma
tonghuima0818@sina.com

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 04 June 2020
Accepted: 18 March 2021
Published: 12 April 2021

Citation:
Li Q, Zhang X, Feng J, Cheng D, Cai L, Dai Z, Zhao S, Li J, Huang J, Fang Y, Zhu H, Wang D, Wang S, Ma T and Lu X (2021) Case Report: Next-Generation Sequencing Reveals Tumor Origin in a Female Patient With Brain Metastases. Front. Oncol. 11:569429. doi: 10.3389/fonc.2021.569429

Background: Brain metastasis mainly originates from lung cancer. Napsin A and TTF-1 factors have frequently been detected in lung adenocarcinoma cases. Brain metastasis tumors with napsin A and TTF-1 positive are easily classified as lung adenocarcinoma origin. However, some thyroid cancers also exhibit these clinical features. Besides, lung is the most common metastasis of undifferential thyroid cancer. Therefore, it requires development of novel diagnostic tools to aid in distinguishing between pulmonary and thyroid origin.

Patient Findings: We reported a case that was initially diagnosed as brain metastatic lung cancer based on immunohistochemistry results. Analysis of next-generation sequencing (NGS) data from the brain lesion revealed that the cancer may have originated from the thyroid. We detected combo mutations in *TERT* promoter mutation, *RET* fusion and *TP53*, which are common in undifferential thyroid cancer (UTC), but rare for lung cancer. These results, coupled with identification of PAX8, indicated that this patient had UTC. Additionally, her three sons, despite being asymptomatic, were all diagnosed with papillary thyroid carcinoma.

Summary: The patient received anlotinib treatment and showed good clinical outcomes. One month after anlotinib treatment, the pulmonary nodules were found to be controlled, and the thyroid tumor drastically reduced, and tracheal compression relieved. She continued anlotinib treatment for the following two months, but died one month later because the treatment stopped owing to financial reasons. All her sons underwent total thyroidectomy with lymph node dissection.

Conclusions: Although NGS has been reported to assist in diagnosis of the origin of some tumors, this is the first evidence of NGS for the determination of the origin of thyroid tumors. To our knowledge, this is the first time that a combination of multiple mutations has been used

to help determine the origin of a tumor, compared with the previous single mutant gene. Moreover, this is the first evidence on the use of anlotinib for treatment of UTC with distant metastasis. Besides, all three sons of the patient had thyroid carcinoma in subsequent examinations, indicating high-risk for familial non-medullary thyroid cancer in UTC patients and necessity for performing thyroid ultrasound testing in other family members.

Keywords: next-generation sequencing, undifferential thyroid cancer, *TERT* promoter mutation, *RET* fusion, anlotinib, tumor origin

INTRODUCTION

Brain metastasis results from various types of cancer, key among them being lung cancer (1). Clinical diagnosis of lung adenocarcinoma results from positive expression of napsin A aspartic peptidase (napsin A), and transcription termination factor-1 (TTF-1) markers (2, 3). In fact, brain metastasis tumors that positively express napsin A and TTF-1 are always classified as lung adenocarcinoma origin.

Undifferential thyroid cancer (UTC) is the least common type accounting for 1-2% of thyroid cancer cases. In fact, advanced distant metastatic disease is the most challenging condition among patients (4, 5). Intrathoracic, neck lymph nodes and lungs are the most common metastatic sites of UTC (6, 7). UTC often expresses PAX8 and loses thyroid cancer-specific marker-thyroglobulin (TG), and some of UTC tumors show positive napsin A and TTF-1 expression (3, 8). Immunophenotyping of the metastases for some UTC tumors is, therefore, similar to that for primary lung adenocarcinoma. Next-generation sequencing (NGS) has been successfully applied in clinical diagnosis, enabling identification of genetic features. Oncogenetic mutations in the MAPK (*BRAF*, *RAS*), PI3K and mismatch repair pathways, *RAC1*, *TP53*, and *TERT* promoters, as well as *CCDC6-RET* fusion have often been identified in UTC. However, mutations in the *TERT* promoter and *RET* fusion have been rarely detected in lung cancer (9, 10). Despite identification of several oncogenic mutations, only one mutation-driven targeted therapy has so far been approved for *BRAF*-mutant UTC (10).

Anlotinib is an oral novel multi-target tyrosine kinase inhibitor that targets RET, vascular endothelial growth factor receptor (VEGFR), fibroblast growth factor receptor (FGFR), platelet-derived growth factor receptors (PGFR) and c-kit (11, 12). When compared to sunitinib, anlotinib exhibits significantly lower side effects. Consequently, anlotinib has currently become an effective compound for inhibition of multi-targeting receptor tyrosine kinases, owing to its efficacy against multiple cancers, including lung cancer (13–16), advanced medullary thyroid cancer (17), soft tissue sarcoma and metastatic renal cell carcinoma (12). However, its potential to treat UTC has not been reported.

In the current study, we studied a case that was initially considered brain metastatic lung cancer based on the immunohistochemistry (IHC) results, but found to be cancer originating from thyroid after NGS of the brain lesions. Specifically, we identified combo mutations in, *TERT* promoter, *RET* fusion and TP53, which are common for UTC but rare for lung cancer. We administered anlotinib treatment to the patient and achieved good clinical outcomes. This is a first study reporting the use of anlotinib for UTC with lung metastasis.

CASE DESCRIPTION

A 63-year-old female patient was presented to hospital, in March 2019, with repeated dizziness and visual impairment. A computerized tomography/magnetic resonance imaging (CT/MRI) revealed a mass lesion in the left occipital lobe. In addition to the brain mass, positron emission tomography (PET)-CT also indicated a hyper-metabolic mass in the left thyroid as well as some opaque mottled shadows and pulmonary nodules in the lung (**Figure 1**). Further ultrasound

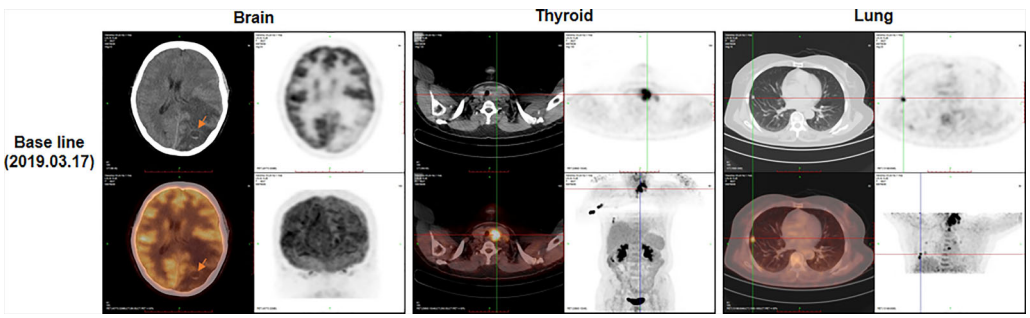


FIGURE 1 | The positron emission tomography-computed tomography (PET-CT) at baseline (before surgery). PET-CT showed masses in brain (arrow), thyroid and lung.

revealed multiple thyroid nodules, the largest calcified one located in the left thyroid and measuring 41*30*44 mm. The patient was subjected to brain tumor resection. IHC studies of the resected brain tumor positively identified cytokeratin 7 (CK7), napsin A and TTF-1, but TG was negative (**Figures 2A–C**; Details on IHC methods were showed in **Supplementary Data**). Fine-needle aspiration (FNA) of the thyroid lesion identified clusters of undifferentiated tumor cells. Based on these results, the pathologist hypothesized that the lesions in brain and thyroid had migrated from undifferentiated lung adenocarcinoma. However, the NGS results from FSZ-Thyroid NGS Panel V1 (Genetronhealth) revealed *TERT* promoter mutation, *RET* fusion and *TP53* mutation in the brain lesion (**Table 1** and **Supplementary Figure 1**). The details of NGS were shown in the **Supplementary Data**. These mutations combo are indicative of undifferential thyroid cancer (9), and rarely occur in lung cancers. Since the patient did not consent to further surgery, we did not obtain any operation samples on the thyroid or lung for further analysis. In addition, we positively detected PAX8 through IHC (**Figure 2D**). This factor is negative in lung adenocarcinoma, but always positive in thyroid cancer (18). Based on these results (positive PAX8, negative TG, cytopathology of thyroid FNA and the special genomic features), we concluded that the patient had undifferential thyroid cancer, with lung and brain metastasis. Due to the patient with brain metastasis and without *BRAF* mutations, RET-targeted BLU-667 clinical trial and BRAF inhibitor were not suitable for this patient. Therefore, indication therapy or off-label treatment were needed for this patient. After comprehensive consideration, we administered anlotinib, a novel multi-kinase inhibitor that has also been found to block RET (11). One month after treatment, the pulmonary nodules were found to be controlled, thyroid tumor drastically reduced and tracheal compression relieved (**Figure 3**). Unfortunately, the patient stopped anlotinib treatment, after three months, owing to financial constraints. Only two weeks after discontinuation of anlotinib, the patient went back to the hospital exhibiting breathlessness and cough. A CT scan revealed a large mass outbreak in her lungs, and she eventually died of respiratory failure.

Given the patient's critical condition, three sons of the proband, who were asymptomatic, were presented to the local hospital and subjected to thyroid ultrasound examinations. They were all found with thyroid lesions. Total thyroidectomy, with lymph node dissection, was performed on all of them. Results revealed a thyroid papillary carcinoma (1.2cm), in the right thyroid lobe and isthmus (without lymph node metastasis) of the 36 years old youngest son. Pathologic diagnosis in the elder son (40 years old) revealed a follicular adenoma in the left lobe as well as a thyroid papillary microcarcinoma in the right lobe and

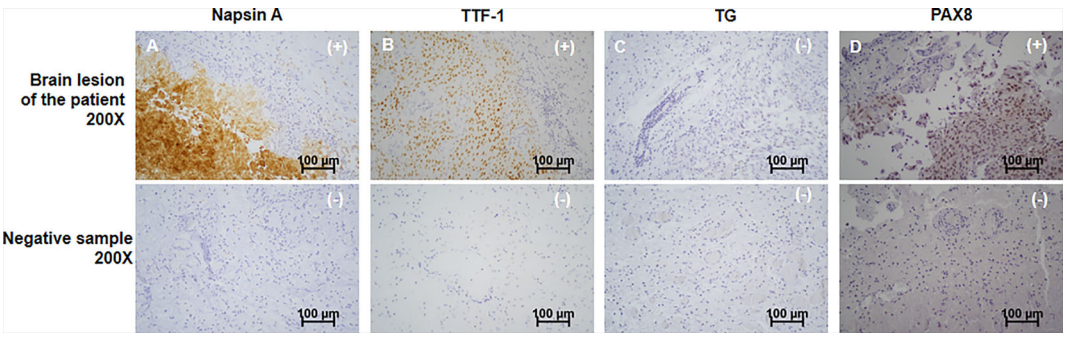


FIGURE 2 | Immunohistochemistry of the patient's brain cancer. From top to bottom, the immunochemistry results of the patient's brain tumor and negative control sample were shown. The patient's tumor showed positive of Napsin A **(A)**, TTF-1 **(B)** and PAX8 **(D)**, showing as "(+)" in top right corner of the graphs. Negative of TG **(C)** was identified, presenting as "(-)" in top right corner of the graphs. Corresponding negative results from negative control sample were also shown. Scale bar was in the lower right corner of each image. All images were magnified by 200 times. Scale bars (100μm) were provided.

TABLE 1 | NGS Gene Mutation Profiling of tumor tissues.

Samples	Gene	Mutation type	DNA _ Change	Protein _ Change
Brain tumor of the patient	*TERT*	Upstream promoter mutation	C228T	
	RET	Gene fusion	*CCDC6{NM_005436.4}:r.1_303+1_RET{NM_020975.4}:r.2137_5617*	
	TP53	Missense mutation	*c.842A>T*	p. D281V
The eldest son	*BRAF*	Missense mutation	*c.1799T>A*	p. V600E
The second son	*BRAF*	Missense mutation	*c.1799T>A*	p. V600E
The youngest son	*BRAF*	Missense mutation	*c.1799T>A*	p. V600E

NGS, next-generation sequencing.

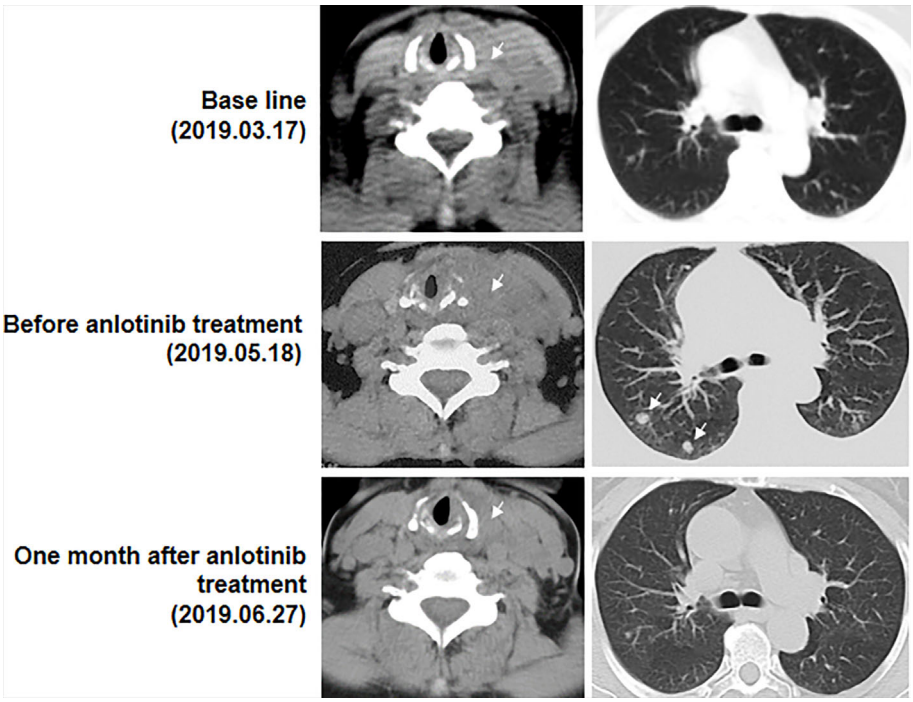


FIGURE 3 | Computed tomography images (CT) of the thyroid and lung before and after anlotinib treatment. Arrows represent the masses. At base line, in addition to brain lesions, there were a mass in the left thyroid and some opaque mottled shadows and pulmonary nodules in the lung. Then the patient experienced brain surgery. Before anlotinib treatment, a huge mass in the left thyroid and two obvious pulmonary nodules was found. One month after anlotinib treatment, the pulmonary nodules were found to be controlled, thyroid tumor drastically reduced and tracheal compression relieved.

isthmus. Lymph node metastasis were found in Level VI of the neck. Unfortunately, a more extensive lymph node metastasis was further detected at Level II, III, IV and VI of the neck in the eldest son (42 years old). All tumor tissues from the three sons were positive of TTF-1, TG and PAX8, negative of Napsin A (**Supplementary Figure 2**). Targeted sequencing by using FSZ-Thyroid NGS Panel V1 (Genetronhealth) was performed on the tumor tissues of the three sons, the results of which showed all the tissues had BRAF V600E mutations (**Table 1** and **Supplementary Figure 3**).

DISCUSSION

We present a case of UTC, with lung and brain metastasis, in which genetic features including *CCDC6-RET* fusion, *TERT* promoter and *TP53* mutations were identified. Based on these features, a multi-targeted reagent, anlotinib, was used to treat the patient and resulted in good clinical outcomes. This is the first study reporting the use of mutation combo and IHC for the confirmation of UTC and the use of anlotinib for the treatment of UTC. The most noteworthy point in the present case is that traditional IHC markers combined with genetic characteristic of tumor are more effective in accurately determining the source of metastatic tumor. Generally, IHC is a major diagnostic tool for determining the origin of metastatic tumors (19, 20). However, this diagnosis is not enough for some special cases. A combined detection of napsin A and TTF-1 through IHC has been considered a reliable procedure for diagnosing adenocarcinoma originating from the lungs (2, 3). However, several studies have recently shown that both of Napsin A and TTF-1 are positive in approximately 15% of UTC cases (8). TG, a specific thyroid cancer marker due to its especial expression in tissues of thyroid origin has been reported (3), but its expression is often lost following dedifferentiation of UTC (8). Therefore, a positive napsin A and TTF-1 detection, as well as negative TG, is considered a misdiagnosis for lung adenocarcinoma. Studies have reported that some genetic alterations can be used to aid in determining the source of metastatic carcinoma, with higher specificity than traditional IHC markers (16). For example, *IDH* has been found to be mutated in low-grade gliomas, chondrosarcomas, hematologic malignancies and bile duct cancer (21, 22). *BRCA1* and *BRCA2* are prone to mutations in breast and ovarian cancers, with evidence of *BRCA* gene mutations indicating the genetic origin of these cancers (23). In addition, *TERT* promoter mutations are frequently identified in melanoma, thyroid cancer, bladder cancer and glioblastoma, while they are rare (2.57%) in lung cancer. On the contrary, EGFR L858R mutation is only found in lung cancer (24). Changes in *RET* expression tend to frequently occur in

thyroid, and invasive breast cancers, as well as pancreatic ductal adenocarcinomas, but are rare in lung cancer. Moreover, the most common *RET* fusions are *CCDC6-RET* in thyroid cancer and *KIF5B-RET* in lung cancer (25). A combination of mutations including *CCDC6-RET* fusion, *TERT*-promoter and *TP53* often occurs in dedifferentiated thyroid cancer (9). However, there is a probability of only 0.005% of this happening in lung adenocarcinoma (24, 25). Combo mutations from NGS in this case clearly confirmed the origin of the thyroid tumors; thus, this is the first case that the determination of the tumor origin by the assist of combo mutations, compared with previous single mutant gene. Moreover, IHC analysis positively detected PAX8, which has been strongly associated with thyroid origin (8), suggesting that this factor is an additional tool for discriminating thyroid from lung cancer in the setting of a double positive of napsin A and TTF-1.

Our results showed that anlotinib is effective in treating UTC, a rare thyroid cancer with a high mortality rate. Conventional treatments, such as surgery, radiotherapy and chemotherapy, have not effectively managed. In fact, the only approved targeted therapy, dabrafenib with trametinib, only works in $BRAF^{V600E}$-positive anaplastic thyroid carcinoma patients. Several novel biological agents, as well as immune checkpoint and aurora kinase inhibitors, are currently under testing for treatment of UTC (26). Case in the current study, a patient with *RET* fusion, who was excluded from RET-targeted BLU-667 clinical trial because of brain metastasis, and the patient without *BRAF* mutation, also excluded from BRAF inhibitor, finally benefited from anlotinib. Anlotinib recently showed durable antitumor activity in lung cancers (13–16), medullary thyroid carcinoma (MTC) (17) and some other cancers (12). To the best of our knowledge, this was the first time that anlotinib has been used to manage UTC. The treatment resulted in good control of primary and lung metastasis lesions. Manageable adverse events were demonstrated in anlotinib-treated lung cancer (15) and MTC (17), here also no serious side effects occurred. These clinical results indicate that anlotinib may develop a new avenue for UTC therapy.

In addition, the patient's sons all diagnosed with thyroid carcinoma, indicating obvious characteristics of familial non-medullary thyroid cancer (FNMTC). Previous studies have implicated FNMTC as an independent risk factor for increased aggressiveness of thyroid cancers (27), which may explain the critical and serious condition of the proband. To the best of our knowledge, no reports have studied FNMTC in UTC patients. It is, therefore, imperative to perform thyroid ultrasound testing on family members of patients diagnosed with UTC.

Overall, this report describes the genetic characteristics of a UTC patient, with distant metastases, and the obvious benefits of targeted therapy using anlotinib. We illustrate that a combination of driver mutations, detected by NGS, could directly guide understanding of the origin of tumors and the corresponding targeted therapy. Taken together, our findings indicate that this treatment could be an acceptable option for the urgent management of UTC.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

Written informed consent was obtained from the patient's relative for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

XL, QL, XZ, DC, LC, ZD, and JH collected this case. SZ and JL provided pathology information of this case. HZ and DW provided bioinformatics analysis of NGS data. QL and XZ wrote the manuscript. JF, YF, SW, TM, and XL modified this paper. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Wenzhou Municipal Science and Technology Bureau of China (Y2020063 and Y20180175).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.569429/full#supplementary-material

Supplementary Figure 1 | The mutation IGV results of the patient. **(A–C)** IGV results of TERT upstream promoter mutation (c.C228T (*Chr5: 1295228*)), CCDC6-RET fusion (CCDC6{NM_005436.4}:r.1_303+1_RET{NM_020975.4}: r.2137_5617) and TP53 mutation (c.842A>T (*Chr17: 7577096*)) of the patient's tumor sample and their corresponding IGV results of negative control sample.

Supplementary Figure 2 | Immunochemistry results of each son's tumor tissue. From top to bottom, the immunochemistry results of the eldest son, second son and youngest son's tumor were shown. The three sons' tumors showed negative of Napsin A, showing as "(-)"in top right corner of the graphs. Positive of TTF-1, PAX8 and TG were identified in all of the three sons' tumors, presenting as "(+)"in top right corner of the graphs. Negative results from negative control samples were also shown. The immunohistochemical tests of the patient and her three sons were repeated simultaneously, so the same negative controls were used. Scale bar was in the lower right corner of each image. All images were magnified by 200 times.

Supplementary Figure 3 | IGV results of BRAF mutation (c.1799T>A *(Chr7: 140453136)*) of the tumor samples of the patient's three sons. Corresponding IGV results of negative control sample were also provided.

REFERENCES

1. Cagney DN, Martin AM, Catalano PJ, Redig AJ, Lin NU, Lee EQ, et al. Incidence and prognosis of patients with brain metastases at diagnosis of systemic malignancy: a population-based study. *Neuro Oncol* (2017) 19:1511–21. doi: 10.1093/neuonc/nox077
2. Travis W, Burke E, Burke AP, Marx A, Nicholson AG. *WHO classification of tumours of the lung, pleura, thymus and heart fourth edition*. Geneva: IARC press (2015).
3. Wu J, Zhang Y, Ding T, Cheng R, Gong W, Guo Y, et al. Napsin A Expression in Subtypes of Thyroid Tumors: Comparison with Lung Adenocarcinomas. *Endocr Pathol* (2019) 31(1):39–45. doi: 10.1007/s12022-019-09600-6
4. Cabanillas ME, McFadden DG, Durante C. Thyroid cancer. *Lancet* (2016) 388:2783–95. doi: 10.1016/S0140-6736(16)30172-6
5. Molinaro E, Romei C, Biagini A, Sabini E, Agate L, Mazzeo S, et al. Anaplastic thyroid carcinoma: from clinicopathology to genetics and advanced therapies. *Nat Rev Endocrinol* (2017) 13:644–60. doi: 10.1038/nrendo.2017.76
6. Pomorski L, Bartos M. Metastasis as the first sign of thyroid cancer. *Neoplasma* (1999) 46:309–12.
7. Rühli FJ, Hilfiker PR. Metastasis of thyroid cancer to the heart. *AJR Am J Roentgenol* (2001) 177:474. doi: 10.2214/ajr.177.2.1770474
8. Chernock RD, El-Mofty SK, Becker N, Lewis JS, Napsin A. Expression in Anaplastic, Poorly Differentiated, and Micropapillary Pattern Thyroid Carcinomas. *Am J Surg Pathology* (2013) 37:1215–22. doi: 10.1097/PAS.0b013e318283b7b2
9. Landa I, Ibrahimpasic T, Boucai L, Sinha R, Knauf JA, Shah RH, et al. Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers. *J Clin Invest* (2016) 126:1052–66. doi: 10.1172/JCI85271
10. Pozdeyev N, Rose MM, Bowles DW, Schweppe RE. Molecular Therapeutics for Anaplastic Thyroid Cancer. *Semin Cancer Biol* (2020) 61:23–9. doi: 10.1016/j.semcancer.2020.01.005
11. Sun Y, Niu W, Du F, Du C, Li S, Wang J, et al. Safety, pharmacokinetics, and antitumor properties of anlotinib, an oral multi-target tyrosine kinase inhibitor, in patients with advanced refractory solid tumors. *J Hematol Oncol* (2016) 9:105. doi: 10.1186/s13045-016-0332-8
12. Shen G, Zheng F, Ren D, Du F, Dong Q, Wang Z, et al. Anlotinib: a novel multi-targeting tyrosine kinase inhibitor in clinical development. *J Hematol Oncol* (2018) 11:120. doi: 10.1186/s13045-018-0664-7
13. Cheng Y, Wang Q, Li K, Shi J, Wu L, Han B, et al. OA13.03 Anlotinib as Third-Line or Further-Line Treatment in Relapsed SCLC: A Multicentre, Randomized, Double-Blind Phase 2 Trial. *J Thoracic Oncol* (2018) 13:S351–2. doi: 10.1016/j.jtho.2018.08.308
14. Han B, Li K, Zhao Y, Li B, Cheng Y, Zhou J, et al. Anlotinib as a third-line therapy in patients with refractory advanced non-small-cell lung cancer: a multicentre, randomised phase II trial (ALTER0302). *Br J Cancer* (2018) 118:654–61. doi: 10.1038/bjc.2017.478
15. Si X, Zhang L, Wang H, Zhang X, Wang M, Han B, et al. Management of anlotinib-related adverse events in patients with advanced non-small cell lung cancer: Experiences in ALTER-0303. *Thorac Cancer* (2019) 10:551–6. doi: 10.1111/1759-7714.12977
16. Han B, Li K, Wang Q, Zhang L, Shi J, Wang Z, et al. Effect of Anlotinib as a Third-Line or Further Treatment on Overall Survival of Patients With Advanced Non-Small Cell Lung Cancer: The ALTER 0303 Phase 3 Randomized Clinical Trial. *JAMA Oncol* (2018) 4:1569–75. doi: 10.1001/jamaoncol.2018.3039
17. Sun Y, Du F, Gao M, Ji Q, Li Z, Zhang Y, et al. Anlotinib for the Treatment of Patients with Locally Advanced or Metastatic Medullary Thyroid Cancer. *Thyroid* (2018) 28:1455–61. doi: 10.1089/thy.2018.0022
18. Chernock RD. Immunohistochemistry of thyroid gland carcinomas: clinical utility and diagnostic pitfalls. *Diagn Histopathol* (2016) 22:184–90. doi: 10.1016/j.mpdhp.2016.04.008
19. Conner JR, Hornick JL. Metastatic carcinoma of unknown primary: diagnostic approach using immunohistochemistry. *Adv Anat Pathol* (2015) 22:149–67. doi: 10.1097/PAP.0000000000000069
20. Stelow EB, Yaziji H. Immunohistochemistry, carcinomas of unknown primary, and incidence rates. *Semin Diagn Pathol* (2018) 35:143–52. doi: 10.1053/j.semdp.2017.11.012
21. Mondesir J, Willekens C, Touat M, de Botton S. IDH1 and IDH2 mutations as novel therapeutic targets: current perspectives. *J Blood Med* (2016) 7:171–80. doi: 10.2147/JBM.S70716
22. Kamath SD, Lin X, Kalyan A. A Case of Metastatic Biliary Tract Cancer Diagnosed Through Identification of an *IDH1* Mutation. *Oncol* (2019) 24:151–6. doi: 10.1634/theoncologist.2018-0210
23. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips K-A, Mooij TM, Roos-Blom M-J, et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for *BRCA1* and *BRCA2* Mutation Carriers. *JAMA* (2017) 317:2402. doi: 10.1001/jama.2017.7112
24. Ma X, Gong R, Wang R, Pan Y, Cai D, Pan B, et al. Recurrent TERT promoter mutations in non-small cell lung cancers. *Lung Cancer* (2014) 86:369–73. doi: 10.1016/j.lungcan.2014.10.009
25. Li AY, McCusker MG, Russo A, Scilla KA, Gittens A, Arensmeyer K, et al. RET fusions in solid tumors. *Cancer Treat Rev* (2019) 81:101911. doi: 10.1016/j.ctrv.2019.101911
26. Ranganath R, Shah MA, Shah AR. Anaplastic thyroid cancer. *Curr Opin Endocrinol Diabetes Obes* (2015) 22:387–91. doi: 10.1097/MED.0000000000000189
27. Tavarelli M, Russo M, Terranova R, Scollo C, Spadaro A, Sapuppo G, et al. Familial Non-Medullary Thyroid Cancer Represents an Independent Risk Factor for Increased Cancer Aggressiveness: A Retrospective Analysis of 74 Families. *Front Endocrinol (Lausanne)* (2015) 6:117. doi: 10.3389/fendo.2015.00117

Conflict of Interest: Authors XZ, JH, YF, HZ, DW, SW, and TM were employed by the company Genetron Health (Beijing) Technology, Co. Ltd., Beijing, China.

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.

frontiers in Oncology

ORIGINAL RESEARCH
published: 07 May 2021
doi: 10.3389/fonc.2021.603564

Distinct Genomic Landscape of Colorectal Mucinous Carcinoma Determined *via* Comprehensive Genomic Profiling: Steps to a New Treatment Strategy

Liang Huang[1,2,3†], Shuanglin Luo[1,2,3†], Xingwei Zhang[1,2,3], Yonghua Cai[1,2,3], Fangqin Xue[4], Huanxin Hu[1,2,3], Ziwei Zeng[1,2,3], Tengjiao Lin[5], Fei Wang[5], Weifeng Wang[5], Sen Zhang[6] and Liang Kang[1,2,3*]*

[1] Department of Colorectal Surgery, The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, [2] Guangdong Institute of Gastroenterology, The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, [3] Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, [4] Department of Gastrointestinal Surgery, Fujian Provincial Hospital, Fuzhou, China, [5] Department of Research and Development, OrigiMed, Shanghai, China, [6] Department of Colorectal Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, China

OPEN ACCESS

Edited by:
Bing Xu,
Xiamen University, China

Reviewed by:
Wei Wu,
Central South University, China
Ravi Manoharan,
University of Madras, India
Kun Zhou,
Boston Children's Hospital, United States

***Correspondence:**
Sen Zhang
zs0771@126.com
Liang Kang
kangl@mail.sysu.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *07 September 2020*
Accepted: *08 April 2021*
Published: *07 May 2021*

Citation:
Huang L, Luo S, Zhang X, Cai Y, Xue F, Hu H, Zeng Z, Lin T, Wang F, Wang W, Zhang S and Kang L (2021) Distinct Genomic Landscape of Colorectal Mucinous Carcinoma Determined via Comprehensive Genomic Profiling: Steps to a New Treatment Strategy. Front. Oncol. 11:603564. doi: 10.3389/fonc.2021.603564

Colorectal mucinous carcinoma (MC) is associated with inferior prognosis and response to treatment compared to adenocarcinoma (AC). The molecular landscapes of MC and adenocarcinoma with mucous composition (AMC) are not well-defined. We aimed to describe the genomic landscape of MC and AMC in a large colorectal cancer cohort. Tumor samples from patients with MC, AMC, or AC were analyzed using next-generation sequencing. MC had a molecular signature distinct from that of AC; genomic features were similar between AMC and MC but not between AMC and AC. *HER2* amplification and *TP53* and *APC* mutation rates were lower, whereas *SMAD4*, *PIK3CA*, *ACVR2A*, *KMT2D*, *LRP1*, *TGFBR2*, *GRIN2A*, *BRAF* V600E, *PTEN*, and *BRCA2* mutation rates were higher in MC than in AC. The mutation frequencies in MAPK, PI3K, and TGF-β pathways were higher, whereas those of cell cycle proteins and Wnt were lower in MC and AMC than in AC. The proportion of hypermutated tumors was significantly higher in MC and AMC than in AC. As MC has a distinct molecular signature from AC, immunotherapy can be potentially applied in treating MC. Similar molecular profiles of AMC and MC suggest that treatment strategies for MC, but not AC, can be used for AMC treatment.

Keywords: colorectal cancer, mucinous adenocarcinoma, adenocarcinoma with mucous composition, next-generation sequencing, hypermutated tumor

INTRODUCTION

According to the 2018 global cancer statistics released by the International Cancer Research Institute of the World Health Organization (WHO), colorectal cancer (CRC) has the third highest incidence rate and second highest mortality rate, and an increasing annual prevalence rate (1, 2). According to the WHO classification, mucinous carcinoma (MC) is a distinct pathological CRC subtype, with a substantial mucous component of more than 50% of the tumor volume, and accounts for 10–15% of all CRC cases (3, 4). MC constitutes a histological subtype with poor differentiation potential and is a predictive factor for poor prognosis (5, 6).

MC is clinically more prevalent among women, frequently located in the proximal colon, and associated with young age, high malignancy grade, tumor infiltration, lymph node metastasis, and peritoneal metastasis (4, 7). Compared with adenocarcinoma not otherwise specified (AC), patients with MC are reportedly less responsive to neoadjuvant radiotherapy and chemotherapy (8). The efficacy of first-line chemotherapy with oxaliplatin or irinotecan is lower among patients with advanced MC than among those with AC. Furthermore, patients with metastatic MC do not benefit from treatment with anti-epidermal growth factor (EGFR) monoclonal antibodies, even in cases with wild-type RAS and BRAF (9). Therefore, it is important to investigate the molecular characteristics of colorectal MC in detail and explore a more effective treatment strategy.

Colorectal MC has unique molecular characteristics. Most early studies focused on protein expression levels and reported that MUC2 and MUC5AC are upregulated in MC tumors (10, 11). Recent genomic analyses have reported that colorectal MC has a higher mutation frequency in Ras/MAPK and PI3K/Akt/mTOR pathways in MC than in AC, with a higher incidence of microsatellite instability (MSI), which is potentially associated with Lynch syndrome and the CpG island methylator phenotype (4). However, owing to limitations in detection technology, previous studies have not revealed the genetic landscape of MC, including comprehensive genomic characteristics, pathway analyses, and biomarkers for immunotherapy. The fraction of mucous composition varies substantially among Colorectal Cancers. Prior studies confirm that the variation of mucous composition in CRC is associated with distinct molecular and clinical features (12, 13). However, adenocarcinomas with relatively low mucous composition (less than 50%, also known as AMC) are usually diagnosed and treated as AC. The somatic mutational landscape of this unique subgroup is less known and the best clinical management of AMC needs to be addressed in the light of the mutational background (4).

In this study, we aimed to perform comprehensive targeted next-generation sequencing (NGS) to detect the two pathological subtypes of CRC, MC and AMC, and gain deep insights into their molecular characteristics, through the evaluation of the landscape of genetic alterations, pathway analysis, and analysis of biomarkers for immunotherapy to provide a molecular basis for the establishment of a precise treatment strategy for MC and AMC.

MATERIALS AND METHODS

Patients and Tumor Selection

Tumor specimens of patients with CRC involved in this study from January 2018 to September 2019 were sent for NGS analysis. Of 2,115 patients with CRC, 1,226 with a confirmed pathological diagnosis of MC, AMC, or AC were selected and recruited. Patients with an uncertain diagnosis of the pathological subtype or those with other special pathological subtypes, such as signet-ring cell carcinoma, undifferentiated carcinoma, and squamous cell carcinoma, were excluded. Of 2,115 patients with CRC, 1,226 with a confirmed pathological diagnosis of MC, AMC, or AC were selected and recruited. Patients with an uncertain diagnosis of the pathological subtype or those with other special pathological subtypes, such as signet-ring cell carcinoma, undifferentiated carcinoma, and squamous cell carcinoma, were excluded. MC was defined as extracellular mucus secretion accounting for >50% of the tumor volume. AMC was defined that accounted for ≤50%. And AC was defined as tumor with no extracellular mucus secretion. All tumor tissues were assessed independently by two experienced pathologists before sample disposal to pathologically confirm the diagnoses.

This study was approved by the Institution Review Board of the Sixth Affiliated Hospital of Sun Yat-sen University in accordance with the Declaration of Helsinki. Written informed consent was obtained from all enrolled patients.

NGS Analysis

NGS analysis was carried out at OrigiMed (Shanghai, China), a College of American Pathologists-accredited and Clinical Laboratory Improvement Amendments-certified laboratory, using a 450-gene comprehensive assay (14). At least 50 ng of DNA was extracted from each 40 mm formalin-fixed paraffin-embedded (FFPE) tumor sample using a DNA Extraction Kit (QIAamp DNA FFPE Tissue Kit) in accordance with the manufacturer's protocols. This panel encompassed all coding exons of 450 cancer-related genes and 64 selected introns of 39 genes that are frequently rearranged in solid tumors. Furthermore, the probe density was increased to ensure high capture efficiency in the conservatively low-read-depth regions. Peripheral blood was sampled from each patient as the normal control sample for genomic profiling. The genes were captured and sequenced with a mean coverage of 900× for FFPE samples and 300× for matched blood samples using an Illumina NextSeq 500 Platform (Illumina Incorporated, San Diego, CA, USA).

Genetic Analysis

All types of genetic alterations, including single-nucleotide variant (SNV), short and long indels, copy number alterations (CNAs), and gene rearrangement, were called using a suite of bioinformatics pipelines. Analysis of SNVs and indels began with the alignment of raw reads to the human genome reference sequence (hg19) with the Burrows-Wheeler Aligner (v0.62; BWA, Cambridge, MA, USA), followed by polymerase chain reaction (PCR) duplicates removal using MarkDuplicates algorithm from Picard (version 1.47; Cambridge, MA, USA). Local realignment and base quality recalibration for SNVs were performed using GATK (v3.1-1; Cambridge, MA, USA) and subsequently called by MUTECT (v1.7; Cambridge, MA, USA). The CNAs included: (1) amplification, defined as an increase in the number of gene segment copies by ≥8, and (2) homozygous deletion, defined as decrease of complete loss of gene segment copies in samples with 20% purity. To identify these alterations, tumor cellularity was estimated by allele frequencies of sequenced single-nucleotide polymorphisms (SNPs). For detection of gene rearrangement, aligned reads with abnormal insert size of 2,000 or zero bp were collected and used as

discordant reads, that is, paired-end reads that could not be closely mapped to a genome reference, with each read of paired reads aligned to the same chromosomes or different chromosomes. Originally, the discordant reads with the distance less than 500 bp formed clusters were further assembled by fermi-lite to identify potential rearrangement breakpoints. The breakpoints were double confirmed by BLAT, and the resulting chimeric gene candidates were annotated. For germline mutations, common single nucleotide polymorphisms, defined as those from the dbSNP database (Version 147), at a frequency of more than 1.5% from the Exome Sequencing Project 6500 (ESP6500), or at a frequency of more than 1.5% from the 1000 Genomes Project, were excluded. Furthermore, the variant allele frequency was adjusted with tumor purity estimated using FACETS.

Tumor Mutational Burden (TMB) and MSI

The TMB was estimated using the method of Chalmers et al. (15). In brief, the somatic, coding, base substitution, and short indel mutations were enumerated. Driver mutations and germline alterations in the dbSNP database were not enumerated. The TMB was determined by dividing the total number of mutations by the size of the coding region. The MSI status was determined in all cases. Based on the MSI score, samples were classified as MSI-high (MSI-H) and microsatellite stable (MSS).

Statistical Analyses

Qualitative variables were assessed using Fisher's exact test. Normally distributed quantitative data were analyzed using the *t*-test and non-normally distributed data were analyzed using the Wilcoxon rank test. All tests were two-tailed and significance was defined as a *P* value less than 0.05. All statistical analyses were performed using R software (Version 3.4.2).

RESULTS

Clinical Characteristics

We defined MC as adenocarcinoma with mucous composition greater than 50% and AC as adenocarcinoma with no mucous composition. Adenocarcinoma with mucous composition but less than 50% is called AMC. **Table 1** summarizes the characteristics of the patients. In total, 1,226 patients with CRC were enrolled in the study and divided into three categories by histological subtype: MC (10.5%), and AMC (8.2%), and AC (81.3%). The median age of patients with MC was less than that of patients with AC (56 *vs.* 59 years, $P = 0.037$), and the incidence of MC in the right colon was higher than that of AC (41.9 *vs.* 24.2%, $P < 0.001$). Patients with MC accounted for a larger proportion of patients with stage III CRC (44.2%) than AC (29.9%, $P < 0.001$) and AMC (29.0%, $P = 0.091$); but for AMC, the difference is only marginally significant. Furthermore, AMC was significantly more common in the right colon than AC was (50 *vs.* 24.2%, $P < 0.001$). No significant differences were observed between AMC and AC with respect to other clinical features.

Comparison of Common Gene Mutations Among MC, AMC, and AC

Comprehensive targeted NGS revealed that the top 10 prevalent mutations in MC were *KRAS* (55.8%), *TP53* (53.5%), *APC* (46.5%), *SMAD4* (34.1%), *ACVR2A* (28.7%), *PIK3CA* (28.7%), *KMT2D* (22.5%), *LRP1* (21.7%), *TGFBR2* (20.2%), and *ARID1A* (19.4%) (**Figure 1A**). In general, the mutation profile of MC was different from that of AC (**Figure 1B**); however, the mutation profiles of MC and AMC did not differ significantly (data not shown). Among the commonly mutated genes in CRC, *TP53* (53.5 *vs.* 79.5%, $P < 0.001$) and *APC* (46.5 *vs.* 75.1%, $P < 0.001$) displayed a significantly lower mutation rate, whereas *SMAD4* (34.1 *vs.* 19.1%, $P < 0.001$), *PIK3CA*

TABLE 1 | Patient and tumor characteristics.

Characteristics	MCN = 129 (%)	AMCN = 100 (%)	ACN = 997 (%)	*P* value MC *vs.* AC	*P* value AMC *vs.* AC	*P* value MC *vs.* AMC
Gender				1.000	0.134	0.227
Female	50 (38.8)	47(47)	390 (39.1)			
Male	79 (61.2)	53(53)	607 (60.9)			
Age				0.037	0.399	0.043
Median	56	62	59			
Range	17–86	28–82	16–96			
Primary Tumor Site				<0.001	<0.001	0.016
Left colon	55 (42.6)	25 (25)	320 (32.1)			
Right colon	54 (41.9)	50 (50)	241 (24.2)			
Rectum	20 (15.5)	24 (24)	427 (42.8)			
NA	0 (0)	1 (1)	9 (0.9)			
Stage at diagnosis[a]				<0.001	0.133	0.091
Stage I	3 (2.3)	5 (5)	58 (5.8)			
Stage II	40 (31)	33 (33)	232 (23.3)			
Stage III	57 (44.2)	29 (29)	298 (29.9)			
Stage IV	27 (20.9)	32 (32)	367 (36.8)			
NA	2 (1.6)	1 (1)	42 (4.2)			
Sample Source				0.415	0.622	0.177
Primary lesion	126 (97.6)	100 (100)	984 (98.7)			
Metastatic lesion	3 (2.4)	0 (0)	13 (1.3)			

[a]*Stage at diagnosis based on AJCC (8th edition). MC, mucinous carcinoma; AMC, adenocarcinoma with mucous composition; AC, adenocarcinoma; AJCC, American Joint Committee on Cancer; NA, not applicable.*

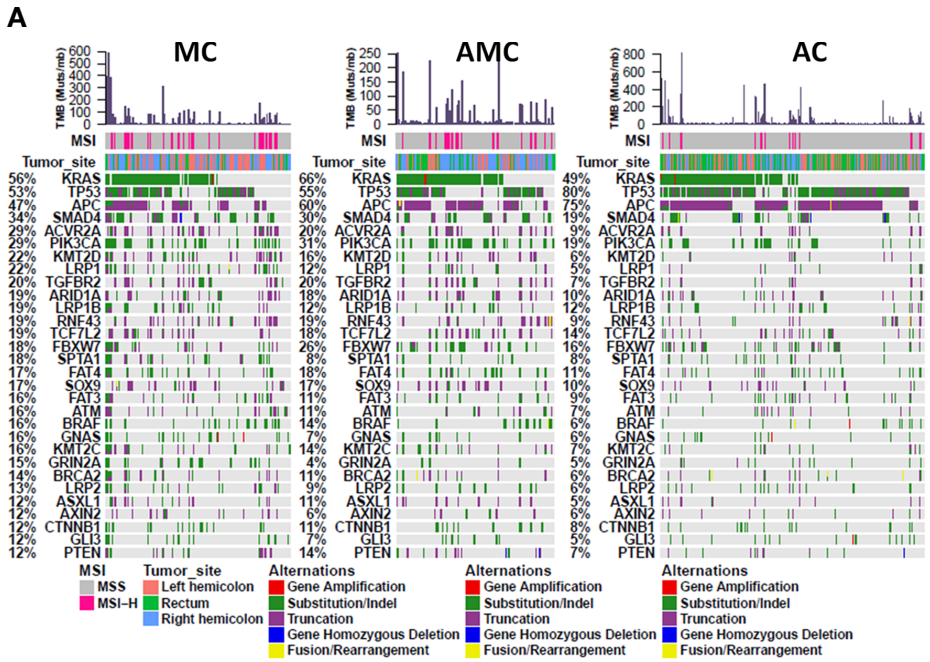


FIGURE 1 | Comparison of common gene mutations in MC, AMC, and AC. **(A)** Genomic landscape showing mutated genes among MC, AMC, and AC. Each column denotes an individual tumor and each row represents the MSI status, tumor site, and individual genes. Colors indicate the type of genetic alterations as indicated in the legend. **(B)** Analysis of the gene alteration frequency showing a higher gene alteration frequency in the MC group. MC, mucinous carcinoma; AMC, adenocarcinoma with mucous composition; AC, adenocarcinoma; MSI, microsatellite instability.

(28.7 *vs.* 19.2% $P = 0.014$), *ACVR2A* (28.7 *vs.* 9.1%, $P < 0.001$), *KMT2D* (22.5 *vs.* 6.4%, $P < 0.001$), *LRP1* (21.7 *vs.* 4.8%, $P < 0.001$), *TGFBR2* (20.2 *vs.* 6.4%, $P < 0.001$), and *GRIN2A* (14.7 *vs.* 4.5%, $P < 0.001$) displayed a significantly higher mutation rate in MC than in AC.

Furthermore, pathway analysis revealed that the mutation frequencies in MAPK, PI3K, and TGFβ pathways were higher, whereas those of cell cycle proteins and the Wnt pathway were lower, in MC and AMC than in AC (**Figure 2**).

Comparison of Clinically Actionable Alterations Among MC, AMC, and AC

The mutation pattern of clinically actionable alterations in MC was different from that in AC but similar to that in AMC. The mutation rates of *BRAF* V600E (10.9 *vs.* 3.3%, $P < 0.001$), *PIK3CA* (28.7 *vs.* 19.2%, $P = 0.014$), *PTEN* (14.7 *vs.* 7.2%, $P = 0.027$), and *BRCA2* (17.8 *vs.* 5.5%, $P < 0.001$) were significantly higher in MC than in AC. Although *HER2* mutation rates were comparable between MC and AC (3.9 *vs.* 6.2%, $P = 0.423$), *HER2* amplification occurred at a rate of 2.1% in AC but was not detected in MC or AMC. The mutation rate of KRAS was significantly higher in AMC than in AC (65.0 *vs.* 49.2%, $P = 0.001$); however, it did not significantly differ between MC and AC or MC and AMC. The mutation frequencies of clinically actionable genes in MC, AMC, and AC are summarized in **Table 2**.

Gene fusions in receptor tyrosine kinases have been recently identified as druggable targets in CRC (16). One patient with an MC tumor in the right colon harbored an *ETV6-NTRK3* fusion and the tumor was identified as MSI-H. *NCOA4-RET* and *FGFR2-PIBF1* fusions were observed in patients with MC and AMC, respectively. The frequency of druggable fusions did not significantly differ among the three CRC pathological subtypes.

Comparison of Immune Biomarkers in MC, AMC, and AC

We defined hypermutated tumors as MSI-H tumors or those harboring *POLE* mutations that result in a dramatic TMB elevation. In general, the proportion of hypermutated tumors was significantly higher in MC and AMC than in AC (MC 27.9% *vs.* AC 8.4%, $P < 0.001$; AMC 18% *vs.* AC 8.4%, $P = 0.003$).

The percentage of MSI-H tumors was significantly higher in MC and AMC than in AC (MC 22.5% *vs.* AC 6.8%, $P < 0.001$; AMC 17% *vs.* AC 6.8%, $P = 0.001$) and comparable between MC and AMC. Although the percentage of all *POLE* mutations among the three subtypes did not differ significantly, the proportion of *POLE* mutations resulting in a high TMB in MSS tumors was significantly higher in MC than in AC (5.4 *vs.* 1.6%, $P = 0.004$). The median TMB and median number of somatic mutations were also significantly higher in MC and AMC than in AC (**Figure 3** and **Table 3**).

Hypermutated Tumors in MC

We further evaluated the relevant immunotherapy indicators in MC, which revealed that 29 of 129 patients harbored MSI-H tumors, among which three harbored *POLE* mutations. Seven of 100 patients harbored MSS tumors with *POLE* mutations (**Figure 4**). Moreover, all *POLE* mutations detected in MSS tumors were located in the exonuclease domain, which led to extremely high levels of TMB. Only one E972G mutation in an MSS tumor was not located in the exonuclease domain and the TMB in this case was relatively lower (79.5 muts/Mb) than that in cases of *POLE* mutations in the exonuclease domain (TMB range, 121.1–595.5 muts/Mb). Furthermore, *POLE* mutations detected in three patients harboring MSI-H tumors were not present in the exonuclease domain. Details of clinical and molecular characteristics of MC with *POLE* mutations are summarized in **Table 4**.

DISCUSSION

This study identified the comprehensive genomic features of MC and AMC by targeted NGS using a large cohort of patients with CRC. We comprehensively compared genetic differences

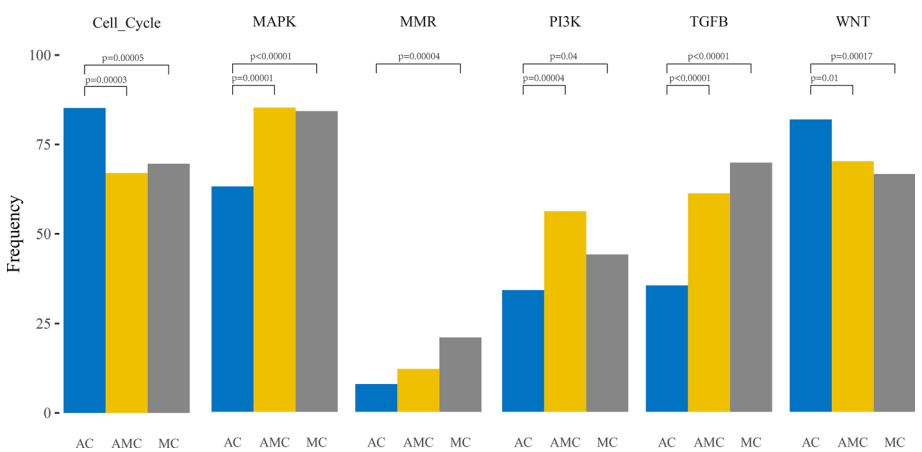


FIGURE 2 | Comparative analysis of the frequencies of genetic alterations based on signaling pathways playing an important role in oncogenesis in colorectal cancer.

TABLE 2 | Comparison of clinically actionable gene alterations in MC, AMC, and AC.

Genes	MCN = 129 (%)	AMCN = 100 (%)	ACN = 997 (%)	P value MC *vs.* AC	P value AMC *vs.* AC	P value MC *vs.* AMC
KRAS	72 (55.8%)	65 (65.0%)	491 (49.2%)	0.135	0.001	0.135
NRAS	4 (3.1%)	2 (2.0%)	36 (3.6%)	1.000	0.570	0.698
VEGFA	2(1.6%)	3(3.0%)	13(1.3%)	0.687	0.173	0.656
EGFR	9(7.0%)	2(2.0%)	100 (10.0%)	0.168	0.231	0.072
BRAF V600E	14 (10.9%)	8 (8.0%)	33 (3.3%)	<0.001	0.027	0.507
BRAF non-V600E	6 (4.7%)	5 (5.0%)	22 (2.2%)	0.123	0.091	1.000
HER2 (ERBB2) amplification	0 (0.0%)	0 (0.0%)	21 (2.1%)	0.158	0.248	NA
HER2 (ERBB2) mutation	5 (3.9%)	4 (4%)	62 (6.2%)	0.423	0.508	1.000
All druggable fusion	2 (1.6%)	1 (1%)	17 (1.7%)	1.000	1.000	1.000
NTRK1 fusion	0 (0%)	0 (0%)	5 (0.5%)	0.039	0.689	NA
NTRK3 fusion	1 (0.8%)	0 (0%)	1 (0.1%)	0.036	1.000	0.506
PIK3CA	37 (28.7%)	31 (31.0%)	191 (19.2%)	0.014	0.008	0.771
AKT1	5 (3.9%)	3 (3%)	22 (2.2%)	0.224	0.491	1.000
PTEN	19 (14.7%)	14 (14%)	72 (7.2%)	0.027	0.013	0.844
BRCA1	6 (4.7%)	5 (5%)	19 (1.9%)	0.057	0.06	1.000
BRCA2	23 (17.8%)	9 (9%)	55 (5.5%)	<0.001	0.043	0.553

MC, mucinous carcinoma; AMC, adenocarcinoma with mucous composition; AC, adenocarcinoma; NA, not applicable.

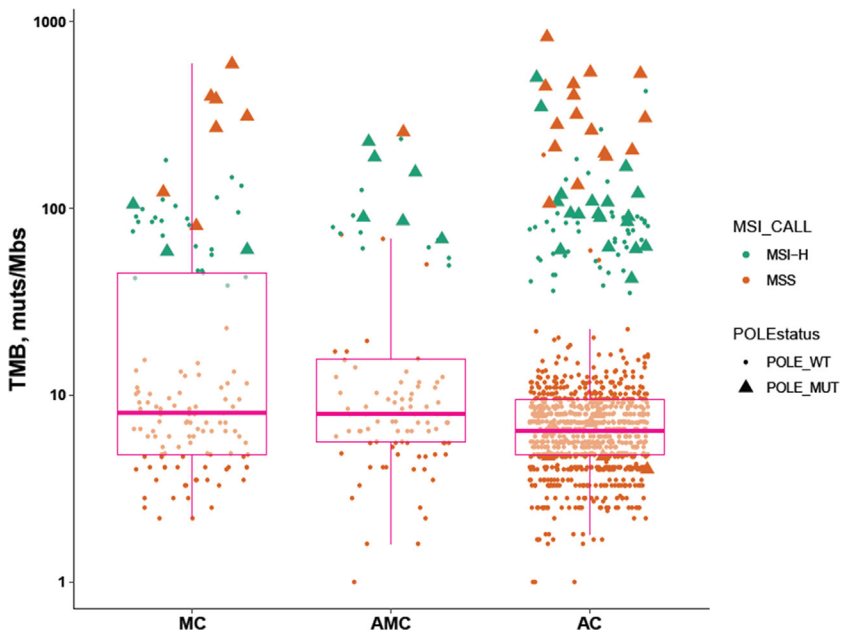


FIGURE 3 | Comparison of immunotherapy-related biomarkers in MC, AMC, and AC. MC, mucinous carcinoma; AMC, adenocarcinoma with mucous composition; AC, adenocarcinoma; MSI, microsatellite instability; TMB, tumor mutational burden; MSS, microsatellite stability.

between MC, AMC, and AC and identified the following major features. In general, MC had a molecular signature that was distinct from that of AC. The genomic features were similar between AMC and MC but different between AMC and AC. MC had a distinguished mutation pattern for prevalent gene mutations and biomarkers used clinically for CRC. Most importantly, the proportion of hypermutated tumors in MC and AMC was significantly higher than that in AC, indicating the higher applicability of immunotherapy for patients with these histological subtypes. Our results support developing more

TABLE 3 | Comparison of immunotherapy-related biomarkers in MC, AMC, and AC.

Characteristics	MCN = 129 (%)	AMCN = 100 (%)	ACN = 997 (%)	P value MC *vs.* AC	P value AMC *vs.* AC	P value MC *vs.* AMC
Hypermutated tumor[a]	36 (27.9)	18 (18.0)	84 (8.4)	<0.001	0.003	0.086
MSI-H tumor	29 (22.5)	17 (17.0)	68 (6.8)	<0.001	0.001	0.324
POLE	10 (7.8)	7 (7.0)	43 (4.3)	0.117	0.209	1.000
ALL mutation						
POLE	7 (5.4)	1 (1)	16 (1.6)	0.004	1.000	0.074
Hypermutation in MSS tumor[b]						
TMB				<0.001	<0.001	0.967
Median (muts/Mb)	7.0	6.9	5.4			
Range	1.2–591.5	0–254.7	0–825.3			
Somatic mutations number				0.002	0.003	0.963
Median (N/tumor)	9	9	8			
Range	2–277	1–160	1–269			

[a]*Hypermutated tumors are defined as MSI-H tumors or tumors harboring POLE mutations, resulting in drastic TMB elevation.*
[b]*Hypermutation in MSS tumors associated with POLE-mutated cases with dramatic TMB elevation in MSS CRC, mostly caused by POLE mutations in the exonuclease domain.*
MC, mucinous carcinoma; AMC, adenocarcinoma with mucous composition; AC, adenocarcinoma; CRC, colorectal cancer; MSI, microsatellite instability; TMB, tumor mutational burden; MSS, microsatellite stability.

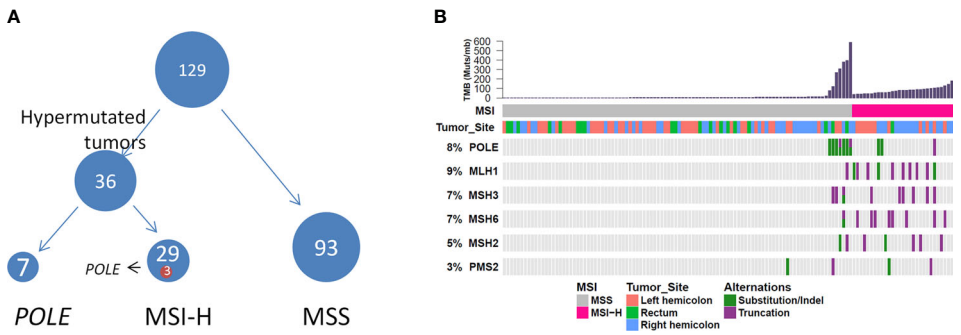


FIGURE 4 | **(A)** Hypermutated tumors include MSI-H tumors and tumors harboring *POLE* mutations. **(B)** Genomic landscape showing associations among MSI-H, microsatellite instability, high TMB, MSI status, and *POLE* mutation. TMB, tumor mutational burden.

TABLE 4 | Clinical and molecular characteristics of MC with *POLE* mutations.

Number	Gene	Sex	Age(years)	Primary tumor site	Stage	TMB(muts/Mb)	MSI	Variation type	DNA change	Amino acid change
1	*POLE*	female	51	Right colon	I	57.7	MSI-H	Substitution	c.5648C>T	p.A1883V
2	*POLE*	male	43	Right colon	III	58.9	MSI-H	Substitution	c.557C>T	p.A186V
3	*POLE*	female	40	Left colon	II	104	MSI-H	Truncation	c.4337_4338del	P.V1446Gfs*3
4	*POLE*	female	37	Rectum	III	396	MSS	Substitution	c.1231G>T	p.V411L
5	*POLE*	female	59	Left colon	II	268.6	MSS	Substitution	c.857C>G	p.P286R
6	*POLE*	female	41	Rectum	III	121.1	MSS	Substitution	c.1231G>C	p.V411L
7	*POLE*	female	56	Right colon	II	79.5	MSS	Substitution	c.2915A>G	p.E972G
8	*POLE*	male	47	Right colon	II	383.4	MSS	Substitution	c.1231G>T	p.V411L
9	*POLE*	female	47	Right colon	II	591.5	MSS	Substitution	c.857C>G	p.P286R
10	*POLE*	male	76	Left colon	I	309.5	MSS	Substitution	c.857C>G	p.P286R

MC, mucinous carcinoma; TMB, tumor mutational burden; MSI, microsatellite instability.

tailored treatment strategies for patients with CRC according to an individual's histological subtype.

Previous studies have suggested that mutations in *SMAD4*, *GNAS*, *BRAF*, and *KRAS* occur at high frequencies in MC, whereas *TP53*, *APC*, and *NRAS* mutations are less common (17, 18). The high frequency of *BRAF* mutations in MC is well-documented in the literature and supported by our findings (19, 20). Patients with metastatic CRC harboring a *BRAF* V600E mutation have a significantly worse prognosis. This study found the *BRAF* V600E mutation rate was significantly higher in MC and AMC than in AC, whereas the mutation rate of *BRAF* (non-V600E) did not significantly differ among the three groups. The *SMAD4* mutation frequency was significantly higher in MC and AMC than in AC. Patients with a *SMAD4* deletion have worse relapse-free survival and are resistant to chemotherapy with 5-fluorouracil (21). MCs were associated with an unsatisfactory

response to neoadjuvant chemotherapy. However, whether *SMAD4* plays a role in chemotherapy resistance mechanisms needs further research. On the other hand, the stage of the cancer is significantly associated with the frequency of specific mutation. For example, *BRAF* (V600E) is more frequent in high-stage MC. The observation suggests that the variation in the mutation rates among the three cancer types is attributed to the different clonal evolution processes, from which MC arises as a unique subtype.

A recent study reported that approximately 5% of patients with CRC harbor a *HER2* mutation. In patients with CRC, nearly half of *HER2* alterations are mutation rather than amplification or protein overexpression. Herein, *HER2* amplification was not observed in MC. However, a proportion of patients with MC harbored *HER2* mutations. Previous animal experiments have reported that the growth of implanted tumors harboring mutant *HER2* can be inhibited by HER2 inhibitors, including trastuzumab, lapatinib, and afatinib, alone and in combination with trastuzumab and tyrosine kinase inhibitors (22–24).

Immune checkpoint inhibitors (ICIs) have recently been widely used in solid and hematological malignancies (25). We defined hypermutated tumors as MSI-H tumors or those harboring *POLE* mutations that result in a dramatic TMB elevation, as there is robust evidence for MSI-H and *POLE* mutations as predictive biomarkers for a good response to immunotherapy in CRC (26, 27). Pembrolizumab has been approved for treating solid tumors with MSI-H/deficient mismatch repair (dMMR) and nivolumab ± ipilimumab has been approved for treating advanced CRC with MSI-H/dMMR (28, 29). Recently, a study on neoadjuvant treatment of CRC was conducted using a combinatorial treatment with an anti-PD-1 antibody and anti-CTLA-4 antibody. The treatment resulted in a pathological response in 20/20 patients and primary pathological remission in 19/20 patients with dMMR tumors (30). MC is significantly more likely to be associated with MSI-H in the colon and rectum (20). The proportion of MSI-H tumors in this study was significantly higher in MC and AMC than in AC, suggesting that immunotherapy is suitable for a larger proportion of patients with MC and AMC. MC was more prevalent in stage III CRC in this study, indicating that patients are more likely to develop local lymph node metastasis and present locally advanced CRC. In some cases of locally advanced CRC, it is challenging for surgeons to perform R0 (margin-negative) resection, which results in a worse prognosis for patients. ICIs in a neoadjuvant setting would be an effective treatment alternative for patients with MC with MSI-H/dMMR; thus, it is necessary to clarify the MSI/MMR status before any treatment.

Immunotherapy in MSS CRC tumors still lacks efficacy; therefore, there is an urgent need to identify biomarkers for immunotherapy in MSS tumors. Hypermutation in MSS CRCs is often associated with *POLE* mutations accompanied by dramatic TMB elevation, owing to the loss of DNA replication fidelity caused by *POLE* mutations in the exonuclease domain (27). Wang et al. summarized the *POLE/POLD1* mutation rate in 47,721 patients with different cancer types and identified that patients harboring *POLE/POLD1* mutations have a significantly higher TMB. When adjusting for cancer types and MSI status for multivariate Cox regression analysis, *POLE/POLD1* mutations were found to be independent risk factors for identifying patients that could benefit from ICI treatment (27). In this study, the frequency of *POLE* mutations resulting in high TMB in MSS tumors was significantly higher in MC than in AC. In addition, the proportion of hypermutated tumors (MSI-H or *POLE* mutations) was 27.9% in MC, suggesting that up to 30% of patients with CRC MC may benefit from immunotherapy. Furthermore, the mutation pattern of *POLE* differed between MSS and MSI-H tumors and *POLE* mutations occurring in the exonuclease domain markedly increased the TMB in MSS tumors.

We acknowledge that the current study had several limitations. First, the retrospective study design could not exclude selection bias. Second, the clinical data of patients' treatments and outcomes were not controlled and collected in the current study; therefore, the clinical impacts of our findings need further confirmation. Finally, on the potential benefit of ICI, the effects of tumor infiltrating lymphocytes need to be evaluated along with the mutational profile as well as MSI and POLE statuses; whereas in the current study still lacks the pathological data for tumor immune microenvironment. On the other hand, the somatic mutational landscape is also affected by the host immune environment, hence serve as a proxy to the activities of the immune cells.

In spite of the limitations, using the large cohorts of MC (n = 129) and AMC (n = 100) *via* a comprehensive targeted NGS panel, our results reveal the molecular landscapes of MC, AMC, and AC, which could lead to tailored treatment for different histological subtypes of CRC. The selection of baseline clinical and pathological characteristics was relatively intact in this study, allowing the analysis of clinical and genomic features. Our findings shed new light on the treatment and management of patients with MC and AMC. Further prospective studies in patients with MC and AMC are warranted to validate our findings, especially regarding the potential use of immunotherapy.

CONCLUSIONS

We identified a distinct genomic landscape in colorectal MC *via* comprehensive genomic profiling for commonly mutated and clinically actionable genes. Hypermutated tumors account for nearly 30% of MC, suggesting that a large proportion of patients with MC may benefit from immunotherapy; therefore, there is a need for comprehensive molecular testing in these patients. AMC has similar genomic features to MC but different from AC, suggesting the potential for the use of MC treatment strategies for treating AMC.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the CNGB Sequence Archive, repository, accession number is CNP0001753.

The reviewer link is http://db.cngb.org/cnsa/review/show/CNP0001753_20210422_a35ddad7.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institution Review Board of the Sixth Affiliated Hospital of Sun Yat-Sen University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Study design: LH, LK, SZ, and TL. Study recruitment, clinical sample, and data acquisition: XZ, YC, FX, HH, ZZ, SZ, and TL. Bioinformatic analysis: WW and FW. Primary results interpretation: LH and SL. Manuscript drafting: SL, LH, LK, and SZ. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Fundamental Research Funds for the Central Universities, grant number 16ykjc25, Sun Yat-Sen University Clinical Research 5010 Program, grant number 2016005 and Province Natural Science Fund of Guangdong (2018A030313621).

ACKNOWLEDGMENTS

We thank Jian Wang and Jicheng Yao from OrigiMed for their contributions to the bioinformatic analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.603564/full#supplementary-material

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) 68:394–424. doi: 10.3322/caac.21492
2. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal Cancer. *Lancet* (2019) 394:1467–80. doi: 10.1016/S0140-6736(19)32319-0
3. Hugen N, Brown G, Glynne-Jones R, de Wilt JH, Nagtegaal ID. Advances in the Care of Patients With Mucinous Colorectal Cancer. *Nat Rev Clin Oncol* (2016) 13:361–9. doi: 10.1038/nrclinonc.2015.140
4. Luo C, Cen S, Ding G, Wu W. Mucinous Colorectal Adenocarcinoma: Clinical Pathology and Treatment Options. *Cancer Commun (Lond)* (2019) 39:13. doi: 10.1186/s40880-019-0361-0
5. Hugen N, van Beek JJ, de Wilt JH, Nagtegaal ID. Insight Into Mucinous Colorectal Carcinoma: Clues From Etiology. *Ann Surg Oncol* (2014) 21:2963–70. doi: 10.1245/s10434-014-3706-6
6. Catalano V, Loupakis F, Graziano F, Torresi U, Bisonni R, Mari D, et al. Mucinous Histology Predicts for Poor Response Rate and Overall Survival of Patients With Colorectal Cancer and Treated With First-Line Oxaliplatin- and/or Irinotecan-Based Chemotherapy. *Br J Cancer* (2009) 100:881–7. doi: 10.1038/sj.bjc.6604955
7. Li C, Zheng H, Jia H, Huang D, Gu W, Cai S, et al. Prognosis of Three Histological Subtypes of Colorectal Adenocarcinoma: A Retrospective Analysis of 8005 Chinese Patients. *Cancer Med* (2019) 8:3411–9. doi: 10.1002/cam4.2234
8. Simha V, Kapoor R, Gupta R, Bahl A, Nada R. Mucinous Adenocarcinoma of the Rectum: A Poor Candidate for Neo-Adjuvant Chemoradiation? *J Gastrointest Oncol* (2014) 5:276–9. doi: 10.3978/j.issn.2078-6891.2014.020
9. Moretto R, Morano F, Ongaro E, Rossini D, Pietrantonio F, Casagrande M, et al. Lack of Benefit From Anti-EGFR Treatment in RAS and BRAF Wild-Type Metastatic Colorectal Cancer With Mucinous Histology or Mucinous Component. *Clin Colorectal Cancer* (2019) 18:116–24. doi: 10.1016/j.clcc.2019.02.007
10. Hugen N, Simons M, Halilovic A, van der Post RS, Bogers AJ, Marijnissen-van Zanten MA, et al. The Molecular Background of Mucinous Carcinoma Beyond MUC2. *J Pathol Clin Res* (2015) 1:3–17. doi: 10.1002/cjp2.1
11. Verhulst J, Ferdinande L, Demetter P, Ceelen W. Mucinous Subtype as Prognostic Factor in Colorectal Cancer: A Systematic Review and Meta-Analysis. *J Clin Pathol* (2012) 65:381–8. doi: 10.1136/jclinpath-2011-200340
12. Ogino S, Brahmandam M, Cantor M, Namgyal C, Kawasaki T, Kirkner G, et al. Distinct Molecular Features of Colorectal Carcinoma With Signet Ring Cell Component and Colorectal Carcinoma With Mucinous Component. *Mod Pathol* (2006) 19:59–68. doi: 10.1038/modpathol.3800482
13. Tozawa E, Ajioka Y, Watanabe H, Nishikura K, Mukai G, Suda T, et al. Practice, Mucin Expression, p53 Overexpression, and Peritumoral Lymphocytic Infiltration of Advanced Colorectal Carcinoma With Mucus Component: Is Mucinous Carcinoma a Distinct Histological Entity? *Pathol Res Pract* (2007) 203:567–74. doi: 10.1016/j.prp.2007.04.013
14. Cao J, Chen L, Li H, Chen H, Yao J, Mu S, et al. An Accurate and Comprehensive Clinical Sequencing Assay for Cancer Targeted and Immunotherapies. *Oncologist* (2019) 24:e1294–302. doi: 10.1634/theoncologist.2019-0236
15. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, et al. Analysis of 100,000 Human Cancer Genomes Reveals the Landscape of Tumor Mutational Burden. *Genome Med* (2017) 9:34. doi: 10.1186/s13073-017-0424-2
16. Garcia-Aranda M, Redondo M. Targeting Receptor Kinases in Colorectal Cancer. *Cancers (Basel)* (2019) 11:433–57. doi: 10.3390/cancers11040433
17. Khan M, Loree JM, Advani SM, Ning J, Li W, Pereira AAL, et al. Prognostic Implications of Mucinous Differentiation in Metastatic Colorectal Carcinoma can Be Explained by Distinct Molecular and Clinicopathologic Characteristics. *Clin Colorectal Cancer* (2018) 17:e699–709. doi: 10.1016/j.clcc.2018.07.005
18. Kim H, Kim BH, Lee D, Shin E. Genomic Alterations in Signet Ring and Mucinous Patterned Colorectal Carcinoma. *Pathol Res Pract* (2019) 215:152566. doi: 10.1016/j.prp.2019.152566
19. Tanaka H, Deng G, Matsuzaki K, Kakar S, Kim GE, Miura S, et al. BRAF Mutation, CpG Island Methylator Phenotype and Microsatellite Instability Occur More Frequently and Concordantly in Mucinous Than non-Mucinous Colorectal Cancer. *Int J Cancer* (2006) 118:2765–71. doi: 10.1002/ijc.21701
20. Reynolds IS, Furney SJ, Kay EW, McNamara DA, Prehn JHM, Burke JP. Meta-Analysis of the Molecular Associations of Mucinous Colorectal Cancer. *Br J Surg* (2019) 106:682–91. doi: 10.1002/bjs.11142
21. Wasserman I, Lee LH, Ogino S, Marco MR, Wu C, Chen X, et al. Smad4 Loss in Colorectal Cancer Patients Correlates With Recurrence, Loss of Immune Infiltrate, and Chemoresistance. *Clin Cancer Res* (2019) 25:1948–56. doi: 10.1158/1078-0432.CCR-18-1726
22. Greally M, Kelly CM, Cercek A. Her2: An Emerging Target in Colorectal Cancer. *Curr Probl Cancer* (2018) 42:560–71. doi: 10.1016/j.currproblcancer.2018.07.001
23. Siena S, Sartore-Bianchi A, Marsoni S, Hurwitz HI, McCall SJ, Penault-Llorca F, et al. Targeting the Human Epidermal Growth Factor Receptor 2 (HER2) Oncogene in Colorectal Cancer. *Ann Oncol* (2018) 29:1108–19. doi: 10.1093/annonc/mdy100
24. Cheng Z, Shao X, Xu M, Wang J, Kuai X, Zhang L, et al. Rab1A Promotes Proliferation and Migration Abilities Via Regulation of the HER2/AKT-independent mTOR/S6K1 Pathway in Colorectal Cancer. *Oncol Rep* (2019) 41:2717–28. doi: 10.3892/or.2019.7071

25. Gibney GT, Weiner LM, Atkins MB. Predictive Biomarkers for Checkpoint Inhibitor-Based Immunotherapy. *Lancet Oncol* (2016) 17:e542–51. doi: 10.1016/S1470-2045(16)30406-5
26. Schrock AB, Ouyang C, Sandhu J, Sokol E, Jin D, Ross JS, et al. Tumor Mutational Burden is Predictive of Response to Immune Checkpoint Inhibitors in MSI-high Metastatic Colorectal Cancer. *Ann Oncol* (2019) 30:1096–103. doi: 10.1093/annonc/mdz134
27. Wang F, Zhao Q, Wang YN, Jin Y, He MM, Liu ZX, et al. Evaluation of POLE and POLD1 Mutations as Biomarkers for Immunotherapy Outcomes Across Multiple Cancer Types. *JAMA Oncol* (2019) 5(10):1504–6. doi: 10.1001/jamaoncol.2019.2963
28. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, et al. Mismatch Repair Deficiency Predicts Response of Solid Tumors to PD-1 Blockade. *Science* (2017) 357:409–13. doi: 10.1126/science.aan6733
29. Overman MJ, Lonardi S, Wong KYM, Lenz HJ, Gelsomino F, Aglietta M, et al. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *J Clin Oncol* (2018) 36:773–9. doi: 10.1200/JCO.2017.76.9901
30. Chalabi M, Fanchi LF, Dijkstra KK, Van den Berg JG, Aalbers AG, Sikorska K, et al. Neoadjuvant Immunotherapy Leads to Pathological Responses in MMR-proficient and MMR-deficient Early-Stage Colon Cancers. *Nat Med* (2020) 26:566–76. doi: 10.1038/s41591-020-0805-8

Conflict of Interest: TL, FW and WW were employed by OrigiMed.

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.

frontiers in Oncology

ORIGINAL RESEARCH
published: 02 June 2021
doi: 10.3389/fonc.2021.641869

Correlation Between the Evolution of Somatic Alterations During Lymphatic Metastasis and Clinical Outcome in Penile Squamous Cell Carcinoma

Jian Cao[1], Chun-He Yang[2], Wei-Qing Han[1], Yu Xie[1], Zhi-Zhong Liu[1] and Shu-Suan Jiang[1]**

[1] Department of Urology, Hunan Cancer Hospital and The Affiliated Cancer Hospital of Xiangya Medicine School, Central South University, Changsha, China, [2] GloriousMed Clinical Laboratory Co., Ltd., Shanghai, China

OPEN ACCESS

Edited by:
Qiyuan Li,
Xiamen University, China

Reviewed by:
Rongshan Yu,
Xiamen University, China
Jinchun Xing,
Xiamen University, China

***Correspondence:**
Jian Cao
caojian@hnca.org.cn
Shu-Suan Jiang
jiangshusuan@hnca.org.cn

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *15 December 2020*
Accepted: *29 April 2021*
Published: *02 June 2021*

Citation:
Cao J, Yang C-H, Han W-Q, Xie Y, Liu Z-Z and Jiang S-S (2021) Correlation Between the Evolution of Somatic Alterations During Lymphatic Metastasis and Clinical Outcome in Penile Squamous Cell Carcinoma. Front. Oncol. 11:641869. doi: 10.3389/fonc.2021.641869

Penile squamous cell carcinoma (PSCC) is a rare malignancy with poor survival after standard treatment. Although genomic alterations of PSCC have been characterized in several latest studies, the association between the formation of somatic landscape and regional lymph node metastasis (LNM), an important predictor for patient survival, has not been comprehensively investigated. Here, we collected formalin-fixed paraffin-embedded tumor tissue and matched normal samples of 32 PSCC patients, including 14 LNM patients and 18 clinically node-negative patients, to implement a whole-exome sequencing. Comparison of genomic features among different lymph node status subgroups was conducted after genomic profiling and its effects on patient survival were explored. Top-ranked recurrent gene mutants in our PSCC cohort were *TP53* (13/32), *NOTCH1* (12/32), *CDKN2A* (11/32), *TTN* (9/32) and *FAT1* (8/32), mainly identified in the Notch, Hippo, cell cycle, TP53, RTK-RAS and PI3K pathways. While *CDKN2A* was confirmed to be the driver gene in all PSCC patients, certain gene mutants were significantly enriched in LNM involved patients, including *TP53* (9/14 *vs.* 4/18, $p = 0.029$) and *GBF1* (4/14 *vs.* 0/18, $p = 0.028$). Overall survival stratification of PSCC patients were found to be significantly correlated with mutations of three genes, including *PIK3CA* (Hazard ratio [HR] = 4.15, $p = 0.029$), *CHD7* (HR = 4.82, $p = 0.032$) and *LAMC3* (HR = 15.9, $p < 0.001$). *PIK3CA* and *LAMC3* held a higher prevalence in patients with LNM compared to those without LNM (*PIK3CA*: 3/14 *vs.* 1/18, *LAMC3*: 2/14 *vs.* 1/18). Our finding demonstrated that genomic divergence exists across PSCC patients with different lymph node statuses, and it may be correlated with their survival outcome. It helps delineate somatic evolution during tumor progression and perfect potential therapeutic intervention in this disease.

Keywords: penile squamous cell carcinoma, lymph node metastasis, somatic alteration, patient survival, whole-exome sequencing

INTRODUCTION

Penile squamous cell carcinoma (PSCC) is a rare cancer with a significantly higher incidence in developing countries compared to developed countries (1), mainly attributed to exposure to human papilloma virus (HPV) (2). For patients with advanced PSCC, standard treatment paradigm is a multimodal approach of chemotherapy combinations followed by surgical procedures (3–6). Unfortunately, more than half of the patients shortly progressed or relapsed after the treatment (6, 7).

To discover novel diagnostic and prognostic biomarkers that are capable to identify patients sensitive to specific therapy, genomic profiling of penile carcinoma has been examined (8–11) and a few PSCC cell lines were established (12). Those studies revealed that somatic alterations are associated with penile carcinogenesis, including frequent mutations in gene *TP53*, *CDKN2A*, *NOTCH1* and *PIK3CA* (13, 14). Associated risk factors were also investigated and prediction models of patient survival were developed in the past years to achieve better management of this malignancy (15–18). In all studies, lymph node involvement was found to be the most evidential factor (19) compared to other predictors including histological subtypes (20) and high expression levels of *TP53*-regulated inhibitor of apoptosis 1 (21). Although it is indicated that lymph node metastasis (LNM) could be roughly inferred from lymph node staging, lymph vascular invasion or sentinel lymph node biopsy combined with sonography (22–24), accurate prediction of lymph node status is still lacking and the connections between lymphatic metastasis and potential genetic biomarkers remain unclear (25).

To investigate the evolution of somatic alterations during the process that tumors transform to a state prone to spread to the lymph node, we characterized the somatic mutation landscape and compared genetic characteristics between PSCC patients with different lymph node statuses with whole-exome sequencing. The performance of predicting patient survival with relevant variants was also evaluated.

MATERIALS AND METHODS

Sample Cohort

Tumor tissue and matched normal blood or tissue samples of 32 PSCC patients were collected for whole-exome sequencing. These patients were diagnosed with PSCC from June 2015 to June 2019 in Hunan Cancer Hospital and underwent surgical resection afterward. Lymph node dissection was performed and the lymph node statuses were assessed in some patients. Clinical information including age, tumor stage, pathological type, lymph node status and survival information were gathered by reviewing the electronic medical records. The study was approved by the ethics committee of Hunan Cancer Hospital and all involved human subjects have signed the informed consent.

Whole-Exome Sequencing

DNA was extracted from formalin-fixed paraffin-embedded tissues and white blood cells using QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany) and Blood Genomic DNA Mini Kit (Cwbiotech, Beijing, China). The whole exome was captured according to the standard procedures of xGen Exome Research Panel v1.0 (Integrated DNA Technologies, Coralville, IA). The captured DNA fragments were then used for library preparation and quantification guided by KAPA Hyper Prep protocols (Kapa Biosystems, Wilmington, MA), followed by purification with AMPure XP (Beckman Coulter, Brea, CA) and quantification using Qubit™ dsDNA HS Assay Kit (ThermoFisher, Waltham, MA). Pooled library was finally sequenced using Novoseq6000 (Illumina, San Diego, CA).

Variant Calling and Annotation

After adapter trimming with Trimmomatic, the sequencing reads were then aligned to the human reference genome hg19 using Burrows-Wheeler Aligner (BWA). Reads were then realigned using Genome Analysis Tool Kit (GATK) after duplicated reads were flagged with Picard. Mutect2 was used to identify somatic mutations, which were then annotated with ANNOVAR. Human identity consistency of paired samples was verified using an in-house script. Somatic mutations were filtered out under the following conditions (1): base quality value under 20; (2) mutation reads depth less than 10; (3) variant allele frequency less than 5%; (4) variant supporting reads more than 4 or variant allele frequency above 2% in the paired normal sample. Then synonymous and benign mutations were removed from the remaining variants. OncodriveCLUSTL was used to detect significant clusters of variation across genomic regions to identify candidate driver genes (26). Visualization of gene alterations in oncogenic signaling pathways was conducted using the PathwayMapper tool (http://pathwaymapper.org).

HPV Genotyping

HPV status of PSCC patients was assessed by HPV genotyping (17 high risk HPV: 16, 18, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73 and 82; 6 low risk HPV: 6, 11, 42, 43, 81 and 83), which was performed with a polymerase chain reaction reverse dot blot (PCR-RDB) approach (Yaneng Bio, Shenzhen, China) using DNA extracted from tumor tissue samples.

Tumor Mutation Burden, Heterogeneity and Genomic Stability

Tumor mutational burden (TMB), heterogeneity and genomic stability were assessed to evaluate the genomic status of their tumor samples for each patient (27). TMB was defined as the number of nonsynonymous somatic mutations per million bases, and heterogeneity was estimated with mutant-allele tumor heterogeneity (MATH) calculated by R package maftools (28, 29). Genomic instability was represented by the weighted genome integrity index (wGII), which denotes the chromosome-weighted proportion of genomic fragments with abnormal copy number (30).

Statistical Analysis

All statistical analyses were performed with R v3.6.0. Prevalence comparison of gene mutant was conducted using Fisher's exact test. TMB, MATH and wGII among lymph node subgroups were compared with Wilcoxon rank-sum test. Kaplan-Meier estimate was implemented for survival analysis and the log-rank test was used to determine the mutated gene that correlated with patient survival. Hazard ratio (HR) was reported by the univariate and multivariable cox proportional hazard regression models. Two-sided P<0.05 was considered statistically significant.

RESULTS

Cohort Characteristics

32 PSCC patients including 14 (43.75%) lymph node-positive patients and 18 (56.25%) negative node patients were enrolled in our investigation, summarized in **Table 1**. The median age of this cohort was 53.5 years (41–78 years). Among all patients, 8 (25%) of them were diagnosed with low, low-to-moderate or moderate grade cancer while 24 (75%) of them were evaluated as moderate-to-high or high grade cancer. 14 (43.75%) patients were assessed at stage III or higher. 28 (87.5%) patients were tested for HPV genotyping, and 16 (50%) of them were found to be HPV-positive. 7 (21.88%) patients experienced metastases or relapse and 9 (28.13%) patients deceased during follow-up.

Somatic Mutation Landscape of PSCC in Chinese Patients

A total of 3,026 somatic mutations were identified in 2,418 genes, including single nucleotide variants and small insertions/deletions. The most common variant type was missense mutation (80.3%), followed by frameshift insertion/deletion (8.5%), nonsense mutation (7.3%), in-frame insertion/deletion (2.2%), splicing mutation (1.5%) and nonstop mutation (0.2%) (**Figure 1A**). *TP53* (13, 40.63%), *NOTCH1* (12, 37.50%), *CDKN2A* (11, 34.38%), *TTN* (9, 28.13%) and *FAT1* (8, 25.00%) were found to be the most common repeatedly mutated genes in this cohort (**Figure 1B**), which have been reported to be frequently mutated in penile carcinoma. *CASP8*, which was previously reported to be frequently altered in the Chinese PSCC population (11), was mutated in 6 (18.75%) patients in this study. Although *CSN1* mutant was reported in a Caucasian cohort (9), it was not found in any patient in our cohort. Other reported frequent gene mutants, like *PIK3CA* and *HRAS*, were mutated in a few patients but not among the top mutated genes.

Pathway Alterations and Driver Mutant

Somatic mutations of 10 commonly altered pathways in cancer were characterized and variants were found in all these pathways with varying frequencies (**Figure 2A**). Notch (20, 62.50%), Hippo (18, 56.25%), TP53 (15, 46.88%), cell cycle (13, 40.63%), RTK-RAS (12, 37.50%) and PI3K (7, 21.88%) were the most frequently mutated pathways. The prevalence of each pathway was contributed by different dominant gene mutants, as exemplified for *FAT1* in the Hippo pathway, *CDKN2A* in the cell cycle pathway and *HRAS* in the RTK-RAS pathway.

To further investigate tumorigenesis-associated pathways in penile carcinoma, somatic mutations were used to identify candidate driver genes with OncodriveCLUSTL, which has been proven to be a state-of-the-art method in the field of driver gene prediction. The only gene that showed significance was *CDKN2A* (adjusted $p < 0.001$, **Figure 2B**), whose two prevailing hotspot mutations enriched in its ankyrin repeat-containing domain (**Figure 2C**), indicating alterations in the cell cycle pathway may be involved in triggering this malignancy.

LNM-Associated Somatic Alterations Correlate With Patient Survival in Penile Carcinoma

Regional lymph node involvement was considered to be a key predictor of patient survival in PSCC. We assessed the overall survival (OS) by lymph node statuses in our cohort and confirmed that the positive lymph node was associated with

TABLE 1 | Patient characteristics.

	All patients (N = 32)	Positive lymph node (N = 14)	Negative lymph node (N = 18)
Age, median (range)	53.5 (41–78)	54.0 (46–66)	52.5 (41–78)
Grade			
Well	43.8% (14/32)	50.0% (7/14)	38.9% (7/18)
Well to moderate	31.3% (10/32)	28.8% (4/14)	33.3% (6/18)
Moderate	18.8% (6/32)	14.3% (2/14)	22.2% (4/18)
Moderate to poor	3.1% (1/32)	–	5.6% (1/18)
Poor	3.1% (1/32)	7.1% (1/14)	–
Stage			
0	18.8% (6/32)	-	33.3% (6/18)
I	15.6% (5/32)	-	27.8% (5/18)
II	21.9% (7/32)	-	38.9% (7/18)
III	18.8% (6/32)	42.9% (6/14)	–
IV	25.0% (8/32)	57.1% (8/14)	–
HPV status			
Negative	37.5% (12/32)	21.4% (3/14)	50.0% (9/18)
Positive	50.0% (16/32)	64.3% (9/14)	38.9% (7/18)
NA	12.5% (4/32)	14.3 (2/14)	11.1% (2/18)

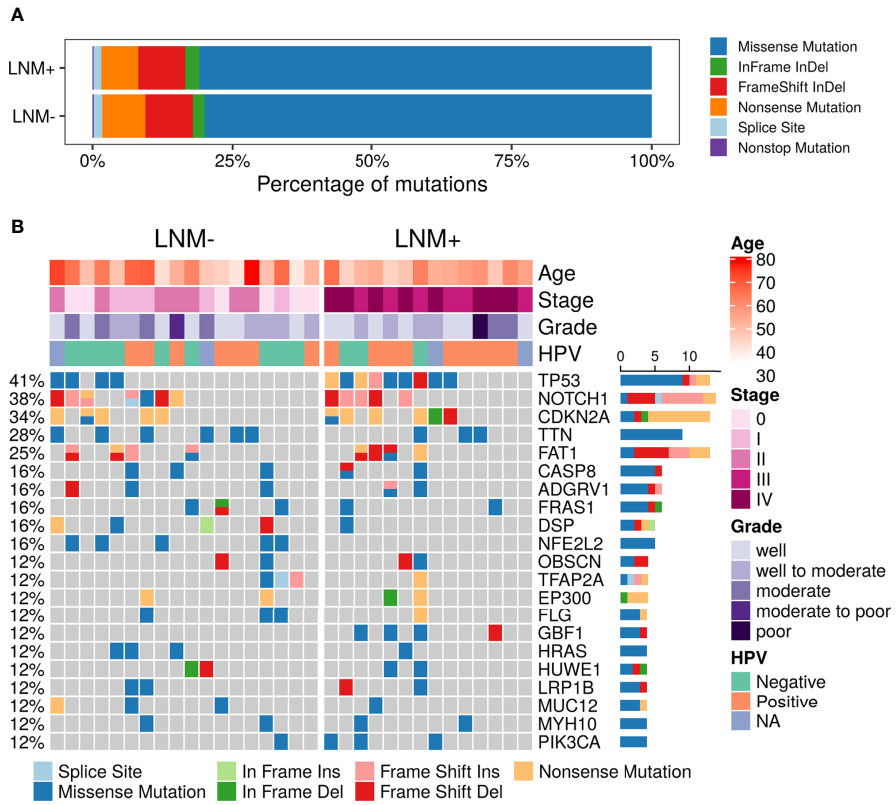


FIGURE 1 | Summary of somatic mutations in 32 PSCC patients. **(A)** The frequencies of different variant types in lymph node metastasis (LNM) involved patients and negative-node patients. **(B)** Mutational landscape of all 32 PSCC patients. Each row represents one gene while each column represents one patient. The frequencies of gene mutants and clinical characteristics are labeled by the side of the heatmap. (NA, not assessed).

shorter survival (HR = 4.92, log-rank p = 0.028; **Figure 3A**). Only 7 genes appeared in the intersection of the first 20 mutated genes in LNM positive and negative patients, including *NOTCH1*, *TTN*, *CDKN2A*, *FAT1*, *TP53*, *CASP8* and *AFDGRV1* (**Supplementary Figure 1**). Furthermore, the prevalence of mutated genes between different lymph node statuses was compared to explore the association between somatic variants and the LNM process in PSCC. The candidate genes for this comparison were limited to 58 genes that altered in at least 3 patients. Alterations of two genes, *TP53* (9/14 *vs.* 4/18, p = 0.029) and *GBF1* (4/14 *vs.* 0/18, p = 0.028), were found to be significantly enriched in lymph node-positive samples (**Figure 3B**), suggesting the occurrence of such genomic events during tumor progression may potentially promote regional lymph node metastasis. Although *NFE2L2* mutations tended to serve a protective role of LNM, no significant difference was observed (0/14 *vs.* 5/18, p = 0.052).

The associations between somatic mutations and survival outcome of PSCC patients were further investigated. Mutants of *TP53* and *GBF1*, which were significantly enriched in positive lymph node patients, indicated shorter OS but the difference was not statistically significant (*TP53*: HR = 1.27; *GBF1*: HR = 1.94; **Supplementary Figure 2**). However, fine performance of stratifying overall survival of patients was observed in three other genes, including *PIK3CA* (HR = 4.15, p = 0.029; **Figure 3C**), *CHD7* (HR = 4.82, p = 0.032; **Figure 3D**) and *LAMC3* (HR = 15.9, p < 0.001; **Figure 3E**). Furthermore, these genes together with age and LNM status, were included as covariates in cox multivariable regression to verify their significance. Independent associations with OS were confirmed in *CHD7* (HR = 29.4, p = 0.009) and *LAMC3* (HR = 11.9, p = 0.003), except for *PIK3CA* (HR = 3.65, p = 0.1) (**Supplementary Table 1**). Moreover, *PIK3CA* and *LAMC3*, held a higher frequency of mutation in patients with LNM but did not reach the significant level (*PIK3CA*: 3/14 *vs.* 1/18, *LAMC3*: 2/14 *vs.* 1/18; **Figure 3B**).

LNM-Related Somatic Alterations in Pathway Level

Further investigations were carried out by exploring correlations between somatic alterations and lymphatic metastasis in the pathway level. TP53 pathway is the only significantly enriched pathway in node-positive patients (p = 0.031, **Supplementary Figure 3**), which is mainly caused by the mutations in tumor suppressor gene *TP53* and *ATM* (**Figure 4A**). There also is a tendency that alterations in RAS pathway preferentially occurred in LNM negative patients (*HRAS* and *BRAF*, **Figure 4B**). In the cell cycle pathway (*CDKN2A* and *E2F3*, **Figure 4C**) and NRF2

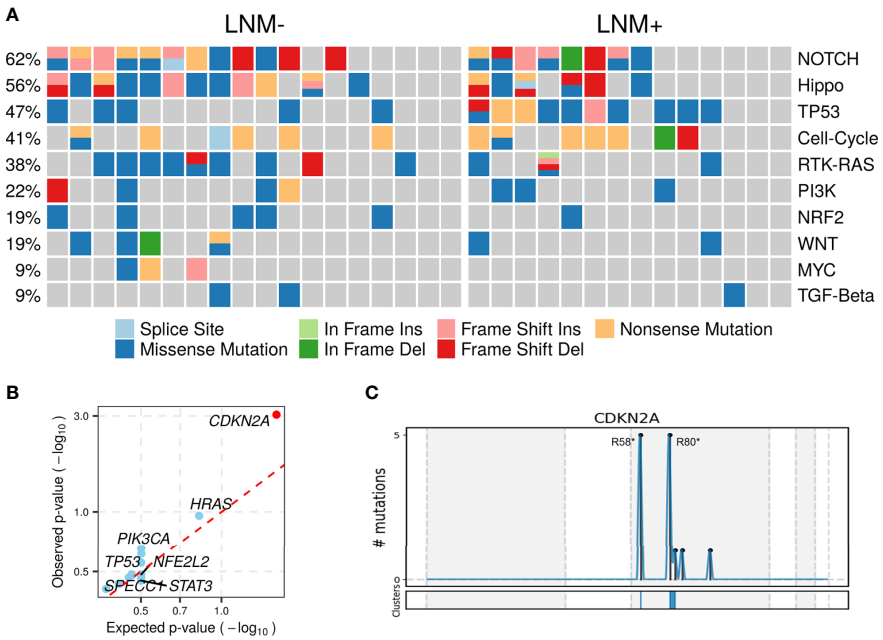


FIGURE 2 | Altered pathways and diver mutant in penile carcinoma. **(A)** The frequencies of ten common cancer-related pathways altered in 32 PSCC patients, **(B)** PSCC candidate driver genes identified by OncodriveCLUSTL. Significant gene (observed p-value < 0.01) is highlighted with red circle. **(C)** Distribution of mutations across CDKN2A region in 32 PSCC patients. The mutations, mainly two hotspots labeled in the figure, are enriched in two clusters (shown at the bottom) that span 1 base and 16 bases respectively.

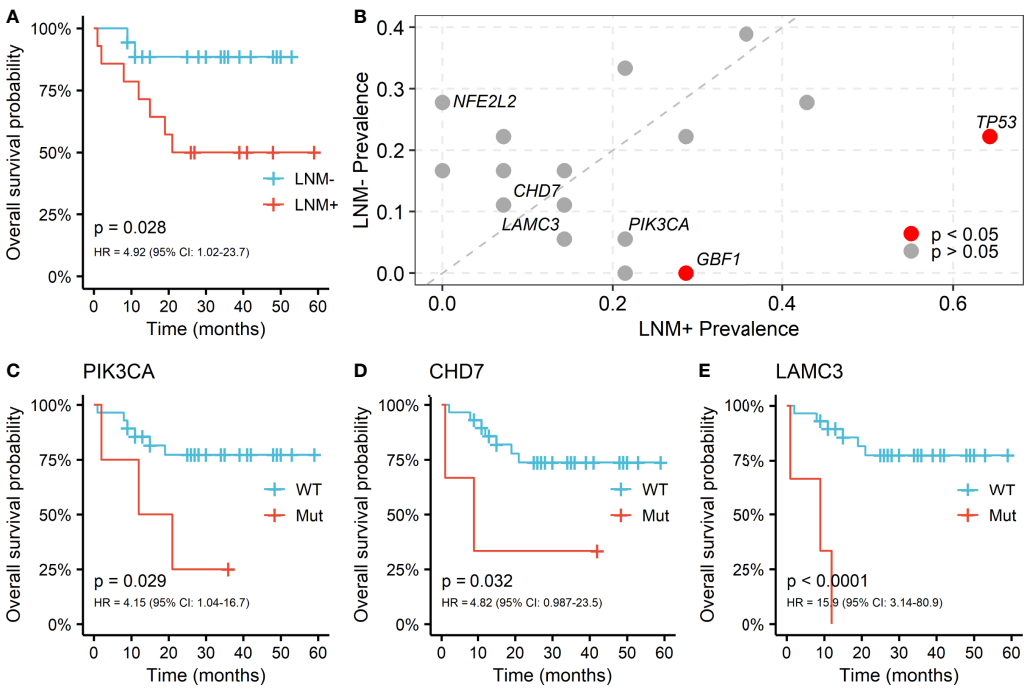


FIGURE 3 | Somatic alterations between different lymph node metastasis (LNM) subgroups and its correlation with OS in PSCC. **(A)** Kaplan-Meier curve of overall survival by lymph node metastasis (LNM). **(B)** Enrichment of somatic alterations by lymph node status. Kaplan-Meier curves of overall survival by the mutation status of gene **(C)** PIK3CA, **(D)** CHD7 and **(E)** LAMC3 in PSCC patients. P values of the log-rank test and hazard ratios are shown at the bottom left for each curve.

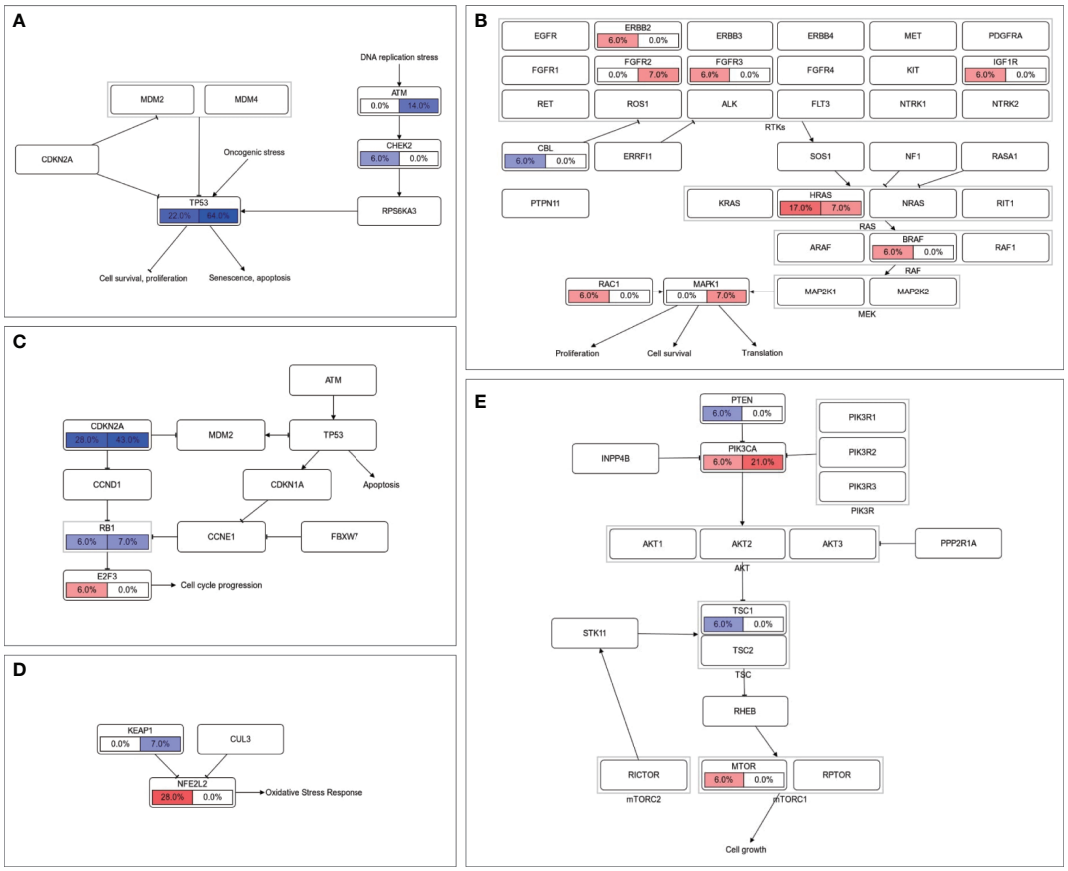


FIGURE 4 | Comparison of somatic alterations between LNM subgroups in signaling pathway level. Mutation frequencies of genes in **(A)** TP53 pathway, **(B)** RTK-RAS pathway, **(C)** cell cycle pathway, **(D)** NRF2 pathway and **(E)** PI3K pathway were labeled with LNM negative on the left and LNM positive on the right. Oncogenes were filled with red color and tumor suppressor genes were filled with blue.

pathway (*NFE2L2* and *KEAP1*, **Figure 4D**), oncogene mutations in lymph node-negative patients and tumor suppressor gene mutations in lymph node-positive patients can be observed in similar patterns. The opposite phenomena were found in the PI3K pathway (*PIK3CA* in node-positive patients and *PTEN* in node-negative patients, **Figure 4E**).

Except for certain mutants, the relationship between LNM and other tumor genomic features including tumor mutational burden, tumor heterogeneity and genomic stability were also investigated (**Supplementary Figure 4**). No significant differences were found between node-negative and positive lymph node patients in TMB (3.4 *vs.* 2.2, p = 0.44), MATH (31.9 *vs.* 21.6, p = 0.67) and wGII (0.15 *vs.* 0.12, p = 0.28). The similarity of different lymph node statuses further indicated that lymph node metastasis may be driven by key alterations rather than advanced tumor status.

DISCUSSION

Risk stratification of lymph node metastasis is essential for both clinical intervention and prognosis prediction of PSCC. Given the high-risk lymph node micrometastases in node-negative tumors (31) and high false-negative rates of modified inguinal lymph nodes dissection and dynamic sentinel lymph node biopsy (32, 33), the molecular drivers of metastasis and novel biomarkers for risk assessment of LNM need to be urgently uncovered.

The advancements within genomic characterization of PSCC were mostly constrained in a form of targeted panel strategy, except for two (9, 11). In this study, we implemented a whole-exome sequencing to perform comprehensive somatic alteration profiling of 32 PSCC patients. The observation that *TP53*, *CDKN2A*, *NOTCH1*, *TTN* and *FAT1* being the most frequently mutated genes was in concordance with previous studies and similar result was found in the pathway level. We also confirmed that *CDKN2A* plays a critical role in tumorigenesis of PSCC, which has been reported to be preferentially occurred in lichen sclerosus-external genital carcinoma (34).

Comparison between different lymph node status subgroups showed that LNM is associated with alterations of certain genes, like *TP53* and *GBF1* in our study. *GBF1* is required for the trans-Golgi network localization of HPV16 infection (35), which inactivates tumor suppressor protein p53 in penile carcinoma.

It has been found that specific mutations or changes in expression of *TP53* are correlated with LNM in various cutaneous squamous cell carcinomas (36–38), and alterations in *TP53* were significantly associated with shorter event-free survival (10). In addition, higher prevalence in positive-node patients along with the tendency towards shorter survival were observed in *PIK3CA* and *LAMC3* mutants. Notably, it has been demonstrated that lymphatic metastasis in PSCC was correlated with the elevated expression of *LAMC2* (39), another heterotrimer of the laminin gamma family. The mutations within all these genes during tumor progression of PSCC could promote its spread to lymph nodes, leading to a poor prognosis.

There were some limitations in this study. Firstly, it is unrealistic to harbor both the pre-LNM sample and samples in a state with positive-node from the same patient. This leads to the lack of direct evidence for our findings, which may need to be resolved after the establishment of animal models. Due to the low incidence of penile carcinoma, partial results did not reach the significance level with a small number of enrolled samples, like in most PSCC studies. It will be further validated by a larger cohort in the upcoming future.

In summary, we reproduced an accordant genomic landscape in penile carcinoma and depicted the formation of somatic alterations while the tumor evolved to the status liable to spread to lymph nodes. The findings also proposed candidate genetic biomarkers for both the management of low-risk primary penile tumors and prognosis prediction of patients with this malignancy.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of restrictions by national legislation/guidelines, specifically the Administrative Regulations of the People's Republic of China on Human Genetic Resources. Requests to access the datasets should be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethics committee of Hunan Cancer Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JC, C-HY, W-QH, YX, and S-SJ conceived and designed the study. JC, C-HY, W-QH, YX, and Z-ZL conducted data collection, analysis, and interpretation. JC and C-HY wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research is supported by Natural Science Foundation of Changsha Science and Technology Bureau (Basic Research Program: kq2004134) and Wu Jie-Ping Medical Foundation (Clinical Research Special Funding: 320.6750.2020-14-14).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.641869/full#supplementary-material

REFERENCES

1. Montes Cardona CE, Garcia-Perdomo HA. Incidence of Penile Cancer Worldwide: Systematic Review and Meta-Analysis. *Rev Panam Salud Publica* (2017) 41:e117. doi: 10.26633/RPSP.2017.117
2. Olesen TB, Sand FL, Rasmussen CL, Albieri V, Toft BG, Norrild B, et al. Prevalence of Human Papillomavirus DNA and p16(INK4a) in Penile Cancer and Penile Intraepithelial Neoplasia: A Systematic Review and Meta-Analysis. *Lancet Oncol* (2019) 20(1):145–58. doi: 10.1016/S1470-2045(18)30682-X
3. Haas GP, Blumenstein BA, Gagliano RG, Russell CA, Rivkin SE, Culkin DJ, et al. Cisplatin, Methotrexate and Bleomycin for the Treatment of Carcinoma of the Penis: A Southwest Oncology Group Study. *J Urol* (1999) 161(6):1823–5. doi: 10.1016/S0022-5347(05)68815-5
4. Hakenberg OW, Drager DL, Erbersdobler A, Naumann CM, Junemann KP, Protzel C. The Diagnosis and Treatment of Penile Cancer. *Dtsch Arztebl Int* (2018) 115(39):646–52. doi: 10.3238/arztebl.2018.0646
5. Pagliaro LC, Williams DL, Daliani D, Williams MB, Osai W, Kincaid M, et al. Neoadjuvant Paclitaxel, Ifosfamide, and Cisplatin Chemotherapy for Metastatic Penile Cancer: A Phase II Study. *J Clin Oncol* (2010) 28 (24):3851–7. doi: 10.1200/JCO.2010.29.5477
6. Nicholson S, Hall E, Harland SJ, Chester JD, Pickering L, Barber J, et al. Phase II Trial of Docetaxel, Cisplatin and 5FU Chemotherapy in Locally Advanced and Metastatic Penis Cancer (CRUK/09/001). *Br J Cancer* (2013) 109 (10):2554–9. doi: 10.1038/bjc.2013.620
7. Wang J, Pettaway CA, Pagliaro LC. Treatment for Metastatic Penile Cancer After First-line Chemotherapy Failure: Analysis of Response and Survival Outcomes. *Urology* (2015) 85(5):1104–10. doi: 10.1016/j.urology.2014.12.049
8. Ali SM, Pal SK, Wang K, Palma NA, Sanford E, Bailey M, et al. Comprehensive Genomic Profiling of Advanced Penile Carcinoma Suggests a High Frequency of Clinically Relevant Genomic Alterations. *Oncologist* (2016) 21(1):33–9. doi: 10.1634/theoncologist.2015-0241
9. Feber A, Worth DC, Chakravarthy A, de Winter P, Shah K, Arya M, et al. Csn1 Somatic Mutations in Penile Squamous Cell Carcinoma. *Cancer Res* (2016) 76(16):4720–7. doi: 10.1158/0008-5472.CAN-15-3134
10. McDaniel AS, Hovelson DH, Cani AK, Liu CJ, Zhai Y, Zhang Y, et al. Genomic Profiling of Penile Squamous Cell Carcinoma Reveals New Opportunities for Targeted Therapy. *Cancer Res* (2015) 75(24):5219–27. doi: 10.1158/0008-5472.CAN-15-1004
11. Wang Y, Wang K, Chen Y, Zhou J, Liang Y, Yang X, et al. Mutational Landscape of Penile Squamous Cell Carcinoma in a Chinese Population. *Int J Cancer* (2019) 145(5):1280–9. doi: 10.1002/ijc.32373
12. Zhou QH, Deng CZ, Li ZS, Chen JP, Yao K, Huang KB, et al. Molecular Characterization and Integrative Genomic Analysis of a Panel of Newly Established Penile Cancer Cell Lines. *Cell Death Dis* (2018) 9(6):684. doi: 10.1038/s41419-018-0736-1
13. Chahoud J, Pickering CR, Pettaway CA. Genetics and Penile Cancer: Recent Developments and Implications. *Curr Opin Urol* (2019) 29(4):364–70. doi: 10.1097/MOU.0000000000000640

14. Aydin AM, Chahoud J, Adashek JJ, Azizi M, Magliocco A, Ross JS, et al. Understanding Genomics and the Immune Environment of Penile Cancer to Improve Therapy. *Nat Rev Urol* (2020) 17(10):555–70. doi: 10.1038/s41585-020-0359-z
15. Reddy JP, Pettaway CA, Levy LB, Pagliaro LC, Tamboli P, Rao P, et al. Factors Associated With Regional Recurrence After Lymph Node Dissection for Penile Squamous Cell Carcinoma. *BJU Int* (2017) 119(4):591–7. doi: 10.1111/bju.13686
16. Zheng W, Li K, Zhu W, Ding Y, Wu Q, Tang Q, et al. Nomogram Prediction of Overall Survival Based on Log Odds of Positive Lymph Nodes for Patients With Penile Squamous Cell Carcinoma. *Cancer Med* (2020) 9(15):5425–35. doi: 10.1002/cam4.3232
17. Necchi A, Lo Vullo S, Mariani L, Zhu Y, Ye DW, Ornellas AA, et al. Nomogram-Based Prediction of Overall Survival After Regional Lymph Node Dissection and the Role of Perioperative Chemotherapy in Penile Squamous Cell Carcinoma: A Retrospective Multicenter Study. *Urol Oncol* (2019) 37(8):531.e7–15. doi: 10.1016/j.urolonc.2019.04.003
18. Bandini M, Spiess PE, Pederzoli F, Marandino L, Brouwer OR, Albersen M, et al. A Risk Calculator Predicting Recurrence in Lymph Node Metastatic Penile Cancer. *BJU Int* (2020) 126(5):577–85. doi: 10.1111/bju.15177
19. Rieken M, Djajadiningrat RS, Kluth LA, Favaretto RL, Xylinas E, Guimaraes GC, et al. Predictors of Cancer-Specific Mortality After Disease Recurrence in Patients With Squamous Cell Carcinoma of the Penis. *Eur Urol* (2014) 66 (5):811–4. doi: 10.1016/j.eururo.2014.05.032
20. Reyes ME, Borges H, Adjao MS, Vijayakumar N, Spiess PE, Schabath MB. Novel Prognostic Models for Patients With Penile Carcinoma. *Cancer Control* (2020) 27(1):1–12. doi: 10.1177/1073274820924728. 1073274820924728.
21. Zhang J, Cong R, Zhao K, Wang Y, Song N, Gu M. High TRIAP1 Expression in Penile Carcinoma Is Associated With High Risk of Recurrence and Poor Survival. *Ann Transl Med* (2019) 7(14):330. doi: 10.21037/atm.2019.06.47
22. Wang JY, Gao MZ, Yu DX, Xie DD, Wang Y, Bi LK, et al. Histological Subtype Is a Significant Predictor for Inguinal Lymph Node Metastasis in Patients With Penile Squamous Cell Carcinoma. *Asian J Androl* (2018) 20(3):265–9. doi: 10.4103/aja.aja_60_17
23. Peak TC, Russell GB, Dutta R, Rothberg MB, Chapple AG, Hemal AK. A National Cancer Database-Based Nomogram to Predict Lymph Node Metastasis in Penile Cancer. *BJU Int* (2019) 123(6):1005–10. doi: 10.1111/bju.14652
24. Lutzen U, Zuhayra M, Marx M, Zhao Y, Colberg C, Knupfer S, et al. Value and Efficiency of Sentinel Lymph Node Diagnostics in Patients With Penile Carcinoma With Palpable Inguinal Lymph Nodes as a New Multimodal, Minimally Invasive Approach. *Eur J Nucl Med Mol Imaging* (2016) 43 (13):2313–23. doi: 10.1007/s00259-016-3482-6
25. Khalil MI, Kamel MH, Dhillon J, Master V, Davis R, Hajiran AJ, et al. What You Need to Know: Updates in Penile Cancer Staging. *World J Urol* (2020). doi: 10.1007/s00345-020-03302-z
26. Arnedo-Pac C, Mularoni L, Muinos F, Gonzalez-Perez A, Lopez-Bigas N. OncodriveCLUSTL: A Sequence-Based Clustering Method to Identify Cancer Drivers. *Bioinformatics* (2019) 35(24):5396. doi: 10.1093/bioinformatics/btz588
27. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, et al. Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *N Engl J Med* (2014) 371(23):2189–99. doi: 10.1056/NEJMoa1406498
28. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118
29. Mroz EA, Rocco JW. MATH, a Novel Measure of Intratumor Genetic Heterogeneity, Is High in Poor-Outcome Classes of Head and Neck Squamous Cell Carcinoma. *Oral Oncol* (2013) 49(3):211–5. doi: 10.1016/j.oraloncology.2012.09.007
30. Burrell RA, McClelland SE, Endesfelder D, Groth P, Weller MC, Shaikh N, et al. Replication Stress Links Structural and Numerical Cancer Chromosomal Instability. *Nature* (2013) 494(7438):492–6. doi: 10.1038/nature11935
31. Graafland NM, Lam W, Leijte JA, Yap T, Gallee MP, Corbishley C, et al. Prognostic Factors for Occult Inguinal Lymph Node Involvement in Penile Carcinoma and Assessment of the High-Risk EAU Subgroup: A Two-Institution Analysis of 342 Clinically Node-Negative Patients. *Eur Urol* (2010) 58(5):742–7. doi: 10.1016/j.eururo.2010.08.015
32. Kirrander P, Andren O, Windahl T. Dynamic Sentinel Node Biopsy in Penile Cancer: Initial Experiences At a Swedish Referral Centre. *BJU Int* (2013) 111(3 Pt B):E48–53. doi: 10.1111/j.1464-410X.2012.11437.x
33. Peyraud F, Allenet C, Gross-Goupil M, Domblides C, Lefort F, Daste A, et al. Current Management and Future Perspectives of Penile Cancer: An Updated Review. *Cancer Treat Rev* (2020) 90:102087. doi: 10.1016/j.ctrv.2020.102087
34. Soufir N, Queille S, Liboutet M, Thibaudeau O, Bachelier F, Delestaing G, et al. Inactivation of the CDKN2A and the p53 Tumour Suppressor Genes in External Genital Carcinomas and Their Precursors. *Br J Dermatol* (2007) 156 (3):448–53. doi: 10.1111/j.1365-2133.2006.07604.x
35. Day PM, Thompson CD, Schowalter RM, Lowy DR, Schiller JT. Identification of a Role for the trans-Golgi Network in Human Papillomavirus 16 Pseudovirus Infection. *J Virol* (2013) 87(7):3862–70. doi: 10.1128/JVI.03222-12
36. Biswas NK, Das C, Das S, Maitra A, Nair S, Gupta T, et al. Lymph Node Metastasis in Oral Cancer Is Strongly Associated With Chromosomal Instability and DNA Repair Defects. *Int J Cancer* (2019) 145(9):2568–79. doi: 10.1002/ijc.32305
37. Lerias S, Esteves S, Silva F, Cunha M, Cochicho D, Martins L, et al. Cd274 (Pd-L1), CDKN2A (P16), TP53, and EGFR Immunohistochemical Profile in Primary, Recurrent and Metastatic Vulvar Cancer. *Mod Pathol* (2020) 33 (5):893–904. doi: 10.1038/s41379-019-0429-z
38. Li YY, Hanna GJ, Laga AC, Haddad RI, Lorch JH, Hammerman PS. Genomic Analysis of Metastatic Cutaneous Squamous Cell Carcinoma. *Clin Cancer Res* (2015) 21(6):1447–56. doi: 10.1158/1078-0432.CCR-14-1773
39. Zhou QH, Deng CZ, Chen JP, Huang KB, Liu TY, Yao K, et al. Elevated Serum LAMC2 Is Associated With Lymph Node Metastasis and Predicts Poor Prognosis in Penile Squamous Cell Carcinoma. *Cancer Manag Res* (2018) 10:2983–95. doi: 10.2147/CMAR.S171912

Conflict of Interest: C-HY was employed by the company GloriousMed Clinical Laboratory (Shanghai) Co., Ltd.

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.

frontiers in Oncology

ORIGINAL RESEARCH
published: 29 June 2021
doi: 10.3389/fonc.2021.707032

TCF11 Has a Potent Tumor-Repressing Effect Than Its Prototypic Nrf1α by Definition of Both Similar Yet Different Regulatory Profiles, With a Striking Disparity From Nrf2

Meng Wang[1], *Yonggang Ren*[1,2], *Shaofan Hu*[1], *Keli Liu*[1], *Lu Qiu*[1,3] *and Yiguo Zhang*[1]*

[1] The Laboratory of Cell Biochemistry and Topogenetic Regulation, College of Bioengineering, Chongqing University, Chongqing, China, [2] Department of Biochemistry, North Sichuan Medical College, Nanchong, China, [3] School of Life Sciences, Zhengzhou University, Zhengzhou, China

OPEN ACCESS

Edited by:
Feng He,
Shanghai University of Traditional Chinese Medicine, China

Reviewed by:
Shanshan Li,
University of California, San Francisco, United States
Jia Hu,
Memorial Sloan Kettering Cancer Center, United States
Yingxiang Li,
University of Michigan, United States
Yanyang Cao,
Washington University School of Medicine in St. Louis, United States

***Correspondence:**
Yiguo Zhang
yiguozhang@cqu.edu.cn;
eaglezhang64@gmail.com

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: 08 May 2021
Accepted: 09 June 2021
Published: 29 June 2021

Citation:
Wang M, Ren Y, Hu S, Liu K, Qiu L and Zhang Y (2021) TCF11 Has a Potent Tumor-Repressing Effect Than Its Prototypic Nrf1α by Definition of Both Similar Yet Different Regulatory Profiles, With a Striking Disparity From Nrf2. Front. Oncol. 11:707032. doi: 10.3389/fonc.2021.707032

Nrf1 and Nrf2, as two principal CNC-bZIP transcription factors, regulate similar but different targets involved in a variety of biological functions for maintaining cell homeostasis and organ integrity. Of note, the unique topobiological behavior of Nrf1 makes its functions more complicated than Nrf2, because it is allowed for alternatively transcribing and selectively splicing to yield multiple isoforms (e.g., TCF11, Nrf1α). In order to gain a better understanding of their similarities and differences in distinct regulatory profiles, all four distinct cell models for stably expressing *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2* have been herein established by an Flp-In™ T-REx™-293 system and then identified by transcriptomic sequencing. Further analysis revealed that Nrf1α and TCF11 have similar yet different regulatory profiles, although both contribute basically to positive regulation of their co-targets, which are disparate from those regulated by Nrf2. Such disparity in those gene regulations by Nrf1 and Nrf2 was further corroborated by scrutinizing comprehensive functional annotation of their specific and/or common target genes. Conversely, the mutant TCF11$^{\Delta N}$, resulting from a deletion of the N-terminal amino acids 2–156 from TCF11, resembles Nrf2 with the largely consistent structure and function. Interestingly, our further experimental evidence demonstrates that TCF11 acts as a potent tumor-repressor relative to Nrf1α, albeit both isoforms possess a congruous capability to prevent malignant growth of tumor and upregulate those genes critical for improving the survival of patients with hepatocellular carcinoma.

Keywords: TCF11, Nrf1α, Nrf2, transcriptomic sequencing, regulatory profiling, hepatocellular carcinoma

INTRODUCTION

In all life forms, a variety of cell identifications with specialized topological shapes are evolutionarily selectively determined by diverse hub sets of transcription factors (TFs)-regulated gene expression profiles. Thereof, activation or inhibition of distinct TFs is essential for regulation of their target gene expression by binding a specific DNA sequence to maintain and perpetuate the normal and

orderly operation of given organisms. In human, there exist over 1,600 known and likely putative TFs; they have comprised nearly forty of distinct families (1). Among them, the basic leucine zipper (bZIP) TFs consist of a larger group of the diverse basic-domain superfamily (in accordance with a new classification of TFs as described at http://www.edgar-wingender.de/huTF_classification.html). Of these bZIP factors, the Cap'n'collar (CNC) subfamily members are characterized by a highly conserved 43-aa CNC domain, located N-terminally to the basic DNA-binding domain (2, 3). All CNC-bZIP orthologues share a highly conservatism of evolution from marine bacteria, ascidian, sea urchin, octopus, fly, hydra, worm, bird, insect, fish, frog to mammals including human (4), for an indispensable role in executing the cytoprotective transcriptional responses to changing environmental stress, and thus preserving the cellular homeostasis during development and growth of distinct life forms. Only a functional CNC-bZIP heterodimer with small Maf or other bZIP factors can be allowed for binding to target genes, containing one of specialized antioxidant/electrophile response elements (AREs/EpREs) or other *cis*-regulatory homologous consensus sequences to drive distinct gene expression profiles and shape relevant biological functions.

Notably, Nrf1 [also called NFE2L1 (nuclear factor, erythroid 2 like 1), with its long TCF11 (transcription factor 11) and short Nrf1β/LCR-F1 (locus control region-1) isoforms, Gene ID: 4779] and Nrf2 [also called NFE2L2 (nuclear factor, erythroid 2 like 2), Gene ID: 4780] are two principal CNC-bZIP transcription factors expressed in various cell types and tissues of mammals, including mouse and human (2, 5). Analysis of a neighbor-joining CNC-bZIP phylogenetic tree has unveiled that the membrane-bound Nrf1 orthologues should have emerged at a more ancient stage of the earlier evolution from marine bacteria to humans, and is thereby considered as a living fossil, than the water-soluble Nrf2 (4). This discovery supports a notion that Nrf1 has a potent capability to fulfill more biological functions far beyond redox regulation that was originally identified. In fact, accumulating evidence reveals that Nrf1 exerts an important role in embryonic development (6, 7), osteoblastogenesis (8, 9), life quality control of proteostasis by proteasome (10, 11) and metabolism (12–14), anti-inflammatory immune response (15), and anti-tumor cytoprotection against hepatoma (16, 17), in addition to redox stress defense (18, 19). Conversely, loss of Nrf1's function by gene-targeting in mice leads to severe oxidative stress and spontaneous development of distinct pathological phenotypes, resembling human non-alcoholic steatohepatitis (NASH) with progressive hepatoma, neurodegenerative diseases or diabetes mellitus (3). By contrast, Nrf2 is dispensable based on the fact that animal development and growth are unaffected by its functional loss, without any pathological phenotypes. Such being the case, Nrf2 is still accepted as a master regulator of redox homeostasis (20), metabolism (21, 22) and DNA repair (23) in order to meet the healthy needs of life. Rather, Nrf2 acts as a two-edged sword to shape significant biological functions in proliferation (24, 25), resistance to apoptosis (26, 27), angiogenesis (28–30), carcinogenesis (31, 32) and invasion (33–35). Overall, these demonstrate that Nrf1 and Nrf2 elicit similar yet different physiological functions. For instance, our previous work revealed an inter-regulatory crosstalk between Nrf1 and Nrf2 at distinct levels (17), implying that they may regulate each other as a competitive player in similar biological process by distinct ways, but the details require to be further identified.

As a matter of fact, so less attention has been paid on Nrf1 than Nrf2, though the indispensable Nrf1 is highly valued as a robust deterministic transcription factor and identified as an important endoplasmic reticulum (ER) sensor for changes in the intracellular redox, glucose, protein and lipid including cholesterol, status (36, 37). Of note, a unique capability of single *Nrf1* gene confers it to be alternatively transcribed and also further subjected to selective splicing to give rise to multiple isoforms with different tempo-spatial topological properties (38). Consequently, distinct lengths of Nrf1 isoforms (e.g., TCF11, TCF11$^{\Delta N}$, Nrf1α, Nrf1β/LCR-F1) are expressed differentially in distinct type of cells, which makes diverse biological functions of Nrf1 more complicated (3, 5, 39). In response to biological cues, the ER-resident Nrf1/TCF11 is topologically dislocated from the lumen into extra-ER compartments, where their glycoproteins are deglycosylated and then subjected to selective juxtamembrane proteolytic processing to yield a mature factor before transactivating target genes (e.g., those encoding proteasomal subunits, antioxidant and cytoprotective proteins). Of note, an isoform longer than Nrf1α was originally referred to as TCF11 (40), consisting of 772 aa, but it is absent in the mouse, while the human prototypic Nrf1α of 742 aa lacks the Neh4L domain, due to alternative splicing of the *TCF11* transcript to remove its exon 4 (**Figure 1A** and **Figures S1A, B**). As such, Nrf1α retains relative complete structural domains from Neh1L to Neh6L, of which equivalents exist in Nrf2 (3, 41). Thereby, it is postulated that Nrf1α and TCF11 should be two main players to exert differential transcriptional regulation of distinct Nrf1-target genes, but this remains to be proved. In addition, the short isoform Nrf1β (42, 43), which was early designated as LCR-F1, lacks the N-terminal domain (NTD) and its adjacent acidic domain 1 (AD1), relative to Nrf1α or TCF11 (**Figure S1A**).

To date, growing evidence indicates that Nrf1 and Nrf2, as two versatile leading players in maintaining cellular homeostasis, are essential for important pathophysiological processes in human diseases. Yet, which specific target genes are regulated by both CNC-bZIP factors, and specific biological processes in which such genes are implicated, require for further in-depth study, albeit two recent reports also revealed different portions of between the indicated NRF-target expression profiles (44, 45). Herein, to refine distinct functions of Nrf1, it remains important to distinguish TCF11 and its N-terminally-truncated TCF11$^{\Delta N}$ (which is derived from deletion of amino acids at the 2nd to 156th positions in TCF11, and can also occur naturally with the reminiscent Nrf2-like structural domains) from Nrf1α and Nrf2. As for this end, we have established four different cell lines stably expressing *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2*, respectively, by using an Flp-In™ T-REx™-293 system (as deciphered in **Figure 1A** and **Figure S2**). When required, this controllable system is turned on by tetracycline to induce each factor-specific

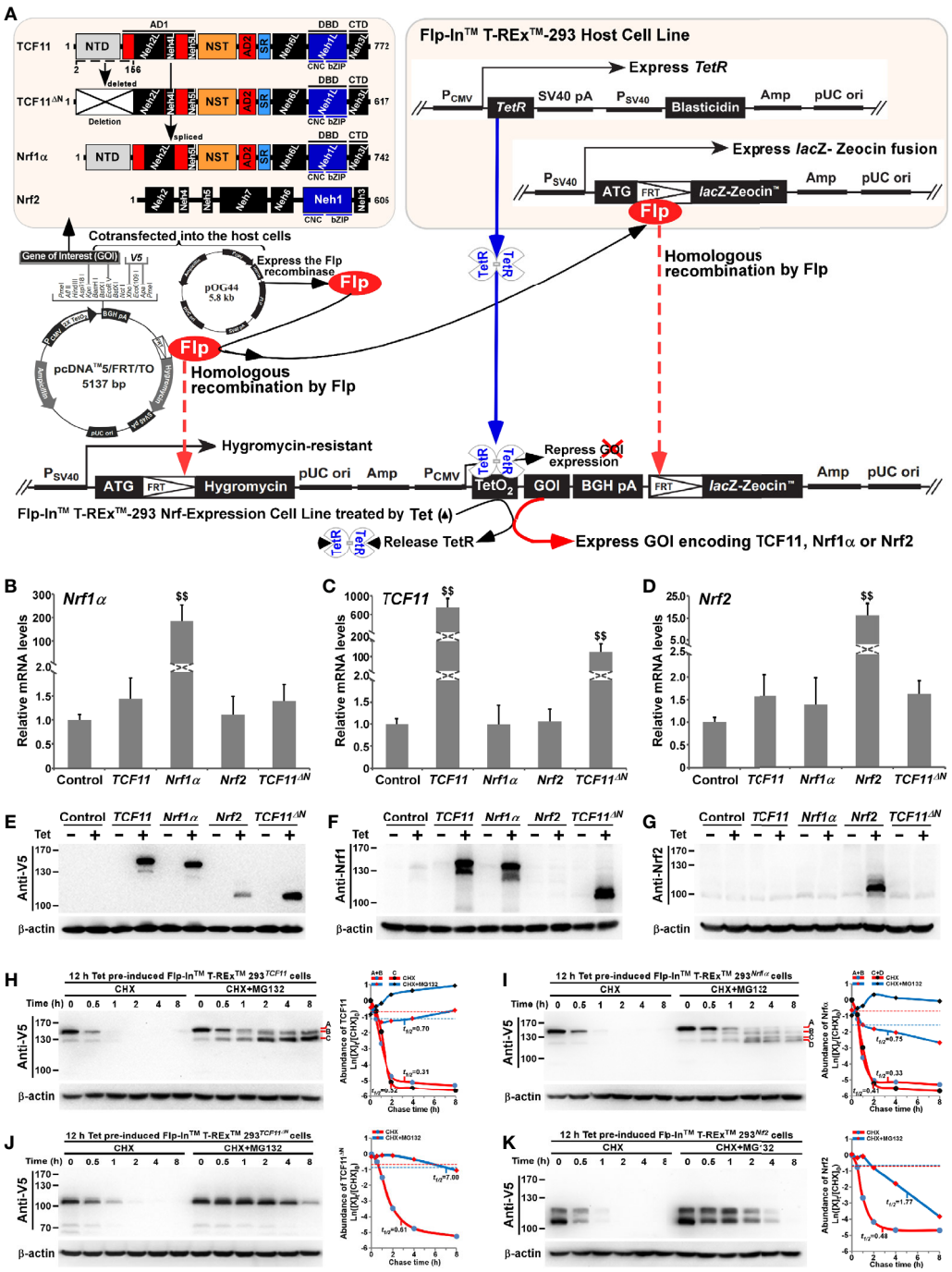


FIGURE 1 | Establishment of four distinct model cell lines to stabilize expression of *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* and *Nrf2*. **(A)** A schematic diagram of the Flp-In™ T-REx™-293 system. The system allows the Flp recombinase-mediated homologous recombination of each indicated pcDNA5/FRT/TO-V5 expression constructs (for TCF11, TCF11$^{\Delta N}$, Nrf1α or Nrf2) with the Flp-In™ T-REx™-293 host cells through the FRT sites. **(B–D)** After incubation of *TCF11*, *Nrf1α*, *Nrf2* or *TCF11*$^{\Delta N}$, as well as control cell lines with 1 μg/ml Tet for 12 h, total RNAs were isolated and then reversely transcribed into the first strand of cDNA. Subsequently, quantitative real-time PCR was employed to identify the mRNA expression levels of *Nrf1α* **(B)**, *TCF11* **(C)**, *TCF11*$^{\Delta N}$ **(C)** and *Nrf2* **(D)** in each of indicated cell lines. The data are shown as mean ± SEM (n = 3 × 3, $$, p <0.01, when compared to the *Control*). **(E–G)** Total lysates of each cell line that had been treated with 1 μg/ml Tet (+) or not (–) were subjected to protein separation by SDS-PAGE gels, and then visualized by immunoblotting with distinct primary antibodies against V5, Nrf1 or Nrf2 to identify the protein levels of TCF11, Nrf1α, Nrf2 and TCF11$^{\Delta N}$. **(H–K)** Total lysates of experimental cells, which had been induced with 1 μg/ml Tet for 12 h before being treated with CHX (50 μg/mL) alone or in combination with MG132 (10 μmol/L) for distinct times as indicated, were resolved by SDS-PAGE and then analyzed by Wester blotting with V5 antibody to identify the stability of TCF11 **(H)**, Nrf1α **(I)**, TCF11$^{\Delta N}$ **(J)** and Nrf2 **(K)** respectively.

transcriptional expression, while all others are unaffected in each of the indicated cell lines. Subsequently, transcriptomic sequencing of these cell modes unraveled that those common genes regulated by Nrf1 and Nrf2 are more responsible for regulating transcriptional expression and signal transduction in response to stimulus and diseases. The differentially expressed genes regulated by Nrf1, but not Nrf2, are preferentially enriched in carbohydrate metabolism and cellular processes, while Nrf2-specific genes strikingly prefer to developmental process. Notably, Nrf1α and TCF11 share similar regulatory patterns, but they are disparate from those of Nrf2 or TCF11$^{\Delta N}$. Besides, TCF11 can also mediate more target genes that are different from those regulated by Nrf1α, displaying distinct biological functions in development and regeneration. This notion is evidenced by our further supportive experiments, demonstrating that TCF11 can serve as a potent tumor-repressor relative to prototypic Nrf1α, albeit both factors possess a congruous capability to prevent tumor growth, as accompanied by up-regulation of those target genes significantly improving the survival rate of patients with hepatocellular carcinoma (HCC).

MATERIALS AND METHODS

Chemicals and Antibodies

All chemicals were of the highest quality commercially available. Hygromycin B and blasticidin were purchased from Invitrogen Ltd, which were employed as double drug-screening to select putative positive clones from those transfected Flp-In™ T-REx™-293 expression cells. Tetracycline from Sangon Biotech Co (Shanghai, China) was utilized as an inducible reagent at a concentration of 1 μg/ml. Both cycloheximide (CHX) and MG132 were purchased from Sigma-Aldrich (St. Louis, MO, USA). The antibody against Nrf1 proteins was acquired from our own lab [as indicated in Zhang's (46)], and all other antibodies were also employed against a V5 epitope (Invitrogen), Nrf2 (Abcam), CHP2 (Sangon Biotech), CPS1 (Abcam), FOXO1 (Cell Signaling Technology), IRS4 (Abcam), NKX2-8 (Sangon Biotech), AKR1B10 (Abcam), EPO (Proteintech Group), MUTYH (Sangon Biotech), PKM (Sangon Biotech), GP73 (Proteintech Group), GPC3 (Proteintech Group), Histone H3 (Bioss) or α-Tubulin (Beyotime), while β-actin and secondary antibodies were from ZSGB-BIO (Beijing, China).

Cell Lines, Cell Culture and Transfection

These cell lines expressing *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α*, *Nrf2*, as well as an empty control, were established by using the Flp-In™ T-REx™-293 system (Invitrogen) (**Figure S2**). Their cDNA fragments, encoding human TCF11, TCF11$^{\Delta N}$ (with a deletion to remove the 2nd to 156th residues prior to the Neh2L subdomain from TCF11), Nrf1α and Nrf2, respectively, were cloned into pcDNA5/FRT/TO-V5 expression vector, before being cotransfected with a Flp-expressing pOG44 plasmid into the Flp-In™ T-REx™-293 host cells. The Flp recombinase was allowed for its homologous recombination at FRT (Flp Recombination Target) sites existing in the host cells with each of the expression vectors. Of note, an empty expression vector was cotransfected into the host cell, to generate a negative control cell line. Then, the positive expression clones, as indicated, were selected by co-treatment of 150 μg/ml hygromycin B and 15 μg/ml blasticidin. All positively-selected cell lines were allowed for stably expression of target genes, beyond the negative control line, and thus referred to simply as *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α*, *Nrf2* or *Control*, respectively. Moreover, HepG2 was obtained originally from ATCC (Zhong Yuan Ltd., Beijing, China); MHCC97H and MHCC97L were obtained originally from the Live Cancer Institute, Fudan University of China; HL-7702, SMMC-7721 and QGY-7701 were obtained originally from National Infrastructure of Cell Line Resource (NICR), and Huh7 was obtained originally from Japanese Collection of Research Bioresources (JCRB), while both *Nrf1α*$^{-/-}$ and *Nrf2*$^{-/-}$ cell lines were created from wild-type HepG2 cells (16, 17). Notably, the fidelity of these cell lines had been conformed to be true by their authentication profiling and STR (short tandem repeat) typing maps (which were carried out by Shanghai Biowing Applied Biotechnology Co., Ltd, Shanghai, China) (37).

All experimental cells, except elsewhere indicated, were allowed for growth in DMEM basic medium (GIBCO, Life technologies), with being supplemented with 10% (v/v) foetal bovine serum (FBS, Biological Industries, Israel) and 100 units/ml of either of penicillin and streptomycin, in the 37 °C incubator with 5% CO_2. Of note, those expression constructs for human Nrf1α, TCF11 and Nrf2 were made by inserting each of their cDNA-encoding sequences into a pcDNA3 vector, respectively. The primer pairs used for this study were provided as shown in **Table S1**. The cell transfection with one of those indicated plasmids, alone or in combination, were carried out by using Lipofectamine®3000 Transfection Kit (Invitrogen) for 8 h, and then allowed for a 24-h recovery from transfection in the fresh medium before being subjected to the indicated experiments.

Quantitative Real-Time PCR

Each of experimental cell lines was subject to its total RNAs isolated by employing an RNA simple Kit (Tiangen Biotech CO. LTD, Beijing, China). Total RNAs (1 μg) were added in a reverse-transcriptase reaction to yield the first strand of cDNAs (by another RevertAid First Strand Synthesis Kit, from Thermo), which served as the template of quantitative PCR in the GoTaq® qPCR Master Mix (Promega). Then, each pairs of all forward and reverse primers (as listed in **Table S1**) were also added in an indicated PCR, that was carried out in the following conditions at 95 °Cfor 3 min, followed by 40 cycles of 15 s at 95 °C, and the last 30 s at 60 °C. The final melting curve was validated to examine the amplification quality, and β-actin at its mRNA expression levels served as an internal control for normalization.

Western Blotting

Each of experimental cell lines was harvested in a lysis buffer (0.5% SDS, 0.04 mol/L DTT, pH 7.5) containing the protease inhibitor EASYpacks (Roche, Germany). The lysates were denatured immediately at 100 °C for 10 min, sonicated sufficiently, and diluted in 3× loading buffer (187.5 nmol/L Tris–HCl, pH 6.8, 6% SDS, 30% Glycerol, 150 nmol/L DTT,

0.3% Bromophenol blue) at 100 °C for 5 min. Subsequently, equal amounts of protein extracts were subjected to separation by SDS-PAGE containing 4–15% polyacrylamide, followed by immunoblotting with each of distinct antibodies as indicated. On some occasions, the blotted membranes were also stripped for 30 min and then re-probed with an additional antibody, while β-actin served as an internal control to verify equal loading of protein in each of electrophoretic wells.

Transcriptome Sequencing Analysis

Total RNAs extracted from each of cell lines, that had incubated with 1 μg/ml of tetracycline for 12 h, were subjected to the transcriptome sequencing by Beijing Genomics Institute (BGI, Shenzhen, China) on an Illumina HiSeq 2000 sequencing system (Illumina, San Diego, CA). All detected mRNAs were fragmented into short fragments (~200 bp). The clean reads were obtained during data filtering to remove the low- quality reads, and subjected to sequence mapping to the reference of human genome (GRCh37/hg19 from UCSC database) by using SOAP2 (47). The resulting expression levels of given genes were calculated by the RPKM method (48). All those differentially expressed genes (DEGs) were identified, with the criteria fold changes ≥2 or ≤0.5 and FDR (false discovery rate) ≤0.001, by the Poisson distribution model method (PossionDis) (49, 50). Such sequencing metadata have also been submitted to NCBI SRA (PRJNA501789). In addition, the DEGs were functional annotated by using the online tool DAVID (https://david.ncifcrf.gov/) to search their involved GO (gene ontology) terms and pathways, which were further classified with QuickGO (https://www.ebi.ac.uk/QuickGO/) and KEGG (https://www.kegg.jp/) databases.

Lentivirus-Mediated Restoration of Nrf1α or TCF11

One of the lentiviral-mediated expression constructs for *Nrf1α* or *TCF11*, that was designed by a help with supplier (GeneCopoeia, Guangzhou, China), together with the GFP-expressing lentiviral control vector, were co-transfected into our $Nrf1\alpha^{-/-}$ cells, to establish *Nrf1α*-restored and *TCF11*-restored cell lines. Briefly, the lentiviral-packaging 293T cells (1×10^6) were seeded in a 10-cm dish and cultured in 10 ml DMEM supplemented with 10% FBS. Then, the mixture of 2.5 μg of each lentiviral ORF expression plasmid and 0.25 μg of the Lenti-Pac HIV plasmid in 15 μl of EndoFectin Lenti was incubated with 200 μl of Opti-MEM® (Invitrogen). The DNA-EndoFectin Lenti complex was directly added into cultured cells before being allowed for overnight incubation at 37 °C in a CO_2 incubator, followed by replaced by fresh medium supplemented with 5% FBS. Subsequently, a 1:500 volume of the TiterBoost reagent was further added to the above cultured media and allowed for continuous culture. The pseudovirus-containing culture media for 48 h post-transfection were collected by centrifuging as 500×*g* for 10 min. Then, the resulting lentivirus titer was estimated, prior to being subjected to efficient transfection of $Nrf1\alpha^{-/-}$ hepatoma cells.

Subcutaneous Tumor Xenografts in Nude Mice

Mouse xenograft models were made by subcutaneous heterotransplantation of wild-type HepG2 (*WT*), $Nrf1\alpha^{-/-}$, *Nrf1α*-restored and *TCF11*-restored cells, respectively, into nude mice as described (51). Each line of experimental cells (1×10^7) was allowed for its exponential growth and then suspended in 0.2 ml of serum-free DMEM, before being inoculated subcutaneously into the right upper back region of male nude mice ($BALB/C^{nu/nu}$, 6 weeks, 18 g, from HFK Bioscience, Beijing) at a single site. The procedure of injection into all mice was complete within 30 min, and subsequent formation of the subcutaneous tumor xenografts was observed. The tumor sizes, after emerged, were measured every two days, until the 42nd day when all those mice were sacrificed and the transplanted tumors were excised. The sizes of growing tumors were calculated by a standard formula (i.e., $V = ab^2/2$) and shown graphically (n = 7 per group). Notably, all the mice were maintained under standard animal housing conditions with a 12-h dark cycle and allowed access *ad libitum* to sterilized water and diet. All relevant studies were carried out on 8-week-old mice (with the license No. PIL60/13167) in accordance with United Kingdom Animal (Scientific Procedures) Act (1986) and the guidelines of the Animal Care and Use Committees of Chongqing University and the Third Military Medical University, both of which were also subjected to the local ethical review (in China). All relevant experimental protocols were approved by the University Laboratory Animal Welfare and Ethics Committee (with two institutional licenses SCXK-PLA-20120011 and SYXK-PLA-20120031). As for additional ethical concerns about the xenograft model mice bearing so big tumors insomuch as to give rise to certain bleeding ulcers, such a bad health condition of mice was only emerged from only day 2 prior to being sacrificed, and also such relevant study was indeed conducted according to the valid ethical regulations that have been approved.

The Colony Formation Assay on Soft Agar

The cell culture plates (each with a diameter of 10 cm) were coated by the basement gel containing 0.6% soft agar mixed in the complete medium, upon which the upper gel containing 0.35% soft agar. Then, experimental cells (2×10^4, that had been growing in the exponential phase) was allowed for two-layer gel formation. Thereafter, the plates were cultured for 2–3 weeks in the incubator at 37 °C with 5% CO_2 before being stained with 1% crystal violet reagent (Sigma) and counted.

The *In Vitro* Scratch Assays

When experimental cells (1×10^5) grown in 6-well plates reached 70% confluency, they were allowed for synchronization by 12-h starvation in serum-free medium and then treated with 1 μg/ml of mitomycin C (from Cayman, USA) for 6 h. Subsequently, a clear 'scratch' in the cell monolayer was created and then allowed for being healed in the continuous culture at 37 °C with 5% CO_2. Thereafter, the cell migration was quantified according to the standard procedures (52).

The Transwell-Based Migration and Invasion Assays

The Transwell-based migration and invasion assays were conducted in the modified Boyden chambers (Transwell; Corning Inc. Lowell, MA, USA) as described previously (53). When the growing cells reached 70% confluency, they were starved for 12 h in serum-free medium. The experimental cells (5×10^3) were suspended in 0.5 ml medium containing 5% FBS and seeded in the upper chamber of a Transwell, which allows the cells to grow on the microporous polycarbonate membrane that is tissue culture-treated to enhance the cellular attachment to the bottom. The cell-seeded Transwells were placed in each well of 24-well plates containing 1 ml of complete medium (i.e., the lower chamber), and then cultured for 24 h in the incubator at 37 °C with 5% CO_2. Of note, the bottom of upper Transwell was pre-coated by matrigel basement matrix (BD, Biosciences, USA), before the cells were placed in the invasion assay. The remaining cells in the upper chamber were removed, and the cells attached to the lower surface of the Transwell membranes were fixed with 4% paraformaldehyde (AR10669, BOSTER) and stained with 1% crystal violet reagent (Sigma) before being counted.

Subcellular Fractionation

Equal numbers (4×10^6) of different cell lines were seeded in each of 10-cm dishes and allowed for growth for 24 h, followed by treatment with Tet (1 μg/ml) for additional 12 h alone or in combination with MG132 (10 μmol/L, within the last 4 h added) before being harvested in an ice-cold Nuclei EZ lysis buffer (Sigma, NUC101-1KT). Then the lysates were subjected to subcellular fractionation by centrifuging at 500×*g* for 5 min at 4 °C. The supernatants were collected as the non-nuclear cytoplasmic fractions, while the sediments were washed twice with the Nuclei EZ lysis buffer, and the resulting nuclear fractions were pelleted by centrifuging at 500×*g* for 5 min at 4 °C. Subsequently, the cytoplasmic and nuclear fractions were evaluated by Western blotting with distinct antibodies.

Immunofluorescence Assay

Experimental cells (2×10^5) were allowed for 24-h growth on a cover glass placed in each of 6-well plates, and treated with Tet (1 μg/ml) for additional 12 h before being fixed with 4% paraformaldehyde for 20 min. The cells were then permeabilized for 10 min with 0.1% Triton X-100 (Beyotime, diluted with PBS) before immunocytochemistry with the primary antibodies against the V5 tag (diluted at 1:100) incubated at 4 °C overnight. The immunostained cells were visualized by incubation with the fluorescein-conjugated goat anti-mouse IgG (ZSGB-BIO, dilution 1:100) for 1 h at room temperature in the dark, followed by DAPI staining (KeyGEN BioTECH, KGA215) of the nuclear DNAs for 5 min. The resulting images were observed and photographed by fluorescence microscope.

Flow Cytometry Analysis of Cell Cycle and Apoptosis

Experimental cells (5×10^5) were allowed for growth in 6-cm dish for 48 h and synchronization by 12-h starvation in a serum-free medium, before being treated with 10 μmol/L BrdU for 12 h. The cells were fixed for 15 min with 100 μl BD Cytofix buffer (containing a mixture of the fixative paraformaldehyde and the detergent saponin) at room temperature and permeabilized for 10 min with 100 μl BD Cytoperm permeabilization buffer (containing fetal bovine serum as a staining enhancer) on ice. Subsequently, the cells were re-fixed and treated with 100 μl DNase (at a dose of 300 μg/ml in DPBS) for 1 h at 37 °C, in order to expose the incorporated BrdU, followed by staining with FITC conjugated anti-BrdU antibody for 1 h at room temperature. Thereafter, the cells were suspended in 20 μl of 7-amino-actinomycin D solution 20 min for the DNA staining, and re-suspended in 0.5 ml of a staining buffer (i.e., 1× DPBS containing 0.09% sodium azide and 3% heat-inactivated FBS), prior to the cell cycle analysis by flow cytometry. Additional fractions of cells (5×10^5) were allowed for 48-h growth in 6-cm dish before being harvested for apoptosis analysis. The cells were pelleted by centrifuging at 500×*g* for 5 min and washed by PBS three times, before being incubated for 15 min with 5 μl of Annexin V-FITC and 10 μl of propidium iodide (PI) in 195 μl of binding buffer, followed by apoptosis analysis with flow cytometry. The results were further analyzed by the FlowJo 7.6.1 sofware.

Hematoxylin–Eosin Staining Assay

Representatives of the above xenograft tumor tissues were fixed with 4% paraformaldehyde and then transferred to 70% ethanol according to the routine protocol. Thereafter, all individual tumor tissues were placed in the processing cassettes, dehydrated through a serial of alcohol gradient, and then embedded in paraffin wax blocks before being sectioned into a series of 5-μm-thick slides. Next, the tissue sections were dewaxed in xylene, and then washed twice in 100% ethanol to eliminate xylene, followed by rehydration in a series of gradient concentrations of ethanol with being distilled. Subsequently, they were stained with the standard hematoxylin and eosin (H&E) and visualized by microscopy.

Statistical Analysis

Statistical significances were determined using either Student's *t*-test (for a comparison between two groups) or two-way ANOVA (for comparison among multiple groups). The relevant data presented herein are shown as a fold changes (mean ± SEM or ± SD) with significant differences that were calculated by the value of $p < 0.05$).

RESULTS

Controllable Model Cell Lines for Stably Expressing TCF11, TCF11$^{\Delta N}$, Nrf1α, or Nrf2 Are Established

As shown in **Figure 1A**, four distinct expression constructs for *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2*, together with a recombinase Flp-expressing plasmid, were co-transfected into the Flp-In™ T-REx™-293 host cells and also integrated into this host cells by Flp-mediated homologous recombination at FRT sites. Then,

putative positive cell clones stably expressing the indicated gene of interest (GOI) were selected and maintained by hygromycin B (150 μg/ml) and blasticidin (15 μg/ml) to establish these cell models for expression of *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2*. Such model cell lines are controllable because their GOI are tightly monitored by interaction of Tet (tetracycline) with its repressor TetR; thus, only after TetR will be released from the $TetO_2$ operator, transcriptional expression of GOI and down-stream target genes can be induced. In subsequent parallel experiments, an additional cell line co-transfected with the empty expression vector served as an internal negative control. In order to further validate such stable controllable expression of *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2*, respectively, all these indicated model cell lines were treated with 1 μg/ml Tet for 12 h and then determined by real-time quantitative PCR (**Figures 1B–D**) and immunoblotting (**Figures 1E–G**).

Notably, a pair of specific TCF11-recognized primers, including part of the Neh4L-coding nucleotides, were designed by distinguishing it from *Nrf1α*, because the Neh4L-missing nucleotides of *Nrf1α* remains present in *TCF11* and also in its N-terminally-truncated *TCF11*$^{\Delta N}$. Thus, the latter two factors-shared same primers were employed for quantitative PCR examinations of *TCF11* and *TCF11*$^{\Delta N}$. As anticipated, real-time qPCR revealed that each specific transcriptional expression of *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2*, respectively, was induced by Tet in their indicated model cell lines (**Figures 1B–D**). Such Tet-induced protein expression abundances of TCF11, TCF11$^{\Delta N}$, Nrf1α and Nrf2 were evaluated by immunoblotting with antibodies against Nrf1/TCF11, Nrf2 and their C-terminally-tagged V5 epitope, respectively. The resulting data (**Figures 1E–G**) demonstrated that those model cell lines had a strong capability to stably express each of interested genes, one of which had almost no effects on all the others examined, though each factor-specific expression was induced under Tet control. Of note, a major protein of TCF11 exhibited a slightly slower mobility, than that of Nrf1α, on electrophoretic gels, whereas the V5-tagged TCF11$^{\Delta N}$ mobility appeared to coincide closely to the electrophoretic band of Nrf2 (**Figure 1E**). Besides, a few of putative C-terminally-truncated isoforms of Nrf1α, TCF11 or TCF11$^{\Delta N}$ were also immunoblotted with Nrf1/TCF11-specfic antibody (**Figure 1F**). Furtherly, the subcellular nucleocytoplasmic fractionation and immunofluorescence experiments unraveled that a certain amount of TCF11, TCF11$^{\Delta N}$, Nrf1α or Nrf2 was allowed to be localized in the cellular nucleus (**Figures S3A–E**). However, such nuclearly-positioning proteins are rapidly degraded, due to this fact that these protein degradations were inhibited by MG132 (at 10 μmol/L), so that obvious increases in their protein expression levels were recovered in the nuclear fractions of MG132-treated cells (**Figures S3A–D**).

For further time-course analysis of *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* and *Nrf2*, their indicated cell lines that had been pre-treated for 12 h with 1 μg/ml Tet were treated with 50 μg/ml cycloheximide (CHX, that inhibits biosynthesis of nascent proteins) alone or plus a proteasomal inhibitor MG132. As shown in **Figures 1H–K**, the N-terminal truncation of TCF11 (to remove both its ER-targeting signal peptide sequence and adjacent juxtamembrane proteolytic degron) caused the resulting TCF11$^{\Delta N}$ isoform to become stabilized relatively. Subsequent calculation of their half-lives in CHX-treated cells suggested that a major TCF11$^{\Delta N}$ protein was conferred with a relative higher stability than those of TCF11, Nrf1 and Nrf2, each with distinct expression isoforms (as illustrated graphically in **Figures 1H–K**). Upon co-treatment of cells with CHX and MG132, all those examined protein half-lives were markedly enhanced. Such being the case, intact TCF11 protein-A became gradually fainter and then disappeared by 2 h after co-treatment (**Figure 1H**). Such disappearance of TCF11 protein-A seemed to be accompanied by gradual emergence and increment of its protein-B and -C until the end of 8-h experimentation. Similar yet different conversion of Nrf1α protein-A into short isoforms-B, -C and -D was observed (**Figure 1I**). However, no similar changes were determined in two cases of Nrf2 and TCF11$^{\Delta N}$, because both proteins gradually decreased with increasing time of co-treatment until they finally disappeared from 4 to 8 h after co-treatment of the cells with CHX and MG132 (**Figures 1J, K**). Collectively, these demonstrate that the absence of the N-terminally ER-targeting sequence in TCF11$^{\Delta N}$, as well as in Nrf2, allows them to display distinguishable behaviors from the ER-resident TCF11 and Nrf1α, both of which are manifested in similar but nuanced ways of topobiological processing within and around this organelle before being dislocated into the nucleus.

Differential Expression Profiles of Genes Regulated by TCF11, TCF11$^{\Delta N}$, Nrf1α and Nrf2 Are Defined

To identify differential expression profiles of genes regulated by TCF11, TCF11$^{\Delta N}$, Nrf1α and/or Nrf2, relevant RNAs extracted from the established model cell lines were subjected to transcriptome sequencing. As a result, all those detectable genes, if upregulated or downregulated respectively with fold changes ≥2 or ≤0.5 plus false discovery rate (FDR) ≤0.001 (**Figure 2A**), were defined as differentially expressed genes (DEGs), by comparison with equivalents measured from control cells. Thereof, 2,845 DEGs were detected in *TCF11*-expressing cells, of which 2,786 target genes were upregulated by TCF11 (**Figure 2A**, and **Table S2**). By contrast, *Nrf1α*-expressing cells only yielded 1,001 DEGs, i.e., 957 upregulated plus 44 downregulated (**Table S3**), whereas *Nrf2*-expressing cells led to a significantly decreased number of upregulated genes (i.e., 276) but as accompanied by down-regulation of 457 DEGs (**Table S4**). Notably, 1,459 DEGs were identified in *TCF11*$^{\Delta N}$-expressing cells, with so many as 989 genes downregulated (**Table S5**). Interestingly, TCF11$^{\Delta N}$ appeared to be endowed with a regulatory trend of its target genes similar to that of Nrf2 (**Figure 2**). Such changed DEGs with distinct trends among different groups were further explicated by scatterplots of gene expression profiles (**Figure S4A**). Together, these data indicate that TCF11 makes a greater impact on the overall gene expression than Nrf1α, albeit both contribute to basically positive regulation of their DEGs, whereas Nrf2 and TCF11$^{\Delta N}$ make more contributions to negative regulation rather than positive regulation of their DEGs.

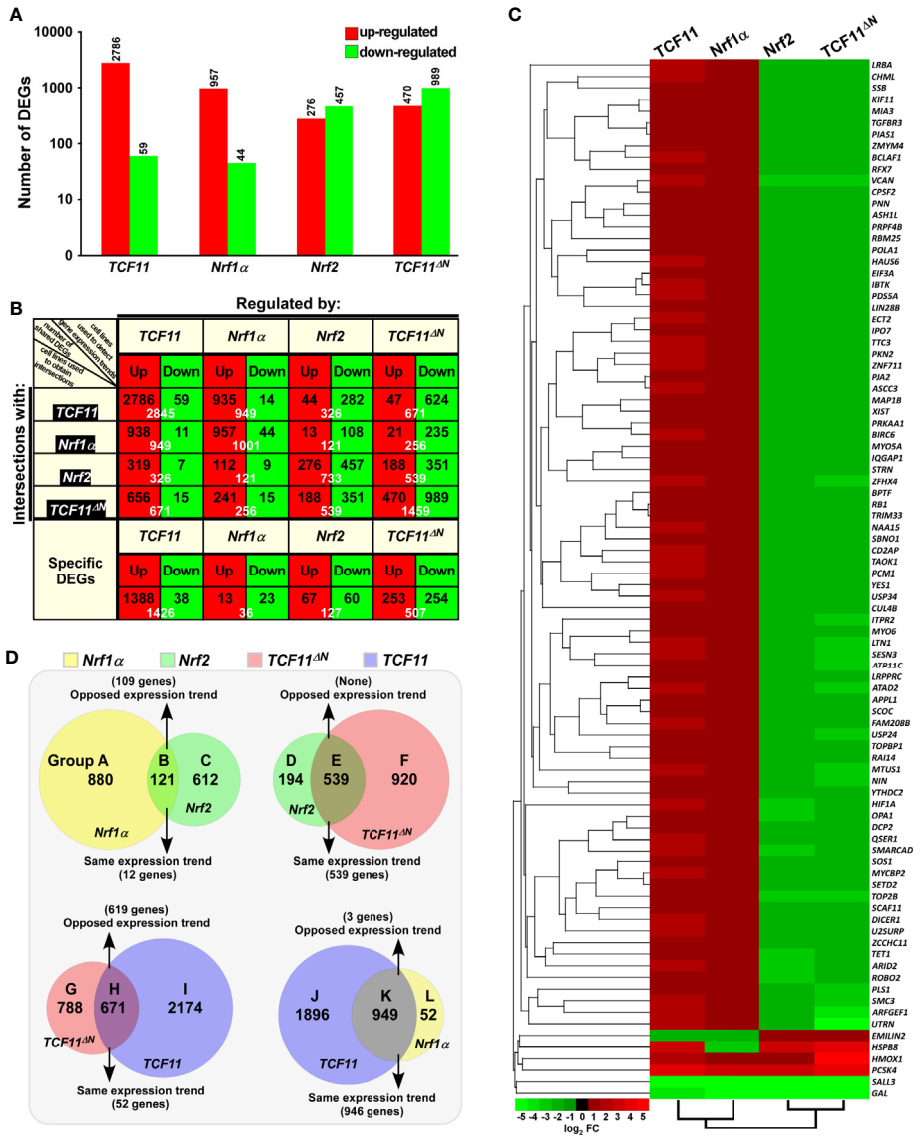


FIGURE 2 | Statistical analysis of the data obtained from transcriptome sequencing. **(A)** Differentially expressed genes (DEGs) in distinct cell lines were analyzed by transcriptome sequencing, and relevant differences in the number of those increased or decreased DEGs are shown in the histogram. The DEGs were selected according to the following criteria: fold change ≥2 or ≤0.5 and FDR ≤0.001 (as compared to the Control cells). **(B)** The specific DEGs in each cell line and their common DEGs between every two cell lines were also counted as indicated in the chart, and the number of increased and decreased DEGs in each group is shown separately in black font, and the total is shown in white. In addition, the change trends of DEGs in each group were indicated in red or green, which represent up-regulated or down-regulated in the cells in the first row, respectively. **(C)** The heatmap with hierarchical clustering of 90 DEGs shared in all four cells lines. **(D)** Distinct groupings of the subsequent functional annotation and also the Venn diagram of DEGs between every two cell lines.

The intersections of these four groups of DEGs regulated by TCF11, TCF11$^{\Delta N}$, Nrf1α and/or Nrf2 were shown by the Venn diagram (**Figure S4B**). Those common DEGs between every two cell groups and each specific DEGs were taken into account (in *orthogonal table*, **Figure 2B**), with distinct regulatory tendencies of DEGs even in each subgroup. Of note, the common DEGs between *Nrf2*-expressing cells and the others were more likely to be downregulated rather than up-regulated by Nrf2, with a roughly similar number of its upregulated genes to down-regulated genes amongst *Nrf2*-specific DEGs. Another similar situation also occurred in *TCF11*$^{\Delta N}$-expressing cells.

Approximately 90 DEGs were identified to be shared among these four cell lines (as shown in the Venn diagram, **Figure S4B**). Differential expression levels of these DEGs were presented by their heatmap with hierarchical clustering (**Figure 2C**). In the shared DEGs, only four genes *HMOX1* (heme oxygenase 1), *PCSK4* (proprotein convertase subtilisin/kexin type 4), *SALL3* (spalt like transcription factor 3) and *GAL* (galanin and

GMAP prepropeptide) had a similar expression trends in all four cell lines. Most of other shared DEGs were up-regulated by TCF11 and Nrf1α, but down-regulated by Nrf2 and TCF11$^{\Delta N}$, besides two exceptions of *EMILIN2* (elastin microfibril interface 2) and *HSPB8* (heat shock protein family B member 8) regulated by opposite ways (**Figure 2C**). In addition, no evident effects of *Nrf2*-expressing cell model on *TCF11*, *TCF11*$^{\Delta N}$ or *Nrf1α* was examined by transcriptome sequencing, but conversely, only a marginal increase of *Nrf2* expression was found in *Nrf1α*-, *TCF11*-, rather than *TCF11*$^{\Delta N}$-expressing cell models (**Figure S4C**). Amongst other CNC-bZIP members, only *BACH1* was up-regulated by TCF11, whilst all *sMAF* partners were up-regulated by Nrf2 and TCF11$^{\Delta N}$, but *Keap1* was unaffected.

In order to gain a further insight into similarities and differences in biological functions of between TCF11, TCF11$^{\Delta N}$, Nrf1α and Nrf2, their common and different regulatory DEGs were scrutinized in distinct combinations of every two groups as indicated by Groups A to L (**Figure 2D**). Similar or opposite trends in those common DEGs of every two cell groups were schematically shown. For an example of Group B, 121 common DEGs between *Nrf1α* and *Nrf2* were subdivided into 109 oppositely-regulated genes and the other 12 genes with the same directional tend. In Group E, all those common DEGs were manifested only with the same directional trend to be regulated by both *Nrf2* and *TCF11*$^{\Delta N}$. By contrast, a largely opposing expression trend in 619 of the common 671 DEGs co-regulated by *TCF11*$^{\Delta N}$ and *TCF11* was shown in Group H, while 946 of the other common 949 DEGs shared by *TCF11* and *Nrf1α* in Groups K showed the same tendency of expression change. Such groups of these DEGs were also subjected to comprehensive analysis of their functional annotations as described below.

Nrf1α and Nrf2 Have Diverse Regulatory Patterns of Their Target Genes

In Group A, 880 DEGs were identified in *Nrf1α*-expressing rather than *Nrf2*-expressing cells, and further subjected to their functional annotation by the DAVID (*d*atabase for *a*nnotation, *v*isualization and *i*ntegrated *d*iscovery) (**Figure 3A**), for the data mining in order to delineate unique biological functionality of Nrf1α in regulating genes preferentially than Nrf2. Besides their shared common 121 DEGs in Group B, the other 612 DEGs in Group C were also annotated to identify those biological functions of Nrf2 that were responsible preferentially than Nrf1α (**Figure 3A**, and **Table S6**). The details of all top significant biological process terms and pathways enriched by DEGs in three different groups were deciphered in histograms and scatterplots (**Figure 3A**). Furtherly, the biological process terms and pathways were classified by using QuickGO and KEGG databases. Such functional annotation analysis implied that DEGs regulated by Nrf1α (Group A) were predominantly involved in cellular metabolic process, response to stimulus, replication and repair, cell growth and death, protein folding, sorting and degradation, signal transduction, immune system and cancers. In Group B, DEGs co-regulated by Nrf1α and Nrf2 took part in distinct cellular metabolic process, RNA processing, regulation of biological process, response to stimulus, apoptotic process, regulation of transcription, translation and cancers. Nrf2-regulated DEGs (in Group C) were also responsible for cellular metabolic process, apoptotic process, response to stimulus, regulation of transcription, development and regeneration, signal transduction, endocrine system, infectious diseases and cancers.

Based on certain association with multiple functions, along with higher expression levels and well significance, 19 DEGs were selectively verified by further quantitative PCR analysis (**Figure 3B**). The results demonstrated that four DEGs, including *BARD1* (*BRCA1 associated RING domain 1*), *CASP3* (*caspase 3*), *HSP90AA1* (*heat shock protein 90 alpha family class A member 1*) and *BPTF* (*bromodomain PHD finger transcription factor*) were upregulated by Nrf1α, but not significantly affected by Nrf2. By contrast, nine DEGs, including *BLM* (*Bloom syndrome*, *RecQ like helicase*), *CCNG1* (*cyclin G1*), *PALB2* (*partner and localizer of BRCA2*), *STAG2* (*stromal antigen 2*), *DICER1* (*dicer 1*, *ribonuclease III*), *PNN* (*pinin*, d*esmosome associated protein*), *RB1* (*RB transcriptional corepressor 1*), *SMARCA1* (*SWI/SNF related*, *matrix associated*, *actin dependent regulator of chromatin*, *subfamily a*, *member 1*) and *SOS1* (*SOS Ras/Rac guanine nucleotide exchange factor 1*) were upregulated by Nrf1α, but downregulated by Nrf2. Another six DEGs, such as *FOXO1* (f*orkhead box O1*), *KRAS* (*KRAS proto-oncogene*, *GTPase*), *MAP2K6* (*mitogen-activated protein kinase kinase 6*), *NF1* (*neurofibromin 1*), *NRP1* (*neuropilin 1*) and *TGFBR1* (*transforming growth factor beta receptor 1*) were reduced in *Nrf2*-expressing cells, but no significant changes of them was detected in *Nrf1α*-expressing cells. In addition, putative functions of the examined genes as well as of their encoding proteins were also extracted (**Figure 3B**, *on the bottom*).

TCF11$^{\Delta N}$ Exhibits a Similar Regulatory Profile to That of Nrf2

All those DEGs regulated by *Nrf2* or *TCF11*$^{\Delta N}$ alone or both were divided into three Groups D, E and F, respectively, and then functionally annotated with the above-described methods (**Figure 4A** and **Table S7**). In Group D, 194 DEGs regulated by Nrf2, but not by TCF11$^{\Delta N}$, were generally involved in cellular metabolic process, localization, regulation of biological process, developmental process, response to stimulus, signal transduction and cancers. In the intersected Group E, 539 DEGs co-regulated by *Nrf2* and *TCF11*$^{\Delta N}$ were also associated with cellular metabolic process, development and regeneration, regulation of transcription, response to stimulus, apoptotic process, endocrine system, signal transduction, infectious diseases and cancers. In Group F, 920 DEGs regulated by *TCF11*$^{\Delta N}$ were significantly enriched in cellular metabolic process, cell motility, cellular community, developmental process, regulation of transcription, signal transduction, signaling molecules and interaction, cardiovascular diseases and cancers. Intriguingly, eight common biological process terms and additional eight common pathways were predicted to exist between the top 10 functions significantly enriched in Group C (including DEGs regulated by *Nrf2* but not by Nrf1α) and Group E (with an intersection of DEGs shared by *Nrf2* and *TCF11*$^{\Delta N}$) (**Figure S4D**).

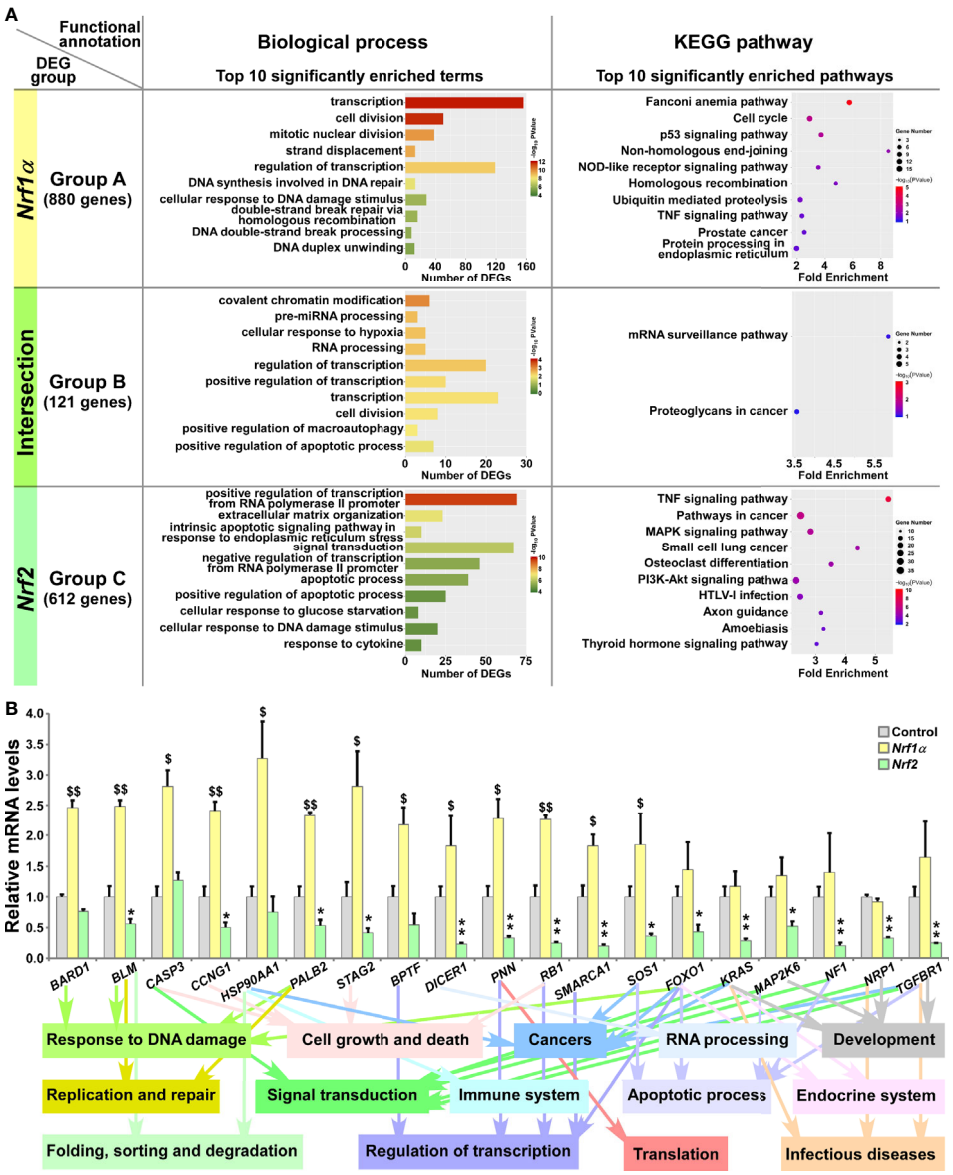


FIGURE 3 | Functional annotation of specific or common DEGs in *Nrf1α* and *Nrf2* cells. **(A)** The top 10 of significant biological process terms and pathways enriched by DEGs in Groups A, B, and C were exhibited in histograms and scatterplots, respectively. **(B)** After *Nrf1α*, *Nrf2* and Control cell lines were incubated with 1 μg/ml Tet for 12 h, total RNAs were isolated and reversely transcribed into the first strand of cDNA. Subsequently, the mRNA levels of DEGs that were associated with more functions, along with high expression levels and well significance in Groups A to C, were determined by quantitative real-time PCR analysis of *Nrf1α*, *Nrf2* and Control cell lines. The data are shown as mean ± SEM (n = 3 × 3, $^{*}p$ 0.05; $^{**}p < 0.01$; $^{\$}p < 0.05$; $^{\$\$}p < 0.01$, when compared to the *Control values*).

This implies that Nrf2 and TCF11$^{\Delta N}$ share the common regulatory profiles, but they are likely differential from those of Nrf1α and Nrf2.

The results of quantitative PCR validation (**Figure 4B**) revealed that expression of *FGF21* (*fibroblast growth factor 21*), *FOSB* (a *subunit of AP-1 transcription factor*) and *JUNB* (*another subunit* of *AP-1*) were upregulated by Nrf2 and TCF11$^{\Delta N}$, with downregulation of *MAP2K6* (*mitogen-activated protein kinase kinase 6*) and *NRP1* (*neuropilin 1*). By contrast, *GLI1* (*GLI family zinc finger 1*) was upregulated by Nrf2, but not significantly altered by TCF11$^{\Delta N}$. Conversely, *TGFBR1* (*transforming growth factor beta receptor 1*) was downregulated by Nrf2, but upregulated by TCF11$^{\Delta N}$. However, reduced expression of *CAV1* (*caveolin 1*) was accompanied by increased *COL6A2* (*collagen type VI alpha 2 chain*) in TCF11$^{\Delta N}$-expressing cells, but almost unaffected by Nrf2. In addition, putative functions of such target genes were mapped, as indicated by the histogram (**Figure 4B**, *on the bottom*).

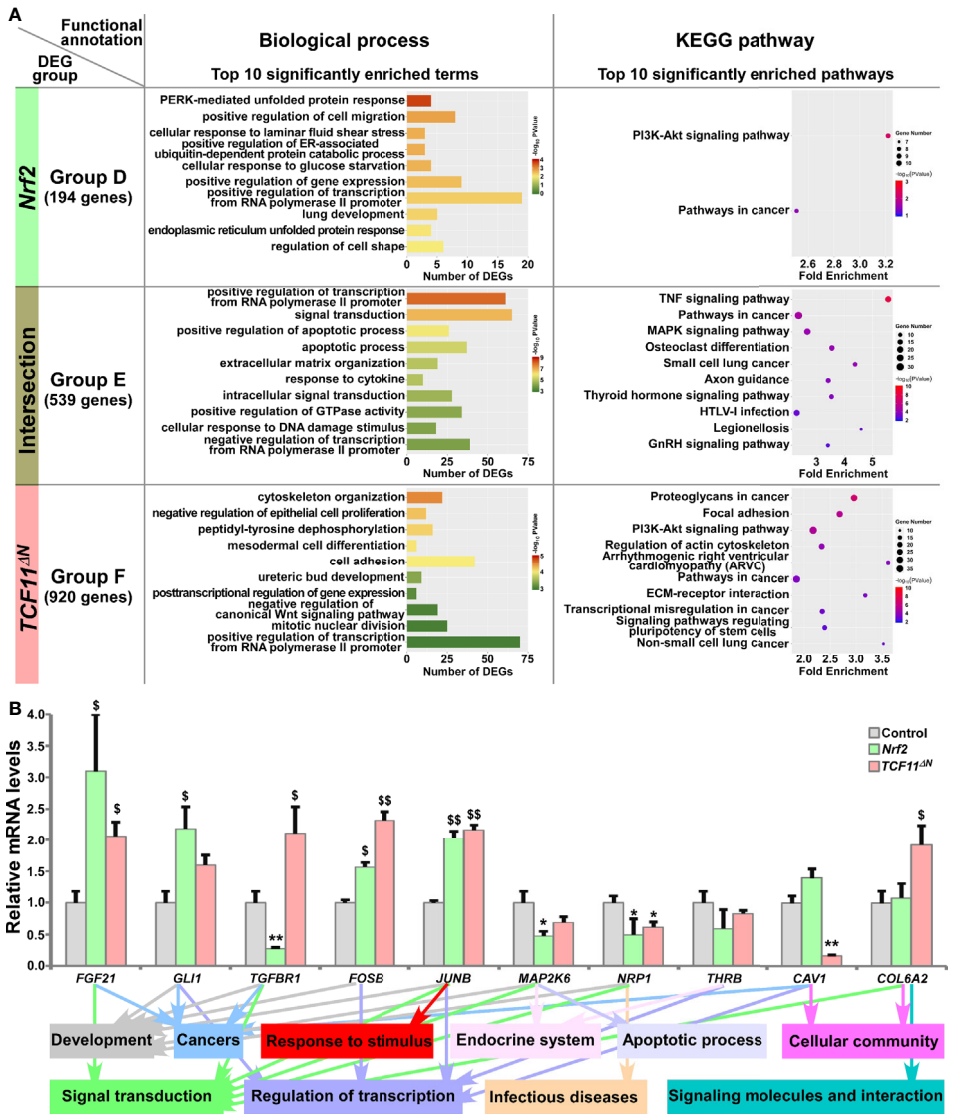


FIGURE 4 | Functional annotation of specific or common DEGs in *Nrf2* and *TCF11*$^{\Delta N}$ cells. **(A)** Top 10 of significant biological process terms and pathways enriched by DEGs in Groups D, E, and F were exhibited in histograms and scatterplots, respectively. **(B)** After induced with 1 μg/ml Tet for 12 h, total RNAs were isolated from *Nrf2*, *TCF11*$^{\Delta N}$ or Control cell lines and then reversely transcribed into the first strand of cDNA. Subsequently, the mRNA levels of DEGs that were associated with more functions as annotated, along with high expression levels and well significance in Groups D to F, were determined by quantitative real-time PCR analysis of *Nrf2*, *TCF11*$^{\Delta N}$ and Control cells. The data are shown as mean ± SEM (n = 3 × 3, $^{*}p < 0.05$; $^{**}p < 0.01$; $^{\$}p < 0.05$; $^{\$\$}p < 0.01$, when compared to the *Control values*).

TCF11 and Its Truncated TCF11$^{\Delta N}$ Regulate Similar yet Different Subsets of Target Genes

The common and distinct target DEGs in *TCF11*$^{\Delta N}$- and/or *TCF11*-expressing cell lines were assigned into three groups, and functionally annotated with the aforementioned method as visualized in bar charts and scatterplots (**Figure 5A** and **Table S8**). In Group G, 788 DEGs were identified as targets of TCF11$^{\Delta N}$, but not of TCF11, and enriched with distinct functions in cellular metabolic process, regulation of biological process, cellular community, apoptotic process, developmental process, development and regeneration, response to stimulus, signal transduction, signaling molecules and interaction, infectious diseases and cancers. Their commonly-shared 671 DEGs in Group H were functionally responsible for cellular metabolic process, cell cycle, cell motility, signal transduction; protein folding, sorting and degradation, transport and catabolism, endocrine system and cancers. In Group I, 2174 DEGs were regulated by *TCF11*, rather than by *TCF11*$^{\Delta N}$, and preferentially functionally associated with cellular metabolic process, cell growth and death, replication and repair, response to stimulus, signal transduction; protein folding, sorting and

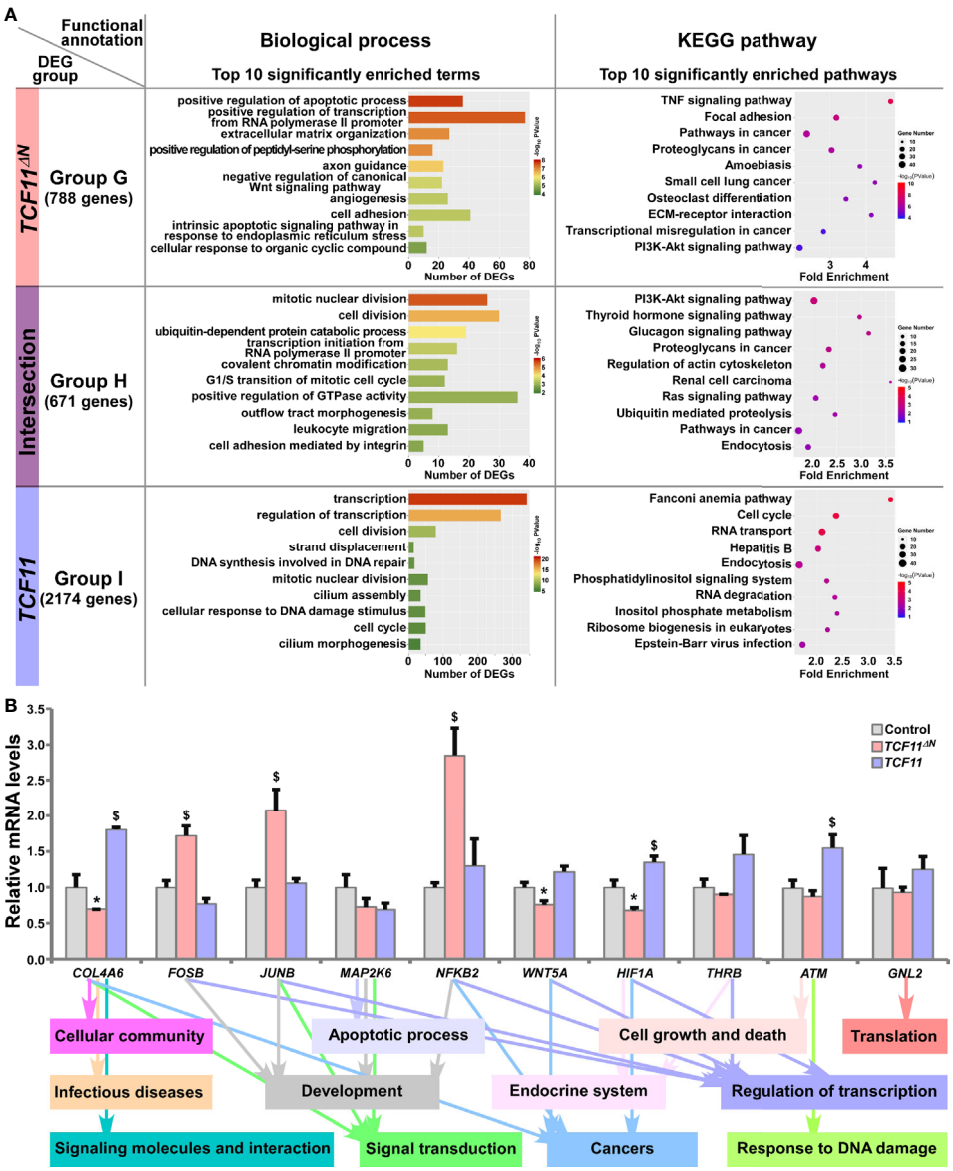


FIGURE 5 | Functional annotation of specific or common DEGs in *TCF11*$^{\Delta N}$ and *TCF11* cells. **(A)** Top 10 of significant biological process terms and pathways enriched by DEGs in Groups G, H, and I were exhibited in histograms and scatterplots, respectively. **(B)** After induced with 1 μg/ml Tet for 12 h, total RNAs were isolated from Control, *TCF11*$^{\Delta N}$ or *TCF11* cell lines before being reversely transcribed into the first strand of cDNA. Subsequently, relative mRNA levels of DEGs that were associated with more functions, along with high expression levels and well significance in Groups G to I, were determined by quantitative real-time PCR in Control, *TCF11*$^{\Delta N}$ and *TCF11* cells. The data are shown as mean ± SEM (n = 3 × 3, $^{*}p < 0.05$; $^{\$}p < 0.05$, when compared to the *Control values*).

degradation, regulation of transcription, translation, transport and catabolism, carbohydrate metabolism and infectious diseases. Notably, six identical pathways were found by comparison between top 10 significantly enriched pathways from Group C (i.e., DEGs regulated by *Nrf2* but not by Nrf1α) and Group G (i.e., DEGs regulated by *TCF11*$^{\Delta N}$ but not by TCF11) (**Figure S4D**). This indicates that TCF11$^{\Delta N}$-regulated genes are much likely to execute somewhat combinational or overlapping functions with Nrf2-target genes.

Subsequently, several unique or common target genes regulated by TCF11 and/or TCF11$^{\Delta N}$ were also validated by quantitative PCR (**Figure 5B**). The results showed that *FOSB*, *JUNB* and *NFKB2* (*nuclear factor kappa B, subunit 2*) were upregulated, while *WNT5A* (*Wnt family member 5A*) was downregulated, by TCF11$^{\Delta N}$, but unaffected by TCF11. Conversely, *COL4A6* (*collagen type IV alpha 6*) and *HIF1A* (*hypoxia inducible factor 1 alpha*) were upregulated by TCF11, but downregulated by TCF11$^{\Delta N}$. Besides, *ATM* (*ATM serine/*

threonine kinase) was also upregulated by TCF11, but roughly unaltered by TCF11$^{\Delta N}$. The putative functions relative to these examined genes were exhibited (as shown in **Figure 5B**, *on the bottom*).

TCF11 and Nrf1α Display Similar but Differential Regulatory Profiles

Those DEGs regulated by TCF11 and/or Nrf1α were grouped by J to L, and then functionally annotated by DAVID, with histograms and scatterplots exhibited (**Figure 6A** and **Table S9**). Comprehensive analysis of the top significantly enriched biological process terms and pathways showed that 1896 DEGs of Group J, by identifying *TCF11*-, but not *Nrf1α* -, expressing cells, were associated with distinct functions in cellular metabolic process, cell cycle, subcellular localization, transport and catabolism, carbohydrate metabolism, regulation of transcription, translation, signal transduction, endocrine system, development and regeneration, infectious diseases and cancers. In Group K, TCF11 and Nrf1α co-regulated 949 DEGs that were involved in cellular metabolic process, cell growth and

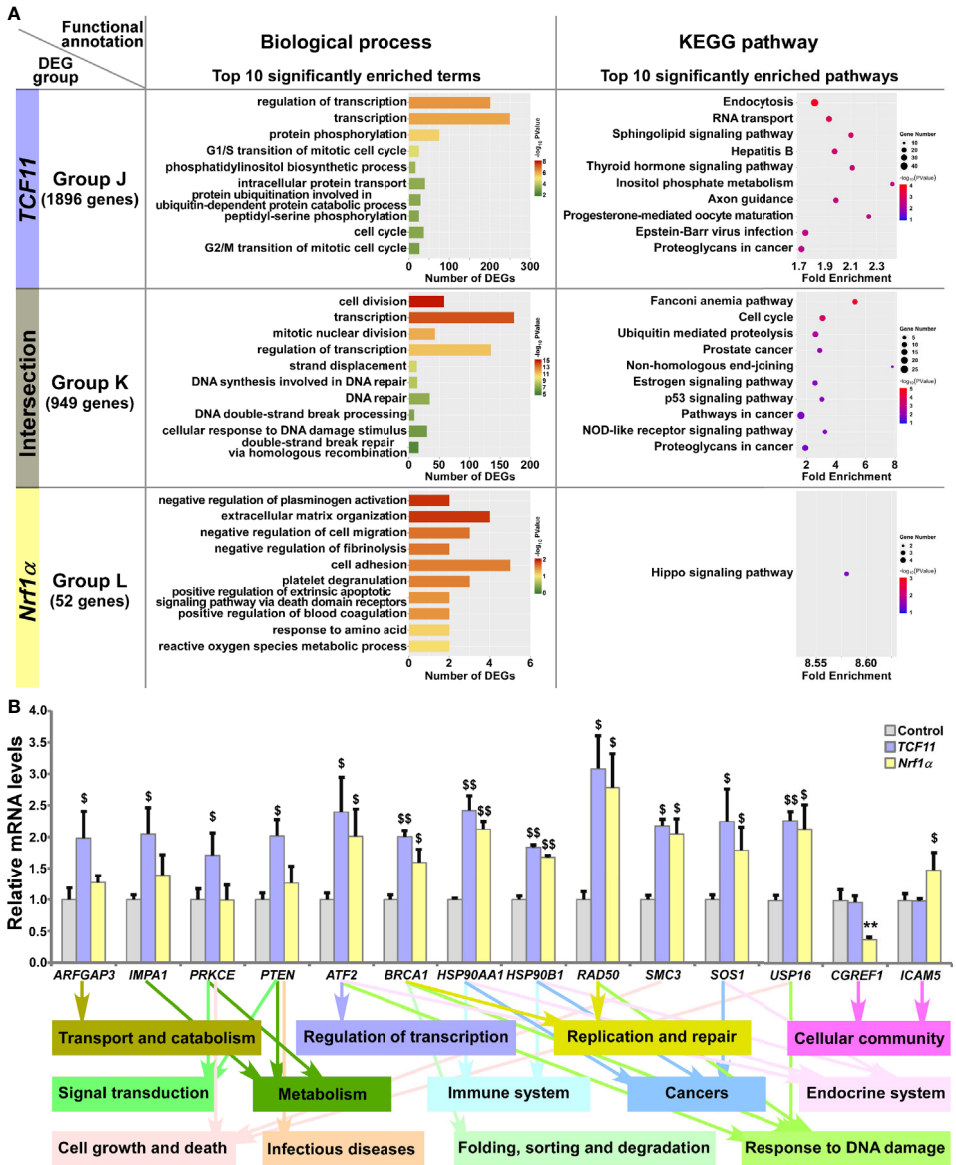


FIGURE 6 | Functional annotation of specific or common DEGs in *TCF11* and *Nrf1α* cells. **(A)** Top 10 of significant biological process terms and pathways enriched by DEGs in Groups J, K, and L were exhibited in histograms and scatterplots, respectively. **(B)** After induced with 1 μg/ml Tet for 12 h, total RNAs were isolated from Control, *TCF11* or *Nrf1α* cell lines and then reversely transcribed into the first strand of cDNA. Subsequently, relevant mRNA levels of DEGs that were associated with more functions as annotated, along with high expression levels and well significance in Groups J to L, were determined by quantitative real-time PCR in Control, *TCF11* and *Nrf1α* cells. The data are shown as mean ± SEM (n = 3 × 3, $^{**}p$ <0.01; $^{\$}p$ <0.05; $^{\$\$}p$ <0.01, when compared to the *Control values*).

death, replication and repair, folding, sorting and degradation, regulation of transcription, response to stimulus, endocrine system, immune system and cancers. In Group L, only 52 DEGs were identified by transcriptional regulation by Nrf1α, but unaffected by TCF11 expression. Their putative functions were associated with cellular community and metabolic process, and subcellular localization, regulation of biological process, signal transduction and response to stimulus. Besides, many of overlapping functions were predicted to exist among distinct combinations of top significantly enriched functions exerted by Group A (i.e., DEGs regulated by Nrf1α but not by Nrf2), Group I (i.e., DEGs regulated TCF11 but not by TCF11$^{\Delta N}$) and Group K (i.e., DEGs co-regulated by TCF11 and Nrf1α) (**Figure S4D**). Thus, it is inferable that TCF11 and Nrf1α display similar regulatory profiles, with a striking disparity from those of TCF11$^{\Delta N}$ or Nrf2.

Amongst the above DEGs, 14 were selected for further quantitation by real-time PCR (**Figure 6B**). The results revealed that eight DEGs were upregulated by both TCF11 and Nrf1α, which included *ATF2* (*activating transcription factor 2*), *BRCA1* (*BRCA1 DNA repair associated*), *HSP90AA1* (*heat shock protein 90 alpha family class A member 1*), *HSP90B1* (*heat shock protein 90 beta family member 1*), *RAD50* (*RAD50 double strand break repair protein*), *SMC3* (*structural maintenance of chromosomes 3*), *SOS1* (*SOS Ras/Rac guanine nucleotide exchange factor 1*) and *USP16* (*ubiquitin specific peptidase 16*). Additional four DEGs, i.e., *ARFGAP3* (*ADP ribosylation factor GTPase activating protein 3*), *IMPA1* (*inositol monophosphatase 1*), *PRKCE* (*protein kinase C epsilon*) and *PTEN* (*phosphatase* and *tensin homolog*) were upregulated by TCF11, but almost unaffected by Nrf1α. Conversely, *Nrf1α* expression caused a decrease in *CGREF1* (*cell growth regulator with EF-hand domain 1*) as accompanied by increased *ICAM5* (*intercellular adhesion molecule 5*), but both genes were unaltered by TCF11. These examined genes were responsible for their putative functions as indicated (**Figure 6B**, *on the bottom*). Altogether, TCF11 and Nrf1α exhibit a similar regulatory profile, but with some of quietly different target genes.

Notably, the relative expression levels of those representative genes from Groups A to L in real-time quantitative PCR are basically consistent with the sequencing data, all with significant positive correlations (as shown in **Figure S5**). Collectively, the overall mapping profile between each of these four transcription factors and the enriched biological functions had been constructed according to the functional annotation and comparative analysis of each group of DEGs (**Figure S6**).

TCF11 Is a More Potent Player Than Nrf1α at Preventing Tumor Xenografts in Nude Mice

As analyzed above, distinct subset of DEGs regulated by Nrf1, TCF11 and/or Nrf2 were annotated for their functional relevancies to cancer development or prevention. In fact, our previous work had revealed that knockdown of Nrf1 caused a significant malignant growth of subcutaneous tumor xenografts in nude mice (54). Thereof, Nrf1α was indicated to act as a dominant tumor-repressor insomuch as to confine oncogenicity of Nrf2 (16, 17). Herein, to corroborate the putative tumor-preventing effects of Nrf1α and TCF11, both CNC-bZIP factors were restored by transfecting the lentivirus expression constructs into HepG2 cells with a specific loss of *Nrf1α*, respectively, as described elsewhere (16). As a consequence, the resulting *Nrf1α*- or *TCF11*-restored cell lines were confirmed by quantitative PCR and immunoblotting to be definitely true (**Figure 7A**).

Subsequently, both *Nrf1α*- and *TCF11*-restored cell lines, alongside with *Nrf1α*$^{-/-}$ cells and wild-type (*WT*) HepG2 cells, were heterotransplanted into distinct groups of those immunodeficient nude mice at their subcutaneous loci as indicated. After tumor formation, the sizes of growing tumors were measured for every two days within ensuing five weeks before the tumor-bearing mice were sacrificed. The resulting data were calculated and shown graphically (**Figure 7B**), as a consequence demonstrating that restoration of *Nrf1α* or *TCF11* enables a significant tumor-preventing effect on the subcutaneous human carcinoma xenografts in nude mice, when compared with those obtained from *Nrf1α*$^{-/-}$ and WT cell lines. In addition, it should be also noted that both *Nrf1α*$^{-/-}$ and WT hepatoma cell lines, as reported previously (16), were herein used as only two co-references in this parallel animal experiments to strengthen their comparability (**Figure 7B**). Furthermore, a series of comparative experiments revealed that *Nrf1α*$^{-/-}$ cell proliferation, migration and invasion were all significantly suppressed by restoration of *Nrf1α* and *TCF11* (**Figures 7C-H**), and the cell-cycle arrest at S-phase, along with the increase in early apoptosis, due to the restoration of them (**Figure S7**). This is further supported by pathohistological results of the hematoxylin and eosin (HE) staining, revealing that malignant progression of *Nrf1α*$^{-/-}$ -derived tumor xenografts was substantially suppressed by restoration of either *Nrf1α* or *TCF11* with complete coagulative necrosis of tumor tissues (**Figure S8**). Collectively, TCF11 acts as a more potent tumor-repressor than Nrf1α, albeit both isoforms are endowed with an intrinsic capability to prevent tumor development and malignant growth.

Nrf1α and TCF11 Regulate Critical Genes for Improving the Survival Rate of HCC Patients

To weigh the practical effects of TCF11, Nrf1α and Nrf2 on human hepatocellular carcinoma (HCC), their relevancies to the overall survival (OS) of HCC patients were firstly investigated (**Figure 8**). For this, HCC-relevant molecules were selected from the UniProt database and their genes were further parsed by the Kaplan–Meier Plotter (55), along with other relevant databases (56–58), to find those markers of being significantly correlated with OS of patients with HCC. As shown in **Figures 8A1**–**A8** and **Table S10**, a better OS was predicted to couple with increased expression of *CHP2* (*calcineurin like EF-hand protein 2*), *CPS1* (*carbamoyl-phosphate synthase 1*), *FOXO1* (*forkhead box O1*), *IRS4* (*insulin receptor substrate 4*) and/or *NKX2-8* (*NK2 homeobox 8*); this was also accompanied by reduced expression of *AKR1B10* (*aldo-keto reductase family 1 member B10*), *EPO*

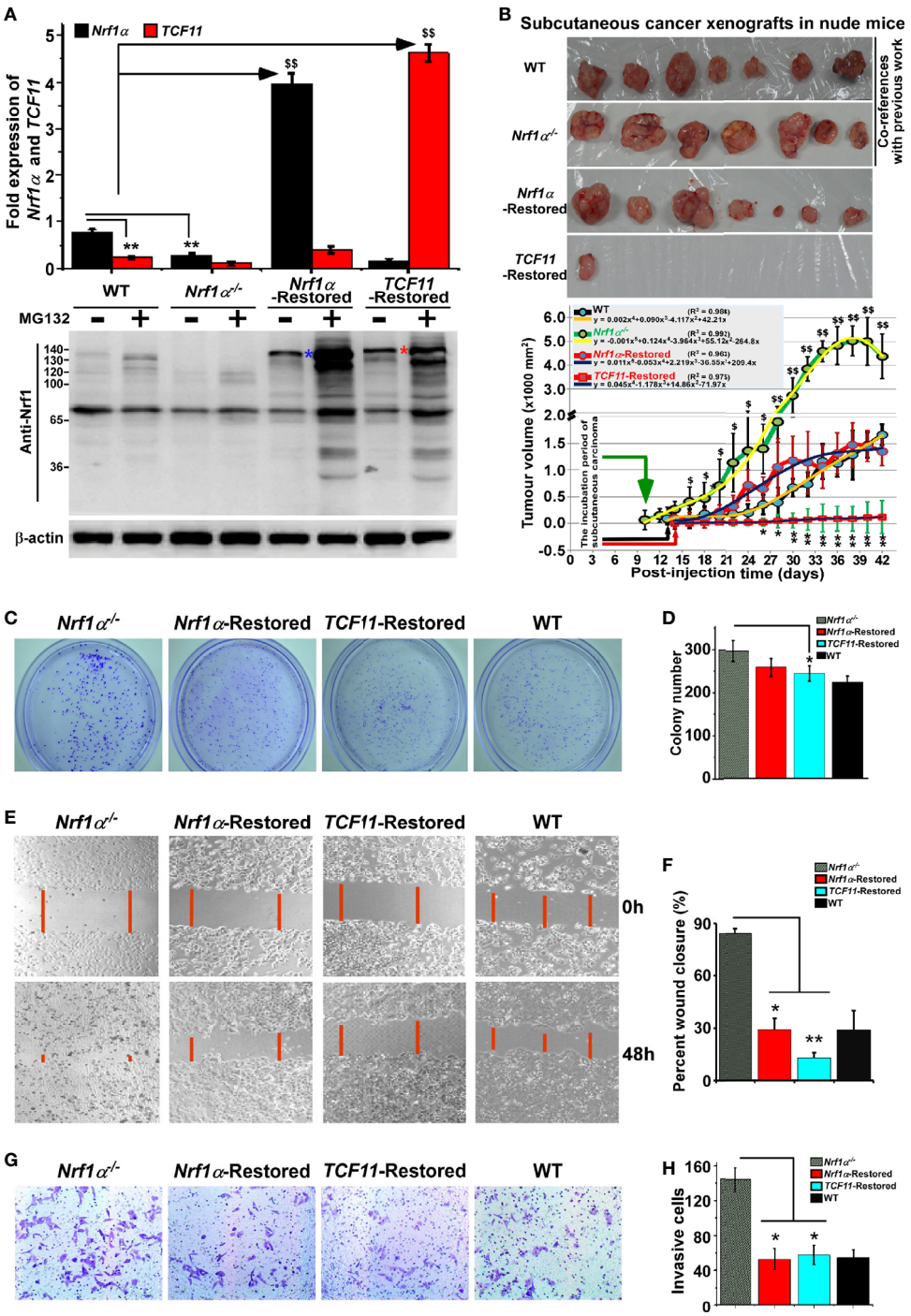


FIGURE 7 | Malgrowth of *Nrf1α*$^{-/-}$-derived hepatoma cells were significantly suppressed by restoration of *Nrf1α* and *TCF11*. **(A)** Both quantitative real-time PCR (*up*) and Western blotting (*down*) were employed to identify the protein and mRNA levels of Nrf1α and TCF11 in *Nrf1α*- and *TCF11*-Restored hepatoma cells. The experimental cells had been treated with or without 10 μmol/L MG132 for 4 h before being harvested for Western blotting. The data are shown as mean ± SEM (n = 3 × 3, $^{**}p < 0.01$; $^{\$\$}p < 0.01$). **(B)** Differences in mouse subcutaneous xenograft tumors derived from wild type HepG2 (WT), *Nrf1α*$^{-/-}$, *Nrf1α*-Restored and *TCF11*-Restored cells were measured in size every two days, before being sacrificed on the 42nd day. The data are shown as mean ± SD (n = 7 per group, $^{*}p < 0.05$; $^{**}p < 0.01$; $^{\$}p < 0.05$; $^{\$\$}p < 0.01$, when compared with the *WT* group). **(C–H)** The soft agar colony formation **(C, D)**, as well as migration **(E, F)** and invasion **(G, H)**, of wild type HepG2 (WT), *Nrf1α*$^{-/-}$, *Nrf1α*-Restored and *TCF11*-Restored cells were examined as described in *Materials and Methods*. The data are shown as mean ± SD (n = 9, $^{*}p < 0.05$, $^{**}p < 0.01$).

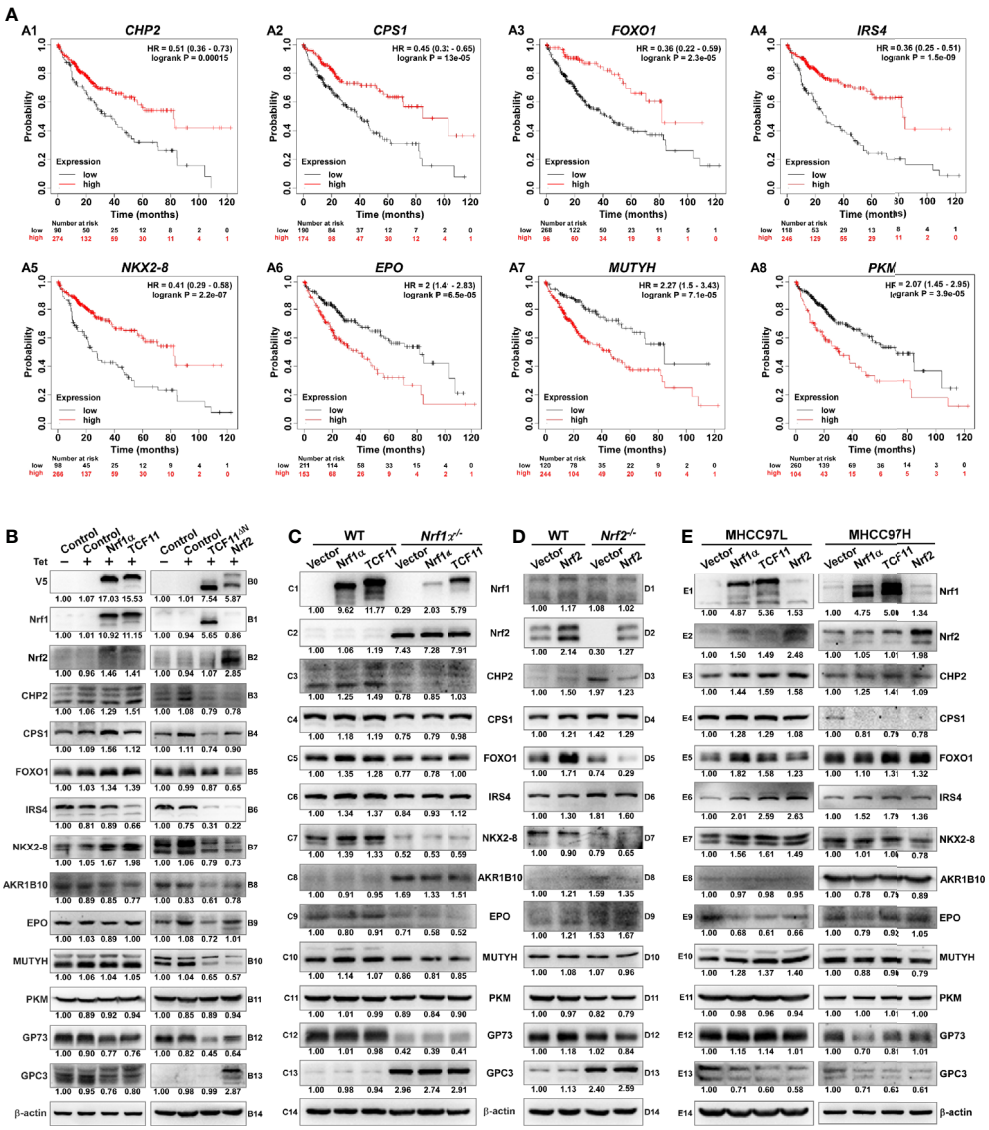


FIGURE 8 | Diverse effects of Nrf1 and Nrf2 on those HCC-relevant proteins that are significantly correlated with the overall survival (OS). **(A)** The correlation analysis between of hepatocellular carcinoma (HCC) associated genes and the relevant OS rates of patients with HCC. **(B)** The protein expression levels of HCC-associated proteins in *Nrf1α*, *TCF11*, *TCF11*$^{\Delta N}$ and *Nrf2* cell models were examined by Western blotting analysis of these experimental cell lines that had been induced with or without 1 μg/ml Tet for 12 h before being harvested. **(C)** Both wild type HepG2 and its derived *Nrf1α*$^{-/-}$ cells were transfected with either Nrf1α or TCF11 expression plasmids, and then allowed for 24-recovery from transfection in the fresh medium before being subjected to Western blotting, to identify the protein levels of HCC-associated proteins. **(D)** Both wild type HepG2 and its derived *Nrf2*$^{-/-}$ cells were transfected with an Nrf2 expression plasmid, and then allowed for 24-recovery from transfection in the fresh medium before being subjected to Western blotting. **(E)** MHCC97L or MHCC97H cells were transfected with each of Nrf1α, TCF11 and Nrf2 expression constructs, and then allowed for 24-recovery from transfection in the fresh medium before being examined by Western blotting to identify abundances of HCC-associated proteins as described above. The intensity of all the immunoblots was calculated and shown on the bottom **(B–E)**.

(*erythropoietin*), *MUTYH* (*mutY DNA glycosylase*) and/or *PKM* (*pyruvate kinase M1/2*). Besides, GP73 (also called GOLM1, golgi membrane protein 1) has been reported to be a potential diagnostic marker for primary HCC (59, 60), while GPC3 (glypican 3) acts as another key biomarker for early diagnosis of human HCC and a rational immunotherapeutic target for HCC (61, 62). Although no significant differences in the expression levels of *Nrf1* or *Nrf2* amongst various tumor tissues from numerous patients with distinct stages of HCC were examined (**Table S10**), such two CNC-bZIP factors had been determined to differentially regulate the progression of HCC (17). Notably, basal expression of Nrf1 in distinct HCC tissues had been shown significantly altered, relative to the equivalent expression in their adjacent para-carcinoma tissues or those expressed in normal liver cells (16). These suggest that Nrf1 and Nrf2 execute distinct functions in the progression of HCC through differentially regulating putative pathophysiological processes. However, discrete isoforms of

Nrf1 were also undistinguished in the above-described databases, but their unique activities are required for being further investigated separately.

The immunoblotting results showed that TCF11 and Nrf1α enabled the normal cells to increase CHP2 and CPS1 abundances, respectively (**Figures 8B3**, **B4**). Although FOXO1 was slightly increased, NKX2-8 was markedly increased, by both TCF11 and Nrf1α (*B5* & *B7*), but IRS4 was down-regulated by TCF11 but not Nrf1α (*B6*). Furtherly, AKR1B10 and EPO were, to greater or less extents, diminished by TCF11 and Nrf1α (*B8* & *B9*). Notably, significant decreases in the two HCC biomarkers GP73 and GPC3 were caused by TCF11 and Nrf1α (*B12* & *B13*). By sharp contrast, most of the above-examined proteins were reduced by TCF11$^{\Delta N}$ and Nrf2 (**Figure 8B**, *right panels*), with an exception that GPC3 was significantly up-regulated by Nrf2, but not TCF11$^{\Delta N}$.

Further examinations revealed that overexpression of TCF11 and Nrf1α in HepG2 cells enabled CHP2, CPS1, FOXO1, IRS4 and NKX2-8 to be enhanced to varying extents, but they were markedly suppressed by loss of *Nrf1α*$^{-/-}$ (**Figure 8C**, *left panels*). Such loss of *Nrf1α* also gave rise to constructive enhancement of Nrf2, AKR1B10 and GPC3, as accompanied by constructive abolishment of EPO and GP73, besides CHP2 and NKX2-8. However, it is intriguing to note that all these constructive changes could not be ameliorated by modest restoration of ectopic TCF11 or Nrf1α into *Nrf1α*$^{-/-}$ cells (**Figure 8C**, *right panels*). By contrast, forced expression of Nrf2 in HepG2 cells caused only a marginal increase in FOXO1 or GP73, but the other examined genes were unaffected (**Figure 8D**, *left panels*). Conversely, loss of *Nrf2* caused a slight increase in CHP2 and GPC3; but the former CHP2 was reduced by ectopic Nrf2 restoration to its basal levels, while the latter GPC3 was not mitigated by ectopic Nrf2 (**Figure 8D**, *right panels*).

The above-described data indicate distinct contributions of TCF11, Nrf1α and Nrf2 to differential or opposite regulation of endogenous expression levels of different, even the same, target genes. This notion was herein substantiated by our further comparative experiments of TCF11, Nrf1α and Nrf2 that had been transfected for their respective overexpression in either MHCC97L or MHCC97H cell lines (**Figure 8E**). As anticipated, abundances of CHP2, CPS1, FOXO1, IRS4, NKX2-8 proteins were increased with their respectively-varying trends by ectopic TCF11, Nrf1α or Nrf2 expression in MHCC97L cells (**Figure 8E**, *left panels*), whereas such events did not occur in MHCC97H cells (**Figure 8E**, *right panels*). Besides, both EPO and GPC3 were suppressed by ectopic expression of TCF11, Nrf1α and Nrf2 in MHCC97L cells, but almost unaffected in MHCC97H cells with hyper-expression of Nrf2. Such discrepancies in these gene expressions are much likely contributable to distinct metastatic potentials and other malignant properties of between MHCC97L and MHCC97H cell lines. Subsequently, an integrated interaction network regulated by Nrf1 and Nrf2 (**Figure 9A**) was built on the basis of the protein interaction database STRING (63), in combination with our experimental results as shown above. The distinct expression levels of those indicated genes that make up the network in stable expression cells (**Figure 9B**) or knockout model cells (**Figure 9C**), were also indicated in the heatmap, respectively. Taken altogether, it is demonstrated that Nrf1α and TCF11, but not Nrf2, are conferred for an intrinsic capability to regulate those genes critical for improving the survival rate of patients with HCC, such that Nrf1α or TCF11 perform a strikingly disparate effect from that of Nrf2 on human hepatoma (**Figure 9D**).

DISCUSSION

In this study, we have established four controllable cell models for stably expressing *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2*, which all occur naturally with their respective intact structural domains, albeit they were tagged C-terminally by a neutral V5 epitope. Of striking note, Nrf1α and TCF11 are two longer isoforms of the ER-resident Nrf1 with changing gears in their overall transcriptional activity, because both proteins are tightly controlled by their unique topobiological rheostat modules from the ER to enter the nucleus. For instance, the NST (Asn/Ser/Thr-rich) domain transactivation of Nrf1α/TCF11, as well as SKN-1 (Skinhead-1) and CncC (Cap′n′collar isoform C), is also required for re-editing of their indicated asparagines to aspartates in this region during dynamic dislocation from the lumen to extra-ER subcellular compartments (64, 65). Upon lack of such amino acid re-editing by peptide: N-glycosidase (PNG1), this causes an evident decrease in the rheostat capacity of Nrf1α/TCF11, and even loss of *PNG1* or its mutations results in inactivation of Nrf1α/TCF11, manifesting inflammation and adrenal insufficiency (66–68). Furtherly, the rheostat capacity of Nrf1α/TCF11 is also monitored by its reversible ubiquitination and deubiquitination during its selective ER-associated proteasome-regulated processing to multiple isoforms (69, 70).

During our manuscript preparation, removal of the N-terminal ER signal peptide-containing 104-aa or 121-aa regions from Nrf1α/TCF11 to yield two artificially-truncated mutants (i.e., Nrf1$^{\Delta N104}$ or Nrf1$^{\Delta N121}$) had been reported by both independent Bollong's and Ooi's laboratories (44, 45). It is, to our great surprise, seen that such artificial mutants Nrf1$^{\Delta N104}$ or Nrf1$^{\Delta N121}$ were asserted to represent two constitutive processed Nrf1 activators, although they each retain a negative PEST region. Thereby, our study of *TCF11*$^{\Delta N}$ showed an N-terminal 2–156 residues-truncated mutant (i.e., TCF11$^{\Delta N2-156}$), which may also arise from a naturally-splicing transcript (41). However, TCF11$^{\Delta N}$ is herein identified as a mimic Nrf2 factor, because both share conserved structural domains (**Figure 1A**). This seems to totally contradict the purpose of Nrf1$^{\Delta N104}$ or Nrf1$^{\Delta N121}$ by both Bollong's and Ooi's colleagues (44, 45). Additional two mutants of caNrf2 to block its Keap1-binding activity were also utilized in their experimental settings. Besides, it should be of crucial importance to note that a strong acidic property of 3×Flag, wherever it was tagged at the N-terminal or C-terminal ends of a given protein, exerts distinctive or even opposing effects on its topobiological folding and dynamic moving in and out of membranes and other subcellular

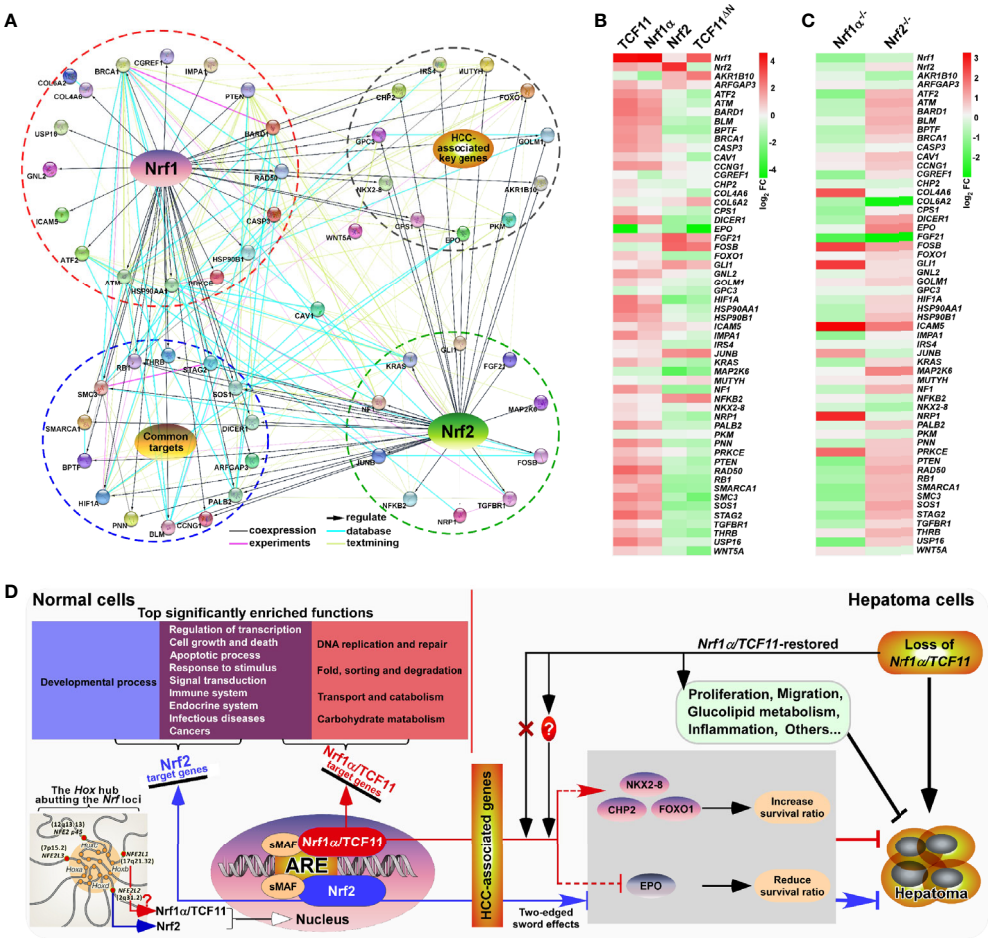


FIGURE 9 | Relationship and difference between the regulation patterns of Nrf1 and Nrf2 on their targets. **(A)** The functional protein association networks of targets regulated by Nrf1 or Nrf2. Of note, the protein-protein associations are determined by various ways, which are thus represented by different colored edges as indicated. **(B)** The heatmap of the sequencing expression of genes, which are composed of the network, with distinct expression levels in *TCF11*, *Nrf1α*, *Nrf2* and *TCF11*$^{\Delta N}$ cell lines. The color of the nodes in the heatmap represents the value of $\log_2$ (fold change) as shown in the color bars, indicating the gene expression trend as compared with the control group (upregulation or downregulation, were marked in red or green, respectively). **(C)** The heatmap of the sequencing expression of genes in *Nrf1α*$^{-/-}$ and *Nrf2*$^{-/-}$ cell lines. The color of the nodes in the heatmap represents the value of $\log_2$ (fold change) as shown in the color bars, indicating the gene expression trend as compared with the wild-type HepG2 cells. **(D)** A comprehensive regulatory model is proposed to reveal the different effects of Nrf1 and Nrf2 on hepatoma (*right panel*). In addition, the *Hox* hub abutting the distinct *Nrf* loci was also indicated (*at the lower left corner*).

compartments (69, 71). Overall, our study has provided four cell tools to evaluate similarities yet differences in between intact TCF11, TCF11$^{\Delta N}$, Nrf1α and Nrf2-regulatory profiles, which are stimulated at the Tet-inducibly controllable levels.

Subsequently, these four cell lines were determined by transcriptomic sequencing in an integrative combination with routine reductionist approaches. As expected, comprehensive functional annotation of TCF11, TCF11$^{\Delta N}$, Nrf1α and Nrf2 unraveled that they are likely to perform diverse biological functions by combinationally overlapping or competitively opposing modules within distinct cellular networks, apart from their unique functions. Thereby, a further understanding of their respectively regulated targets and functions should be achieved only by making one of them function alone. For this end, our established cell lines stably expressing *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2* respectively, are an invaluable tool to gain insights into distinguishing their regulatory profiles, while they are allowed to function alone. Consequently, the regulatory patterns of Nrf1α and TCF11 are similar yet different, which are strikingly disparate from that of Nrf2, even with certain opposite effects on the same targets. Such distinctions between Nrf1α/TCF11 and Nrf2 are attributable to differences in their primary structures and functional subcellular locations. This is due to the objective fact that the ER-associated domains of Nrf1α/TCF11 determine its unique membrane topobiology and post-synthetic processing mechanisms, which enables it to be distinguishable from Nrf2. Yet, once the ER-targeting signal peptide-adjoining region were deleted to yield TCF11$^{\Delta N}$, this N-terminally truncated isoform does exhibit a similar regulatory pattern to that of Nrf2. However, this notion appears to totally

contradict the conclusions drawn from two so-called constitutive activators Nrf1$^{\Delta N104}$ or Nrf1$^{\Delta N121}$ by Bollong's and Ooi's groups (44, 45). As such, their analyses by combining ChIP-seq and RNA-seq data revealed that Nrf1$^{\Delta N121}$ (or Nrf1$^{\Delta N104}$) had 3.19-fold numbers of target-binding peaks than those of caNrf2-binding targets. Of note, over 75% of Nrf1$^{\Delta N121}$-binding targets are focused preferentially on the consensus ARE sequence (5′-TGAC/GnnnGC-3′) flanked by AT-enriched motifs, while only less than 45% of caNrf2-binding targets are recruited on ARE-like sequences flanked by GC-enriched motifs (44). It is inferable that Nrf1 exerts its unique biological functions by predominantly regulating ARE-driven cognate genes, whereas Nrf2 has a widely-varying capacity to elicit its promiscuous roles in diversely mediating non-ARE-battery genes. Altogether, with our previous work (17, 37) demonstrating that both Nrf1α/TCF11 and Nrf2 can mutually influence each other by their inter-regulating at distinct levels, this leads to the formation of a steady-state regulatory network system finely monitored by a negative feedback loop to maintain the cell homeostasis and organ integrity.

Since early discovery of Nrf2 required for regulating antioxidant and detoxification genes (72), the overwhelming majority of investigations in this field have been focused disproportionately on this CNC-bZIP factor and its negative regulator Keap1 in response to redox stress, but also provided myriad insights into its promiscuous roles in biology and medicine, as well as drug development (73–75). However, aside from these studies by routine reductionist approaches, the whole genome-widely integrative analyses uncovered the underlying facts that Nrf1, but not Nrf2, is essentially required for regulating important homeostatic and metabolic genes involved in the normal growth and development throughout life process. This is fully consistent with the experimental evidence showing that loss of Nrf1 results in mouse embryonic lethality and also causes adult pathological phenotypes varying within its gene-targeting mutant organs [reviewed by (70)]. Such bad consequences are attributable to severe endogenous oxidative stress and fatal defects in redox metabolism reprogramming and relevant constitutive gene expression profiles of *Nrf1*-deficient cells [this study and (12)]. Collectively, these facts demonstrate that Nrf1 is a predominant determinist factor of cellular constitutive redox metabolic homeostasis with organ integrity. That is to say, Nrf1 has a potent capacity to contribute to the steady-state robustness of cell physiological homeostasis. By contrast, Nrf2 is much likely to make a major contribution to the sensitive plasticity of cell homeostasis, because it has been accepted as a master regulator in response to diverse stresses. This is also supported by the facts of no obvious phenotypes resulting from global loss of *Nrf2* in mice, but its deficiency leads to more susceptibility to chemical carcinogens and diverse stresses than wild-type controls (76, 77).

Such differential contributions of Nrf1 and Nrf2 to cell homeostasis robustness and plasticity are also likely resulted from a striking evolutionary conservation of their CNC-bZIP family in distinct species during nature selection. Comparative genomics analyses revealed that, although this CNC-bZIP family expansion and diversification have occurred in vertebrates, Nrf1 is viewed as a living fossil that is more ancient than Nrf2, because it is conserved closely to the only CNC in *Drosophila melanogaster*, SKN-1 in *Caenorhabditis elegans* and each of newly-identified Nach factors in simple multicellular eukaryotes, but not in unicellular protozoans or other prokaryotes beyond *Endozoicomonas* (4). This implies that they have evolved for multicellular cooperative selection. For instance, the *C. elegans* SKN-1, like Nrf1, is selectively processed by post-transcriptional and post-translational ways to yield distinct isoforms, one of which is N-terminally truncated (by analogy of TCF11$^{\Delta N}$) but manifested the Nrf2-like functions, albeit with no existence of Keap1-like orthologues (65, 78). Such the highly evolutional conservativity of Nrf1, but not of Nrf2, strongly demonstrates that it is indispensable for basal constitutive contributions to orchestrating the ensemble of critical gene regulatory networks, such that the normal cell homeostasis and organ integrity have been perpetuated in the entire course of life.

In fact, the CNC-bZIP family members (p45, Nrf1, Nrf2 and Nrf3) are, though diversified in vertebrates, topologically organized together with the developmental *Hox* gene clusters (as shown in **Figure 9D**), upon forming the PcG hub to maintain their genes in a silent-but-poised state (79). A recent ChIP data analysis revealed that most of Nrf1-binding sites are focused closely to the canonical ARE sequences that are widely located in its cognate gene promoter and distal intergenic enhancer regions (i.e., dynamic genome-topology as a functional primer is built by spatial interactions of distal enhancers with the proximal promoters in the same genes or even between different genes in chromosomal architecture), whereas Nrf2-binding sites are promiscuously loosed to ARE-like or no-ARE sequences that are, however, constrained narrowly in its target gene promoters (44). Collectively, the interplay between selective transcription factor (e.g., CNC-bZIP)-regulated gene programming and genome conformation is surmised as a driving force for cell-fate decision with distinct type-featured identifications.

Notably, the versatile Nrf2 acts *de facto* as a promiscuous, not essential, player in its biology, because it is dispensable for normal growth and development in mice with no phenotypes (e.g., cancer) resulting from its genetic loss (74, 77). Contrarily, accumulating evidence clearly demonstrates that hyperactive Nrf2 promoted cancer development and malignance, because it is relevant to most of cancer hallmarks (73). This implies that, except that Nrf2 can exert a significant cytoprotective function against diverse stresses so as to confer the cells to be acquired for the adaptive responses, its long-term hyper-activation can also critically inspire potential cancerous cells to be extricated from being rigidly controlled and confined by inevitability of the host multicellular cooperative evolution. The notion is supported by transcriptome sequencing of *Nrf2*-deficient cells, because this loss of Nrf2 causes hepatoma to be significantly ameliorated or completely prevented (17). In the other way round, aberrant accumulation of hyperactive Nrf2 in *Nrf1*-deficient cells (or livers) leads to further malignant transformation of human hepatocellular carcinoma (HCC).

In particular, the spontaneous development of NASH-based inflammation and hepatoma are resulted from severe endogenous oxidative stress and fatal defects in basal constitutive gene

expression affected by its genetic instability in mouse *Nrf1*-deficient livers (80–82). Of note, TCF11 is absent in mice and has a low proportion in all human hepatoma cell lines, but it, together with Nrf1α at a 1:1 ratio, exists in human normal cells (**Figure S9**). Upon *Nrf1α*-specific knockout from hepatoma cells, this results in cancer malignant growth and metastasis to the lungs in xenograft model mice (16, 54). The *Nrf1α*$^{-/-}$-derived cancer deterioration is definitely resulted from severe oxidative stress, fateful defects in the redox metabolism reprogramming, and marked dysfunctions of fate-decisive gene regulatory networks, along with critical aberrant signaling transduction networks (12, 17). However, it is, to our great surprise, found that *Nrf1α*$^{-/-}$-led malignant transformation is markedly alleviated and prevented by silencing of Nrf2; this is also accompanied by blocking oxidative stress upon Nrf2 knockdown (12, 17). More excitingly, restoration of *Nrf1α* is allowed for significant mitigation of *Nrf1α*$^{-/-}$-exacerbated tumor to similar wild-type extents, while its malignant growth is further suppressed or even abolished by *TCF11* restoration, with complete coagulative necrosis of tumor tissues (**Figures 7B** and **S8**). This demonstrates that TCF11 is a potent tumor-suppressor than Nrf1α, while Nrf2 is a tumor-promotor, particularly in the *Nrf1α*$^{-/-}$ case. Altogether, loss of Nrf1α/TCF11's function, with hyperactive Nrf2 accumulation, results in liver cancer initiation and progression. This consequence is likely originated from endogenous oxidative stress-induced damages of mutant cells in aberrant metabolic inflammatory microenvironments. This gives rise to an evolvable selection force in so much of *Darwinian* dynamics, so that a clade of the mutant cancerous-prone cells is endowed with a self-defined fitness function and also specified to acquire for an independent of the host team optimum cooperative confinements, in order to behave themselves with own properties of cell division, proliferation and migration during carcinogenesis and progression.

Moreover, this work further demonstrates that the tumor-preventing effect of Nrf1α and TCF11 is accompanied by the constitutive activation or repression of critical genes for improving the overall survival of patients with hepatocellular carcinoma (**Figure 9**). This is because these changes can be ameliorated by either Nrf1α- or TCF11-restored lines (as mentioned in **Figure 7**). However, their activation or repression of some genes could not be ameliorated by compensation of ectopic *Nrf1α* and *TCF11* transfection into those *Nrf1α/TCF11*-deficient cancer cells. This is surmised to be the relevance of those given genes closely to the contexts within genome topological conformations, enabling ectopic Nrf1α/TCF11 factors to be allowed or forbidden for direct access to target genes (as shown in the lower left corner of **Figure 9D**).

Lastly, it should also be noted that our subcellular fractionation and immunofluorescence results have unraveled that a certain fraction of TCF11, TCF11$^{\Delta N}$, Nrf1α or Nrf2 can be allowed for spatial translocation from the cytoplasmic to the nuclear compartments, in which they gain access to target genes (**Figure S3**). Interestingly, these nuclearly-located fractions of TCF11, TCF11$^{\Delta N}$, Nrf1α or Nrf2 are rapidly degraded. This is due to this fact that inhibition of their proteasomal degradation by MG132 causes an obvious increase in each protein expression level of TCF11, TCF11$^{\Delta N}$, Nrf1α or Nrf2, and they were also markedly recovered in their nuclear fractions of MG132-treated cells (**Figure S3**). Altogether, these demonstrate that only after TCF11, TCF11$^{\Delta N}$, Nrf1α or Nrf2 enter the nucleus and stay for a given time in this subcellular compartment, they are conferred to exert their putative physiological functions by regulating transcriptional expression profiles of distinct subsets of target genes. On the contrary, it is inferable that if they are degraded totally, as reported previously, in the extra-nuclear cytoplasmic compartments, just under normal physiological conditions, none of their remaining proteins can be allowed for dynamic repositioning into the nucleus, so that no physiological functions of cognate genes would be regulated by these CNC-bZIP factors. However, this notion cannot hold true in fact, albeit the detailed mechanisms are required to be further explored.

CONCLUSIONS

In summary, four useful model cell lines stably expressing *TCF11*, *TCF11*$^{\Delta N}$, *Nrf1α* or *Nrf2*, not mutants, are yielded herein. Their transcriptional expression levels are controlled by a tetracycline-inducible switch, but not monitored by one of redox inducers (e.g., sulforaphane or *tert*-Butylhydroquinone) or ER stressors (e.g., tunicamycin or thapsigargin). These cells were subjected to the integrative systems biology analyses of their omics data with routine reductionist approaches. To explore the essential distinctions between Nrf1α/TCF11 and Nrf2 in their contributions to critical constitutive gene expression profiles for basal redox metabolism, normal growth, development, cell homeostasis and organ integrity, we have provided holistic relevant data as much as possible, in the present digital network era. Notably, some seemingly-paradoxical data are still retained, because they may serve as vital nodes of a few negative feedback (and feedforward) regulatory circuits existing in the self-organizing systems of life. This is for the sake of an objective truth in life. Significantly, it is demonstrated that TCF11 serves as a more potent tumor-suppressor than Nrf1α at preventing cancer development and progression. This is defined by similar yet different regulatory profiles of both isoforms, with a striking disparity from Nrf2. Rather, a naturally-spliced mutant TCF11$^{\Delta N}$ resembles Nrf2 with largely consistent structure and function in regulating similar sets of target genes. Interestingly, the tumor-preventing effect of Nrf1α/TCF11 seems to be accompanied by certain constitutive activation or repression of critical genes for improving the overall survival rates of patients with hepatoma. Once loss of Nrf1α/TCF11's function, with hyperactive Nrf2 accumulated, this leads to severe endogenous oxidative damages, aberrant redox metabolic inflammation, and ultimate spontaneous hepatoma. Such genetic and nongenetic drivers could be integrated as a selection force in *Darwinian* dynamics to enable for stochastic speciation of *Nrf1*-deficient cells during carcinogenesis and ensuing progression, albeit this is required for deeply studies. Taken together, this study provides a holistic perspective to give a better understanding of essential differences between Nrf1α/TCF11 and Nrf2 in biology and medicine. Thereby, this facilitates drug discovery to induce Nrf1α/TCF11 as a new potent chemoprevention target against cancer.

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

The animal study was reviewed and approved by the University Laboratory Animal Welfare and Ethics Committee (with two institutional licenses SCXK-PLA-20120011 and SYXK-PLA-20120031).

AUTHOR CONTRIBUTIONS

MW designed and performed most of the experiments and bioinformatic analyses, made all figures and wrote this manuscript draft. YR constructed the *Nrf1α*- and *TCF11*-restored cell lines and performed the subcutaneous tumor xenografts experiment. Both SH and KL helped MW with some molecular cloning to create several expression constructs and performed western blotting. LQ established both $Nrf1\alpha^{-/-}$ and $Nrf2^{-/-}$ cell lines. Lastly, YZ designed and supervised this study, analyzed and parsed all the data, helped to prepare all figures, wrote and revised this 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 (projects 81872336 and 82073079) awarded to Prof. Yiguo Zhang (University of Chongqing, China).

ACKNOWLEDGMENTS

We are greatly thankful to all the other present and past members of Prof. Zhang's laboratory for giving critical discussion and invaluable help with this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.707032/full#supplementary-material

Supplementary Figure 1 | Structural comparison of distinct Nrf1 isoforms with Nrf2.

Supplementary Figure 2 | A detailed diagram of Flp-In system.

Supplementary Figure 3 | The distribution of *TCF11*, $TCF11^{\Delta N}$, *Nrf1α* and *Nrf2* in each of their stably-expressing cells.

Supplementary Figure 4 | Statistical analysis of the transcriptome sequencing data.

Supplementary Figure 5 | Identification of a significant correlation between real-time quantitative PCR results and transcriptome sequencing data.

Supplementary Figure 6 | Similarities and differences amongst those mainly regulatory profiles of TCF11, $TCF11^{\Delta N}$, Nrf1α and Nrf2.

Supplementary Figure 7 | Alterations in the cell cycle and apoptosis caused by restoration of *Nrf1α* and *TCF11*.

Supplementary Figure 8 | The HE staining results of tumors in each group.

Supplementary Figure 9 | The relative proportion of *Nrf1α* and *TCF11* in different liver cancer cells and normal cells.

Supplementary Table 1 | The primers for construction or qPCR used in this paper.

Supplementary Table 2 | The differentially expressed genes regulated by TCF11.

Supplementary Table 3 | The differentially expressed genes regulated by Nrf1α.

Supplementary Table 4 | The differentially expressed genes regulated by Nrf2.

Supplementary Table 5 | The differentially expressed genes regulated by $TCF11^{\Delta N}$.

Supplementary Table 6 | The functional annotation of DEGs from group A to C.

Supplementary Table 7 | The functional annotation of DEGs from group D to F.

Supplementary Table 8 | The functional annotation of DEGs from group G to I.

Supplementary Table 9 | The functional annotation of DEGs from group J to L.

Supplementary Table 10 | Analysis of the significance of the potential HCC-associated proteins in relation of liver cancer in different databases.

REFERENCES

1. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The Human Transcription Factors. *Cell* (2018) 175:598–9. doi: 10.1016/j.cell.2018.09.045
2. Sykiotis GP, Bohmann D. Stress-Activated Cap'n'collar Transcription Factors in Aging and Human Disease. *Sci Signal* (2010) 3:re3. doi: 10.1126/scisignal.3112re3
3. Zhang Y, Xiang Y. Molecular and Cellular Basis for the Unique Functioning of Nrf1, an Indispensable Transcription Factor for Maintaining Cell Homoeostasis and Organ Integrity. *Biochem J* (2016) 473:961–1000. doi: 10.1042/BJ20151182
4. Zhu YP, Wang M, Xiang Y, Qiu L, Hu S, Zhang Z, et al. Nach Is a Novel Subgroup at an Early Evolutionary Stage of the CNC-bZIP Subfamily Transcription Factors From the Marine Bacteria to Humans. *Int J Mol Sci* (2018) 19. doi: 10.3390/ijms19102927
5. Kim HM, Han JW, Chan JY. Nuclear Factor Erythroid-2 Like 1 (NFE2L1): Structure, Function and Regulation. *Gene* (2016) 584:17–25. doi: 10.1016/j.gene.2016.03.002
6. Chan JY, Kwong M, Lu R, Chang J, Wang B, Yen TS, et al. Targeted Disruption of the Ubiquitous CNC-bZIP Transcription Factor, Nrf-1,

Results in Anemia and Embryonic Lethality in Mice. *EMBO J* (1998) 17:1779–87. doi: 10.1093/emboj/17.6.1779
7. Leung L, Kwong M, Hou S, Lee C, Chan JY. Deficiency of the Nrf1 and Nrf2 Transcription Factors Results in Early Embryonic Lethality and Severe Oxidative Stress. *J Biol Chem* (2003) 278:48021–9. doi: 10.1074/jbc.M308439200
8. Kim J, Xing W, Wergedal J, Chan JY, Mohan S. Targeted Disruption of Nuclear Factor Erythroid-Derived 2-Like 1 in Osteoblasts Reduces Bone Size and Bone Formation in Mice. *Physiol Genomics* (2010) 40:100–10. doi: 10.1152/physiolgenomics.00105.2009
9. Xing W, Singgih A, Kapoor A, Alarcon CM, Baylink DJ, Mohan S. Nuclear factor-E2-related Factor-1 Mediates Ascorbic Acid Induction of Osterix Expression *Via* Interaction With Antioxidant-Responsive Element in Bone Cells. *J Biol Chem* (2007) 282:22052–61. doi: 10.1074/jbc.M702614200
10. Lee CS, Lee C, Hu T, Nguyen JM, Zhang J, Martin MV, et al. Loss of Nuclear Factor E2-related Factor 1 in the Brain Leads to Dysregulation of Proteasome Gene Expression and Neurodegeneration. *Proc Natl Acad Sci U S A* (2011) 108:8408–13. doi: 10.1073/pnas.1019209108
11. Lee CS, Ho DV, Chan JY. Nuclear Factor-Erythroid 2-Related Factor 1 Regulates Expression of Proteasome Genes in Hepatocytes and Protects Against Endoplasmic Reticulum Stress and Steatosis in Mice. *FEBS J* (2013) 280:3609–20. doi: 10.1111/febs.12350
12. Zhu YP, Zheng Z, Xiang Y, Zhang Y. Glucose Starvation-Induced Rapid Death of Nrf1alpha-Deficient, But Not Nrf2-Deficient, Hepatoma Cells Results From Its Fatal Defects in the Redox Metabolism Reprogramming. *Oxid Med Cell Longev* (2020) 2020:4959821. doi: 10.1155/2020/4959821
13. Hirotsu Y, Hataya N, Katsuoka F, Yamamoto M. NF-E2-related Factor 1 (Nrf1) Serves as a Novel Regulator of Hepatic Lipid Metabolism Through Regulation of the Lipin1 and PGC-1Beta Genes. *Mol Cell Biol* (2012) 32:2760–70. doi: 10.1128/MCB.06706-11
14. Hirotsu Y, Higashi C, Fukutomi T, Katsuoka F, Tsujita T, Yagishita Y, et al. Transcription Factor NF-E2-Related Factor 1 Impairs Glucose Metabolism in Mice. *Genes Cells* (2014) 19:650–65. doi: 10.1111/gtc.12165
15. Berg DT, Gupta A, Richardson MA, O'Brien LA, Calnek D, Grinnell BW. Negative Regulation of Inducible Nitric-Oxide Synthase Expression Mediated Through Transforming Growth Factor-Beta-Dependent Modulation of Transcription Factor TCF11. *J Biol Chem* (2007) 282:36837–44. doi: 10.1074/jbc.M706909200
16. Ren YG, Qiu L, Lu FL, Ru XF, Li SJ, Xiang YC, et al. TALENs-Directed Knockout of the Full-Length Transcription Factor Nrf1 Alpha That Represses Malignant Behaviour of Human Hepatocellular Carcinoma (HepG2) Cells. *Sci Rep-Uk* (2016) 6:23775. doi: 10.1038/Srep23775
17. Qiu L, Wang M, Hu SF, Ru XF, Ren YG, Zhang ZW, et al. Oncogenic Activation of Nrf2, Though as a Master Antioxidant Transcription Factor, Liberated by Specific Knockout of the Full-Length Nrf1 Alpha That Acts as a Dominant Tumor Repressor. *Cancers* (2018) 10:520. doi: 10.3390/cancers10120520
18. Kwong M, Kan YW, Chan JY. The CNC Basic Leucine Zipper Factor, Nrf1, is Essential for Cell Survival in Response to Oxidative Stress-Inducing Agents. Role for Nrf1 in Gamma-Gcs(L) and Gss Expression in Mouse Fibroblasts. *J Biol Chem* (1999) 274:37491–8. doi: 10.1074/jbc.274.52.37491
19. Chen L, Kwong M, Lu R, Ginzinger D, Lee C, Leung L, et al. Nrf1 is Critical for Redox Balance and Survival of Liver Cells During Development. *Mol Cell Biol* (2003) 23:4673–86. doi: 10.1128/mcb.23.13.4673-4686.2003
20. Ma Q. Role of Nrf2 in Oxidative Stress and Toxicity. *Annu Rev Pharmacol Toxicol* (2013) 53:401–26. doi: 10.1146/annurev-pharmtox-011112-140320
21. Mitsuishi Y, Taguchi K, Kawatani Y, Shibata T, Nukiwa T, Aburatani H, et al. Nrf2 Redirects Glucose and Glutamine Into Anabolic Pathways in Metabolic Reprogramming. *Cancer Cell* (2012) 22:66–79. doi: 10.1016/j.ccr.2012.05.016
22. Lee SB, Sellers BN, DeNicola GM. The Regulation of NRF2 by Nutrient-Responsive Signaling and Its Role in Anabolic Cancer Metabolism. *Antioxid Redox Signal* (2018) 29:1774–91. doi: 10.1089/ars.2017.7356
23. Jayakumar S, Pal D, Sandur SK. Nrf2 Facilitates Repair of Radiation Induced DNA Damage Through Homologous Recombination Repair Pathway in a ROS Independent Manner in Cancer Cells. *Mutat Res* (2015) 779:33–45. doi: 10.1016/j.mrfmmm.2015.06.007
24. Fan Z, Wirth AK, Chen D, Wruck CJ, Rauh M, Buchfelder M, et al. Nrf2-Keap1 Pathway Promotes Cell Proliferation and Diminishes Ferroptosis. *Oncogenesis* (2017) 6:e371. doi: 10.1038/oncsis.2017.65
25. Murakami S, Motohashi H. Roles of Nrf2 in Cell Proliferation and Differentiation. *Free Radical Biol Med* (2015) 88:168–78. doi: 10.1016/j.freeradbiomed.2015.06.030
26. Torrente L, Sanchez C, Moreno R, Chowdhry S, Cabello P, Isono K, et al. Crosstalk Between NRF2 and HIPK2 Shapes Cytoprotective Responses. *Oncogene* (2017) 36:6204–12. doi: 10.1038/onc.2017.221
27. Niture SK, Jaiswal AK. Nrf2-induced Antiapoptotic Bcl-xL Protein Enhances Cell Survival and Drug Resistance. *Free Radical Biol Med* (2013) 57:119–31. doi: 10.1016/j.freeradbiomed.2012.12.014
28. Li L, Pan H, Wang H, Li X, Bu X, Wang Q, et al. Interplay Between VEGF and Nrf2 Regulates Angiogenesis Due to Intracranial Venous Hypertension. *Sci Rep* (2016) 6:37338. doi: 10.1038/srep37338
29. Ji X, Wang H, Zhu J, Zhu L, Pan H, Li W, et al. Knockdown of Nrf2 Suppresses Glioblastoma Angiogenesis by Inhibiting Hypoxia-Induced Activation of HIF-1alpha. *Int J Cancer* (2014) 135:574–84. doi: 10.1002/ijc.28699
30. Kim TH, Hur EG, Kang SJ, Kim JA, Thapa D, Lee YM, et al. NRF2 Blockade Suppresses Colon Tumor Angiogenesis by Inhibiting Hypoxia-Induced Activation of HIF-1alpha. *Cancer Res* (2011) 71:2260–75. doi: 10.1158/0008-5472.CAN-10-3007
31. Bauer AK, Cho HY, Miller-Degraff L, Walker C, Helms K, Fostel J, et al. Targeted Deletion of Nrf2 Reduces Urethane-Induced Lung Tumor Development in Mice. *PloS One* (2011) 6:e26590. doi: 10.1371/journal.pone.0026590
32. Menegon S, Columbano A, Giordano S. The Dual Roles of NRF2 in Cancer. *Trends Mol Med* (2016) 22:578–93. doi: 10.1016/j.molmed.2016.05.002
33. Zhang L, Wang N, Zhou S, Ye W, Jing G, Zhang M. Propofol Induces Proliferation and Invasion of Gallbladder Cancer Cells Through Activation of Nrf2. *J Exp Clin Cancer Res* (2012) 31:66. doi: 10.1186/1756-9966-31-66
34. Pan H, Wang H, Zhu L, Mao L, Qiao L, Su X. The Role of Nrf2 in Migration and Invasion of Human Glioma Cell U251. *World Neurosurg* (2013) 80:363–70. doi: 10.1016/j.wneu.2011.06.063
35. Shen H, Yang Y, Xia S, Rao B, Zhang J, Wang J. Blockage of Nrf2 Suppresses the Migration and Invasion of Esophageal Squamous Cell Carcinoma Cells in Hypoxic Microenvironment. *Dis Esophagus* (2014) 27:685–92. doi: 10.1111/dote.12124
36. Widenmaier SB, Snyder NA, Nguyen TB, Arduini A, Lee GY, Arruda AP, et al. NRF1 Is an ER Membrane Sensor That Is Central to Cholesterol Homeostasis. *Cell* (2017) 171:1094–109. doi: 10.1016/j.cell.2017.10.003
37. Zhu Y-P, Zheng Z, Hu S, Ru X, Fan Z, Qiu L, et al. Unification of Opposites Between Two Antioxidant Transcription Factors Nrf1 and Nrf2 in Mediating Distinct Cellular Responses to the Endoplasmic Reticulum Stressor Tunicamycin. *Antioxidants* (2019) 9:4. doi: 10.3390/antiox9010004
38. Luna L, Johnsen O, Skartlien AH, Pedeutour F, Turc-Carel C, Prydz H, et al. Molecular Cloning of a Putative Novel Human bZIP Transcription Factor on Chromosome 17q22. *Genomics* (1994) 22:553–62. doi: 10.1006/geno.1994.1428
39. Wang M, Qiu L, Ru X, Song Y, Zhang Y. Distinct Isoforms of Nrf1 Diversely Regulate Different Subsets of its Cognate Target Genes. *Sci Rep* (2019) 9:2960. doi: 10.1038/s41598-019-39536-0
40. Luna L, Skammelsrud N, Johnsen O, Abel KJ, Weber BL, Prydz H, et al. Structural Organization and Mapping of the Human TCF11 Gene. *Genomics* (1995) 27:237–44. doi: 10.1006/geno.1995.1037
41. Zhang Y, Hayes JD. The Membrane-Topogenic Vectorial Behaviour of Nrf1 Controls its Post-Translational Modification and Transactivation Activity. *Sci Rep* (2013) 3:2006. doi: 10.1038/srep02006
42. Husberg C, Murphy P, Martin E, Kolsto AB. Two Domains of the Human bZIP Transcription Factor TCF11 Are Necessary for Transactivation. *J Biol Chem* (2001) 276:17641–52. doi: 10.1074/jbc.M007951200
43. Zhang Y, Qiu L, Li S, Xiang Y, Chen J, Ren Y. The C-terminal Domain of Nrf1 Negatively Regulates the Full-Length CNC-bZIP Factor and Its Shorter Isoform LCR-F1/Nrf1beta; Both are Also Inhibited by the Small Dominant-Negative Nrf1gamma/delta Isoforms That Down-Regulate ARE-Battery Gene Expression. *PloS One* (2014) 9:e109159. doi: 10.1371/journal.pone.0109159
44. Liu P, Kerins MJ, Tian W, Neupane D, Zhang DD, Ooi A. Differential and Overlapping Targets of the Transcriptional Regulators NRF1, NRF2, and NRF3 in Human Cells. *J Biol Chem* (2019) 294:18131–49. doi: 10.1074/jbc.RA119.009591
45. Ibrahim L, Mesgarzadeh J, Xu I, Powers ET, Wiseman RL, Bollong MJ. Defining the Functional Targets of Cap'n'collar Transcription Factors NRF1,

NRF2, and NRF3. *Antioxidants (Basel)* (2020) 9:1025. doi: 10.3390/antiox9101025
46. Zhang Y, Hayes JD. Identification of Topological Determinants in the N-terminal Domain of Transcription Factor Nrf1 That Control Its Orientation in the Endoplasmic Reticulum Membrane. *Biochem J* (2010) 430:497–510. doi: 10.1042/BJ20100471
47. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, et al. SOAP2: An Improved Ultrafast Tool for Short Read Alignment. *Bioinformatics* (2009) 25:1966–7. doi: 10.1093/bioinformatics/btp336
48. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and Quantifying Mammalian Transcriptomes by RNA-Seq. *Nat Methods* (2008) 5:621–8. doi: 10.1038/nmeth.1226
49. Benjamini Y, Yekutieli D. The Control of the False Discovery Rate in Multiple Testing Under Dependency. *Ann Stat* (2001) 29:1165–88. doi: 10.1214/aos/1013699998
50. Audic S, Claverie JM. The Significance of Digital Gene Expression Profiles. *Genome Res* (1997) 7:986–95. doi: 10.1101/gr.7.10.986
51. Morton CL, Houghton PJ. Establishment of Human Tumor Xenografts in Immunodeficient Mice. *Nat Protoc* (2007) 2:247–50. doi: 10.1038/nprot.2007.25
52. Liang CC, Park AY, Guan JL. In Vitro Scratch Assay: A Convenient and Inexpensive Method for Analysis of Cell Migration *In Vitro*. *Nat Protoc* (2007) 2:329–33. doi: 10.1038/nprot.2007.30
53. Berginc K, Kristl A. Transwell-Grown HepG2 Cell Monolayers as In Vitro Permeability Model to Study Drug-Drug or Drug-Food Interactions. *J Med Food* (2011) 14:135–9. doi: 10.1089/jmf.2010.0041
54. Chen J, Wang M, Xiang Y, Ru X, Ren Y, Liu X, et al. Nrf1 Is Endowed With a Dominant Tumor-Repressing Effect Onto the Wnt/beta-Catenin-Dependent and Wnt/beta-Catenin-Independent Signaling Networks in the Human Liver Cancer. *Oxid Med Cell Longev* (2020) 2020:5138539. doi: 10.1155/2020/5138539
55. Menyhart O, Nagy A, Gyorffy B. Determining Consistent Prognostic Biomarkers of Overall Survival and Vascular Invasion in Hepatocellular Carcinoma. *R Soc Open Sci* (2018) 5:181006. doi: 10.1098/rsos.181006
56. 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* (2017) 19:649–58. doi: 10.1016/j.neo.2017.05.002
57. Anaya J. Oncolnc: Linking TCGA Survival Data to mRNAs, miRNAs, and LncRNAs. *PeerJ Comput Sci* (2016) 2:e67. doi: 10.7717/peerj-cs.67
58. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhori G, et al. A Pathology Atlas of the Human Cancer Transcriptome. *Science* (2017) 357: aan2507. doi: 10.1126/science.aan2507
59. Liang R, Liu Z, Piao X, Zuo M, Zhang J, Liu Z, et al. Research Progress on GP73 in Malignant Tumors. *Onco Targets Ther* (2018) 11:7417–21. doi: 10.2147/OTT.S181239
60. Ye JZ, Yan SM, Yuan CL, Wu HN, Zhang JY, Liu ZH, et al. GP73 Level Determines Chemotherapeutic Resistance in Human Hepatocellular Carcinoma Cells. *J Cancer* (2018) 9:415–23. doi: 10.7150/jca.19185
61. Yu M, Luo H, Fan M, Wu X, Shi B, Di S, et al. Development of GPC3-Specific Chimeric Antigen Receptor-Engineered Natural Killer Cells for the Treatment of Hepatocellular Carcinoma. *Mol Ther* (2018) 26:366–78. doi: 10.1016/j.ymthe.2017.12.012
62. Zhang Q, Han Z, Tao J, Zhao M, Zhang W, Li P, et al. An Innovative Peptide With High Affinity to GPC3 for Hepatocellular Carcinoma Diagnosis. *Biomater Sci* (2018) 7:159–67. doi: 10.1039/c8bm01016a
63. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: Protein-Protein Interaction Networks, Integrated Over the Tree of Life. *Nucleic Acids Res* (2015) 43:D447–452. doi: 10.1093/nar/gku1003
64. Zhang Y, Ren Y, Li S, Hayes JD. Transcription Factor Nrf1 is Topologically Repartitioned Across Membranes to Enable Target Gene Transactivation Through Its Acidic Glucose-Responsive Domains. *PloS One* (2014) 9: e93458. doi: 10.1371/journal.pone.0093458
65. Lehrbach NJ, Breen PC, Ruvkun G. Protein Sequence Editing of SKN-1A/Nrf1 by Peptide:N-Glycanase Controls Proteasome Gene Expression. *Cell* (2019) 177:737–750 e715. doi: 10.1016/j.cell.2019.03.035
66. Tomlin FM, Gerling-Driessen UIM, Liu YC, Flynn RA, Vangala JR, Lentz CS, et al. Inhibition of NGLY1 Inactivates the Transcription Factor Nrf1 and Potentiates Proteasome Inhibitor Cytotoxicity. *ACS Cent Sci* (2017) 3:1143–55. doi: 10.1021/acscentsci.7b00224
67. Owings KG, Lowry JB, Bi Y, Might M, Chow CY. Transcriptome and Functional Analysis in a Drosophila Model of NGLY1 Deficiency Provides Insight Into Therapeutic Approaches. *Hum Mol Genet* (2018) 27:1055–66. doi: 10.1093/hmg/ddy026
68. van Keulen BJ, Rotteveel J, Finken MJJ. Unexplained Death in Patients With NGLY1 Mutations may be Explained by Adrenal Insufficiency. *Physiol Rep* (2019) 7:e13979. doi: 10.14814/phy2.13979
69. Xiang Y, Halin J, Fan Z, Hu S, Wang M, Qiu L, et al. Topovectorial Mechanisms Control the Juxtamembrane Proteolytic Processing of Nrf1 to Remove Its N-terminal Polypeptides During Maturation of the CNC-bZIP Factor. *Toxicol Appl Pharmacol* (2018) 360:160–84. doi: 10.1016/j.taap.2018.09.039
70. Xiang Y, Wang M, Hu S, Qiu L, Yang F, Zhang Z, et al. Mechanisms Controlling the Multistage Post-Translational Processing of Endogenous Nrf1alpha/TCF11 Proteins to Yield Distinct Isoforms Within the Coupled Positive and Negative Feedback Circuits. *Toxicol Appl Pharmacol* (2018) 360:212–35. doi: 10.1016/j.taap.2018.09.036
71. Zhang Y, Li S, Xiang Y, Qiu L, Zhao H, Hayes JD. The Selective Post-Translational Processing of Transcription Factor Nrf1 Yields Distinct Isoforms That Dictate its Ability to Differentially Regulate Gene Expression. *Sci Rep* (2015) 5:12983. doi: 10.1038/srep12983
72. Itoh K, Chiba T, Takahashi S, Ishii T, Igarashi K, Katoh Y, et al. An Nrf2/small Maf Heterodimer Mediates the Induction of Phase II Detoxifying Enzyme Genes Through Antioxidant Response Elements. *Biochem Biophys Res Commun* (1997) 236:313–22. doi: 10.1006/bbrc.1997.6943
73. Rojo de la Vega M, Chapman E, Zhang DD. NRF2 and the Hallmarks of Cancer. *Cancer Cell* (2018) 34:21–43. doi: 10.1016/j.ccell.2018.03.022
74. Yamamoto M, Kensler TW, Motohashi H. The KEAP1-NRF2 System: A Thiol-Based Sensor-Effector Apparatus for Maintaining Redox Homeostasis. *Physiol Rev* (2018) 98:1169–203. doi: 10.1152/physrev.00023.2017
75. Cuadrado A, Rojo AI, Wells G, Hayes JD, Cousin SP, Rumsey WL, et al. Therapeutic Targeting of the NRF2 and KEAP1 Partnership in Chronic Diseases. Nature Reviews. *Drug Discov* (2019) 18:295–317. doi: 10.1038/s41573-018-0008-x
76. Xu C, Huang MT, Shen G, Yuan X, Lin W, Khor TO, et al. Inhibition of 7,12-Dimethylbenz(a)Anthracene-Induced Skin Tumorigenesis in C57BL/6 Mice by Sulforaphane Is Mediated by Nuclear Factor E2-Related Factor 2. *Cancer Res* (2006) 66:8293–6. doi: 10.1158/0008-5472.CAN-06-0300
77. Chan K, Lu R, Chang JC, Kan YW. NRF2, a Member of the NFE2 Family of Transcription Factors, Is Not Essential for Murine Erythropoiesis, Growth, and Development. *Proc Natl Acad Sci U S A* (1996) 93:13943–8. doi: 10.1073/pnas.93.24.13943
78. Lehrbach NJ, Ruvkun G. Endoplasmic Reticulum-Associated SKN-1A/Nrf1 Mediates a Cytoplasmic Unfolded Protein Response and Promotes Longevity. *eLife* (2019) 8. doi: 10.7554/eLife.44425
79. Schoenfelder S, Sugar R, Dimond A, Javierre BM, Armstrong H, Mifsud B, et al. Polycomb Repressive Complex PRC1 Spatially Constrains the Mouse Embryonic Stem Cell Genome. *Nat Genet* (2015) 47:1179–86. doi: 10.1038/ng.3393
80. Ohtsuji M, Katsuoka F, Kobayashi A, Aburatani H, Hayes JD, Yamamoto M. Nrf1 and Nrf2 Play Distinct Roles in Activation of Antioxidant Response Element-Dependent Genes. *J Biol Chem* (2008) 283:33554–62. doi: 10.1074/jbc.M804597200
81. Xu Z, Chen L, Leung L, Yen TS, Lee C, Chan JY. Liver-Specific Inactivation of the Nrf1 Gene in Adult Mouse Leads to Nonalcoholic Steatohepatitis and Hepatic Neoplasia. *Proc Natl Acad Sci U S A* (2005) 102:4120–5. doi: 10.1073/pnas.0500660102
82. Oh DH, Rigas D, Cho A, Chan JY. Deficiency in the Nuclear-Related Factor Erythroid 2 Transcription Factor (Nrf1) Leads to Genetic Instability. *FEBS J* (2012) 279:4121–30. doi: 10.1111/febs.12005

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 Cell and Developmental Biology

ORIGINAL RESEARCH
published: 12 January 2022
doi: 10.3389/fcell.2021.795600

A Novel Platelet-Related Gene Signature for Predicting the Prognosis of Triple-Negative Breast Cancer

Jindong Xie[1†], Yutian Zou[1†], Feng Ye[1†], Wanzhen Zhao[1,2], Xinhua Xie[1], Xueqi Ou[1], Xiaoming Xie[1] and Weidong Wei[1]**

[1]Department of Breast Oncology, State Key Laboratory of Oncology in South China, Sun Yat-sen University Cancer Center, Collaborative Innovation Center for Cancer Medicine, Guangzhou, China, [2]Department of Radiotherapy, The First Affiliated Hospital, Hengyang Medical School, University of South China, Hengyang, China

Regarded as the most invasive subtype, triple-negative breast cancer (TNBC) lacks the expression of estrogen receptors (ERs), progesterone receptors (PRs), and human epidermal growth factor receptor 2 (HER2) proteins. Platelets have recently been shown to be associated with metastasis of malignant tumors. Nevertheless, the status of platelet-related genes in TNBC and their correlation with patient prognosis remain unknown. In this study, the expression and variation levels of platelet-related genes were identified and patients with TNBC were divided into three subtypes. We collected cohorts from The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases. By applying the least absolute shrinkage and selection operator (LASSO) Cox regression method, we constructed a seven-gene signature which classified the two cohorts of patients with TNBC into low- or high-risk groups. Patients in the high-risk group were more likely to have lower survival rates than those in the low-risk group. The risk score, incorporated with the clinical features, was confirmed as an independent factor for predicting the overall survival (OS) time. Functional enrichment analyses revealed the involvement of a variety of vital biological processes and classical cancer-related pathways that could be important to the ultimate prognosis of TNBC. We then built a nomogram that performed well. Moreover, we tested the model in other cohorts and obtained positive outcomes. In conclusion, platelet-related genes were closely related to TNBC, and this novel signature could serve as a tool for the assessment of clinical prognosis.

Keywords: triple-negative breast cancer, platelet, mRNA signature, prediction, prognosis

OPEN ACCESS

Edited by:
Bing Xu,
Xiamen University, China

Reviewed by:
Federica Tomao,
European Institute of Oncology (IEO), Italy
Xiyun Deng,
Hunan Normal University, China

****Correspondence:***
Xiaoming Xie
xiexm@sysucc.org.cn
Weidong Wei
weiwd@sysucc.org.cn

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Cell and Developmental Biology

Received: *15 October 2021*
Accepted: *13 December 2021*
Published: *12 January 2022*

Citation:
Xie J, Zou Y, Ye F, Zhao W, Xie X, Ou X, Xie X and Wei W (2022) A Novel Platelet-Related Gene Signature for Predicting the Prognosis of Triple-Negative Breast Cancer. Front. Cell Dev. Biol. 9:795600. doi: 10.3389/fcell.2021.795600

INTRODUCTION

Breast cancer (BC) remains a primary disease burden for women worldwide (Britt et al., 2020). According to the expression of estrogen receptors (ERs), progesterone receptors (PRs), and human epidermal growth factor receptor 2 (HER2) proteins, BC can be divided into several subtypes, and each lead to certain therapeutic sensitivities and prognoses (Heer et al., 2020). Among these subtypes, triple-negative breast cancer (TNBC) accounts for 15–20% of all breast cancer cases (Carey et al., 2010). Fewer than 30% of metastatic TNBC patients survive for more than 5 years (Johnson et al., 2013). Due to its chemoresistance and unfavorable prognosis, treatment of TNBC is still a major challenge and considered to be a "black hole" compared to other BC subtypes (Zheng et al., 2020a). Given the limitations of TNBC treatments, there is an urgent need to explore novel targets to

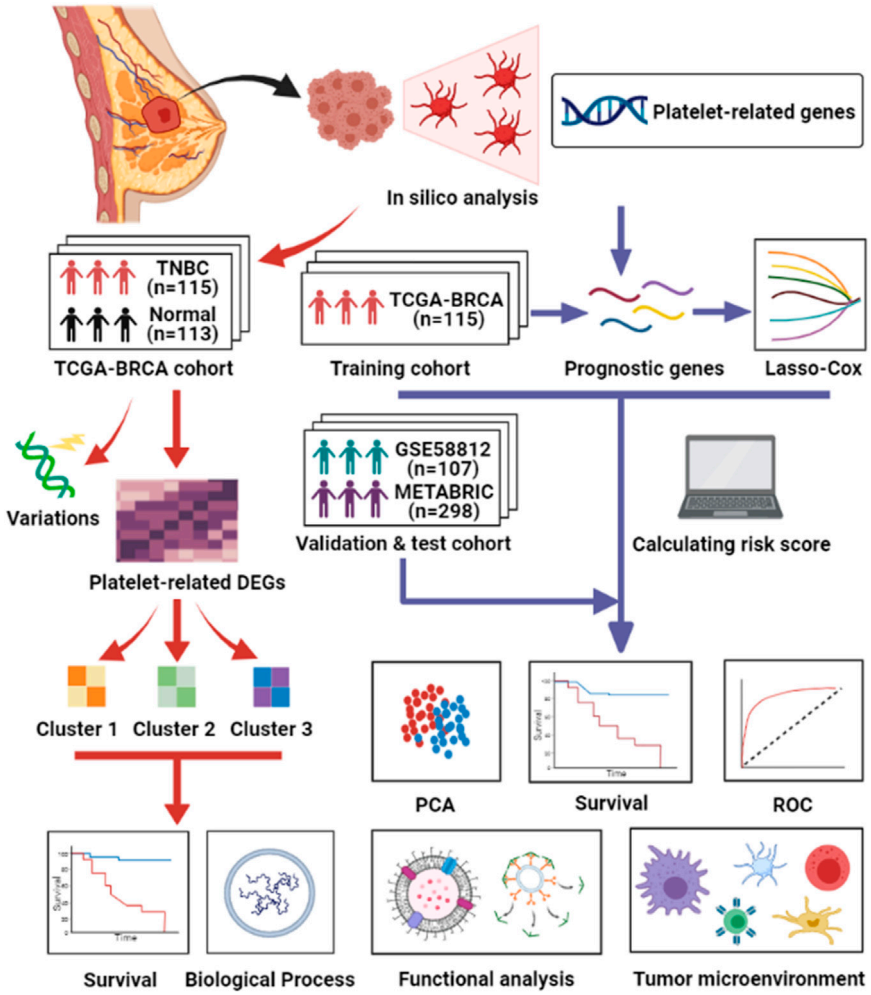


FIGURE 1 | Workflow of data analysis in our study.

ameliorate the prognosis of TNBC, and effective models are imperative to make targeted therapy more feasible.

Platelets, which regulate hemostasis and thrombosis, are one of the three main types of blood cells in the human body (Holinstat, 2017). They are created in the bone marrow and circulate through the bloodstream. When bleeding or injury occurs, they quickly reach the wound site, form a plug, and recruit more platelets simultaneously. They form a clot that coalesces with other clotting factors to help stop bleeding (van der Meijden and Heemskerk, 2019). A reduction of platelets can result in bleeding diathesis, causing spot bleeding, bruising, and purple spots. When the platelet count lowers even more, hematencephalon or pneumorrhagia can occur, which lead to death (Vinholt, 2019). In addition to playing an indispensable role in clotting and maintaining hemostasis, platelets are also directly implicated in cancer. It is well known that platelets participate substantially in the growth and metastasis of cancers. In 1872, Leopold Riess found that thrombocytosis was common in solid tumors (Riess, 1872). In many patients with cancer, platelet counts can greatly increase (Haemmerle et al., 2018). Platelets are involved in angiogenesis and tumor progression. It has been reported that thrombocytosis is directly associated with a shorter survival time in various types of cancer (Buergy et al., 2012). Platelets have an impact on the treatment efficacy of patients and take part in multiple steps of cancer metastasis. Tumor cells cannot survive unless they avoid attack by natural killer (NK) cells. Platelets encase the tumor cells in a thrombus so that they escape from NK cell monitoring, which allows them to circulate through the bloodstream (Nieswandt et al., 1999). Additionally, platelets can secrete TGF-β, which enhances metastasis and reduces the cytotoxicity of NK cells and the production of IFN-γ (Kopp et al., 2009). Platelets can also store and release growth factors such as the vascular endothelial growth factor (VEGF) and platelet-derived growth factor (PDGF), which are crucial for tumor growth and vascular stability, when stimulated by external sources (Wojtukiewicz et al., 2017). Moreover, platelets can prevent chemotherapy-induced apoptosis in cancer cells. It has been proven that thrombocytosis is closely related to poor response to chemotherapy during *in vitro* and *in vivo* experiments, which

indicates that the efficacy of chemotherapy drugs could be improved if we suppress the amount or activity of platelets (Bottsford-Miller et al., 2015).

Based on the existing findings above, we hypothesized that platelets are closely related to the proliferation and metastasis of cancers. However, there are few studies on the detailed functions of platelets in TNBC. In addition, array-based databases to recognize survival-associated genes are valuable and urgently needed to guide tailored therapies for TNBC patients (Xie et al., 2019). Therefore, we conducted a comprehensive study to examine the status of platelet-related genes in normal breast and TNBC tissues, determine their prognostic value, and determine the relationship between platelets and the tumor immune microenvironment. **Figure 1** summarizes the workflow of the data analysis. In summary, our study could contribute to discovering the heterogeneity of TNBC and offers a method to select suitable patients for immunotherapy.

MATERIALS AND METHODS

Selection of Platelet-Related Genes

The platelet-related genes list was collected from gene set enrichment analysis (GSEA) gene sets (https://www.gsea-msigdb.org/gsea/index.jsp/) by the keyword "platelet." Finally, 480 genes related to platelets were included in the analysis (**Supplementary Table S1**).

Data Collection

We acquired the raw transcriptome count data and the normalized converted RNA-sequencing (RNA-seq) profile fragments per kilobase of exon per million reads mapped (FPKM) of 115 TNBC patients and 113 normal tissues from the University of California Santa Cruz (UCSC) database. Clinical characteristics and copy number variation (CNV) information were also downloaded (https://xenabrowser.net/datapages/). Masked somatic mutation information was downloaded from the GDC Data Portal (https://portal.gdc.cancer.gov/). The log2-converted chip-seq data and clinical information of the external 107 TNBC patients' validation cohort were obtained from the Gene Expression Omnibus (GEO) database (ID: GSE58812). We also collected DNA microarray data and clinical information of 298 TNBC patients from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) dataset as a test cohort (http://molonc.bccrc.ca/aparicio-lab/research/metabric/). Another dataset, GSE25066, which was selected as a neoadjuvant therapy cohort, was also downloaded from the GEO database.

Identification of Differentially Expressed Platelet-Related Genes

Raw transcriptome count data of 115 TNBC patients and 113 compared normal samples in the TCGA–BRCA cohort were prepared to identify the differentially expressed genes (DEGs). The "edgeR" package was used subsequently to screen out DEGs with a p value < 0.05 and the absolute value of the log2 fold change (log2FC) > 1 (Robinson et al., 2010). Based on these DEGs, Kyoto Encyclopedia of Genes and Genomes (KEGG) and gene ontology (GO) analyses were performed (R package "clusterProfiler") (Yu et al., 2012).

Identification of Variant Characteristics of Platelet-Related Genes

We applied the "maftools" package to analyze masked somatic mutation information of TNBC patients (Mayakonda et al., 2018). CNV values of the platelet-related genes were screened out and those less than −0.2 were deemed as "loss" while greater than 0.2 were regarded as "gain". The varied characteristics of platelet-related genes were shown in a circus plot with the help of the "RCircos" package (Zhang et al., 2013).

Unsupervised Clustering of Platelet-Related DEGs

According to the DEGs, we used the "ConsensusClusterPlus" package to complete consensus clustering (CC) in order to identify the unidentified subtypes of TNBC (Wilkerson and Hayes, 2010). The CC parameter "maxK" was selected as "10," "clusterAlg" was selected as "pam," and "distance" was selected as "pearson." The "gsva" package was applied to perform the single-sample gene set enrichment analysis (ssGSEA), and the "limma" package was used to find out different active pathways among three clusters. The reference database was "h.hallmark.v7.4.symbols.gmt" (Hänzelmann et al., 2013; Ritchie et al., 2015).

Construction of the Platelet-Related Gene Signature

Univariate Cox regression was utilized to assess whether these genes made a difference to the survival status in both the TCGA cohort and the GSE58812 cohort. To avoid omissions, we adjusted the cutoff p value to 0.1. The least absolute shrinkage and selection operator (LASSO) Cox regression method (R package "glmnet") was further used to shrink the candidates in order to construct the most suitable signature, with the selection of "lambda. min" (Friedman et al., 2010). Ultimately, the model exported the risk score of each patient by the formula below:

$$Risk\ score = \sum_{i=1}^{7} \beta i * Ei$$

(βi represents the coefficient index, and Ei represents the gene expression level).

To make plots more intuitionistic, we adjusted the risk score by using a linear transformation. The calculated risk score subtracted the minimum and divided it by the maximum, which mapped these exponentials to the range of 0–1. The formula was as follows:

$$Ad\ just\ Risk\ score = \frac{Risk\ score - \mathbf{min}\,(Risk\ score)}{\mathbf{max}\,(Risk\ score) - \mathbf{min}\,(Risk\ score)}$$

On the grounds of the median value of the risk score, TNBC patients in each cohort were classified into low-risk and high-risk groups. We used the "stats" R package to perform principal component analysis (PCA). The prognostic difference between the two groups was investigated *via* Kaplan–Meier analysis (R package "survival"). Besides, the "survminer," "rms," and "timeROC" R packages were applied to finish receiver operating characteristic (ROC) analysis (Blanche et al., 2013).

Independent Prognostic Analysis of the Risk Score

We collected the clinical features (age, pathologic T, pathologic N, and stage) of TNBC patients in the TCGA cohort, and they were analyzed together with the risk score by means of univariate and multivariable Cox regression.

Functional Enrichment Analyses Between Risk Groups

Gene set enrichment analysis (GSEA, https://software.broadinstitute.org/gsea/index.jsp/) was used to find out the different biological functions and signaling pathways between the two groups. We set "c2.cp.kegg.v7.4.symbols.gmt" and "c5.go.bp.v7.4. symbols.gmt" as the reference database, and the cutoff criteria were set to $|NES| > 1.5$ and $Q < 0.25$. Gene set variation analysis (GSVA) was employed with the "c2.cp.reactome.v7.4.symbols.gmt" database, and their correlations with the risk score were also calculated by the means of "gsva" and "corrplot" packages (Hänzelmann et al., 2013).

The Nomogram Establishing

Age, pathologic T, pathologic N, stage, and the risk score were set together. Multivariable Cox regression and stepwise regression analyses were employed to establish a prognostic nomogram. The nomogram plot was shown by the "regplot" package. Calibration plots and decision curve analysis (DCA) were used to evaluate the efficacy of the nomogram (R package "caret" and "rmda").

Tumor Immune Microenvironment Analysis

We collected gene expression data in the three cohorts and downloaded the LM22 signatures (Newman et al., 2015). With the help of the online platform CIBERSORTx (https://cibersortx.stanford.edu/), we analyzed whether there was a relationship between the risk groups and the tumor immune microenvironment.

Statistical Analysis

All statistical analyses were presented *via* R 4.1.0 (https://www.r-project.org/). The Wilcoxon test was employed to evaluate the difference of expression levels of the TNBC samples and normal tissues, as well as low- and high-risk groups. The relationship between the risk groups of patients and response to neoadjuvant therapy was assessed by the chi-square test.

RESULTS

Identification of DEGs Between TNBC and Normal Tissue

The expression levels of 480 platelet-related genes were compared between samples from 115 patients with TNBC and 113 normal samples in the TCGA-BRCA cohort, and 177 DEGs were identified. In the TNBC group, 64 genes were downregulated, whereas 113 other genes were upregulated (**Supplementary Table S2**). The heatmap of DEGs is shown in **Figure 2A**, and the volcano plot of DEGs is presented in **Figure 2B**. We then performed functional enrichment analyses including KEGG and GO, and the results showed that DEGs were directly related to the platelets (**Figures 2C,D**), including platelet activation, degranulation, and aggregation (**Supplementary Table S3**).

Variant Landscape of Platelet-Related Genes

Variations in platelet-related genes were also evaluated in patients with TNBC from the TCGA cohort. The result showed there was approximately 92% (92/100) of TNBC patients who experienced mutations. The top 10 mutations of platelet-related genes are displayed in **Figure 2E** with TP53 as the most frequent (89%) and nine other mutations ranging in frequency from 4 to 21%. Meanwhile, CNV status analysis indicated frequent alterations in platelet-related genes. We noted that YWHAZ and ZFPM2 possessed the most significant copy number amplification, whereas F2RL2 had the most extensive CNV deletion (**Supplementary Figure S1**). The location, expression, CNV values, and correlation of each platelet-related gene are shown in **Figure 2F**.

TNBC Platelet-Related Subtypes Based on the DEGs

To explore unidentified subtypes of TNBC, we used the expression of platelet-related DEGs to perform CC analysis of the TCGA cohort. We found that when $k = 3$, the differences among subgroups were the most obvious, which indicated that 115 patients with TNBC could be classified into three clusters (**Figures 3A,B**). It is worth noting that there were obvious differences between the overall survival (OS) time and the three clusters ($p = 0.043$, **Figure 3C**). Cluster three was associated with a favorable prognosis, while cluster two was associated with a poor prognosis, and cluster one was between them. However, we found no substantial differences in the clinical features among the three clusters. We then performed ssGSEA and found that different clusters were enriched in certain pathways. For example, the worst-prognosis cluster two was rich in the KRAS signal pathway, angiogenesis, and coagulation, while it was downregulated in DNA repair and the G2M checkpoint compared to the best-prognosis cluster three (**Figure 3D**).

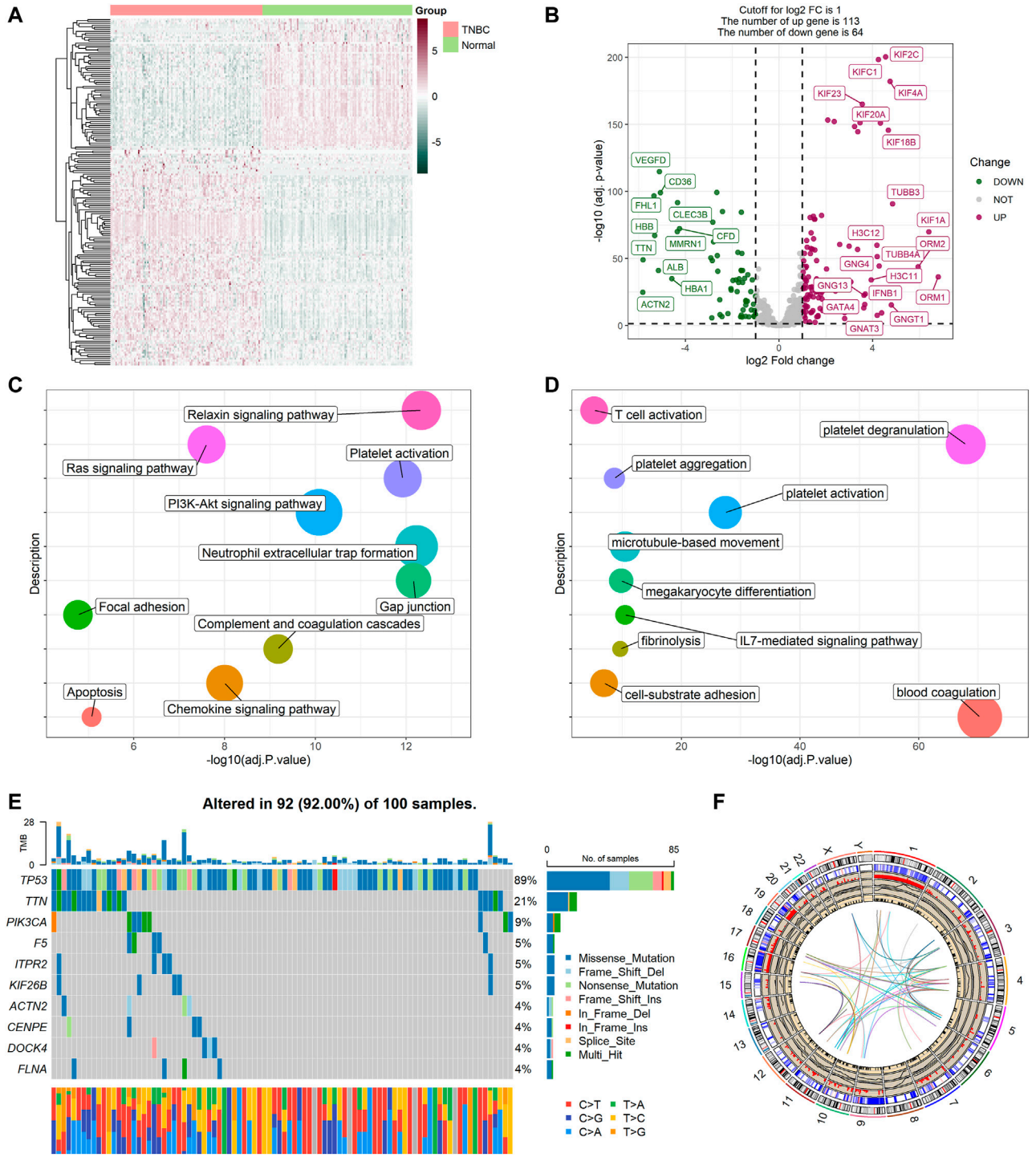


FIGURE 2 | Expressions of the 177 platelet-related DEGs and the functional enrichment analyses. **(A)** Heatmap (atrovirens: low expression level; brick red: high expression level) of the platelet-related DEGs between TNBC (brilliant red) and the normal tissues (brilliant green). **(B)** Volcano plot of the platelet-related DEGs (green: downregulated DEGs; red: upregulated DEGs; gray: unchanged genes). Points with labels are obvious DEGs with p value < 0.001 and |log2FC| > 3.5. **(C)** Bubble plot for KEGG enrichment (different colors represent different descriptions, and sizes reflect the enrichment numbers of genes). **(D)** Bubble plot for GO enrichment (biological processes). **(E)** The variation of platelet-related genes in the training cohort. **(F)** The location, expression, CNV values, and correlation of platelet-related genes in TCGA cohort.

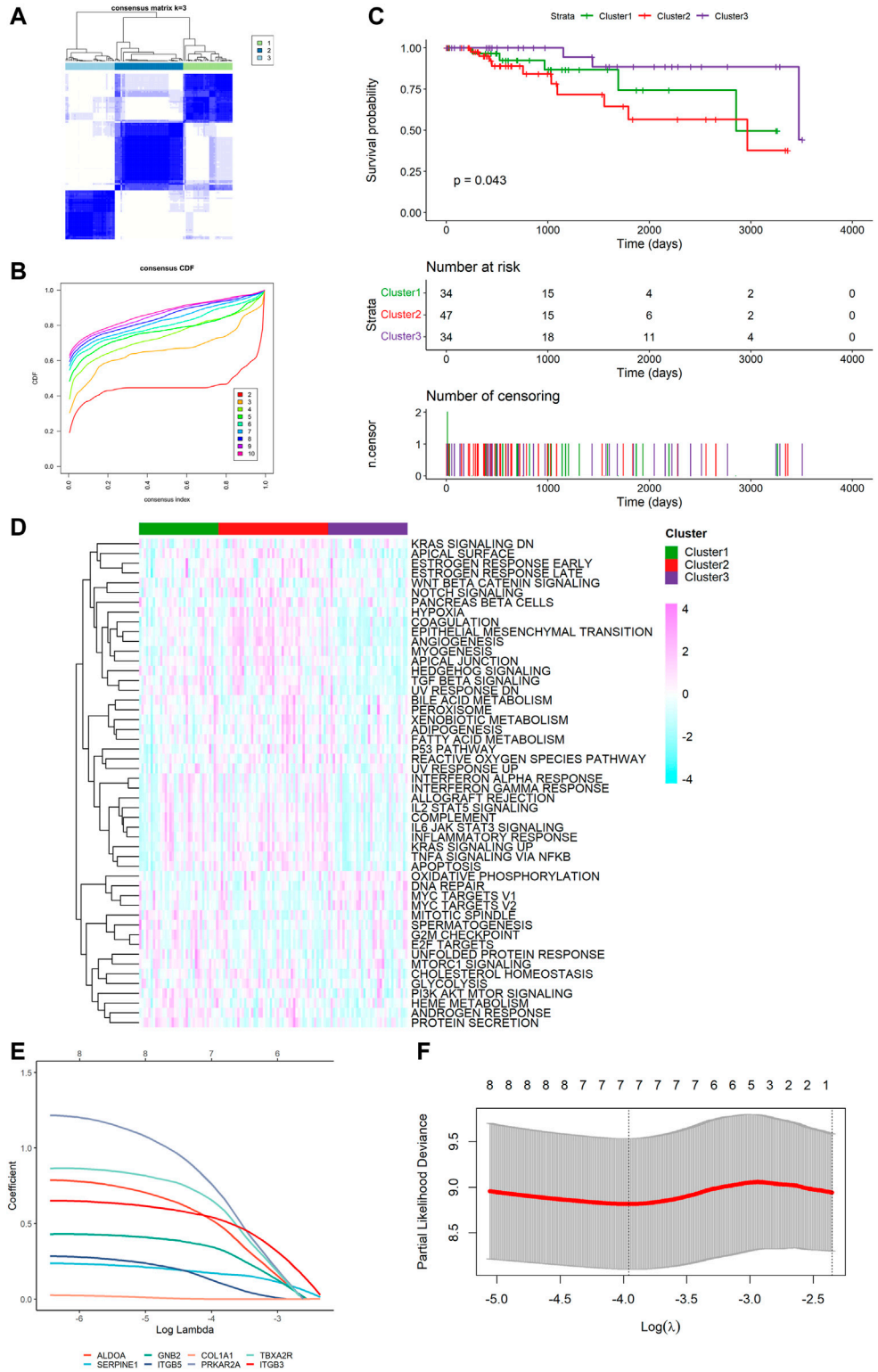


FIGURE 3 | Tumor classification based on the platelet-related DEGs and construction of risk signature in TCGA cohort. **(A)** 115 TNBC patients were divided into three subgroups (k = 3). **(B)** Consensus Cumulative Distribution Function (CDF) Plot under k = 2–10. **(C)** Kaplan–Meier OS curves for the three clusters. **(D)** Heatmap (light blue: low expression level; pink: high expression level) of the hallmark gene set among three clusters (green: cluster 1, red: cluster 2, purple: cluster 3). **(E)** LASSO Cox regression of the 7 model genes. **(F)** Cross-validation for the LASSO Cox regression.

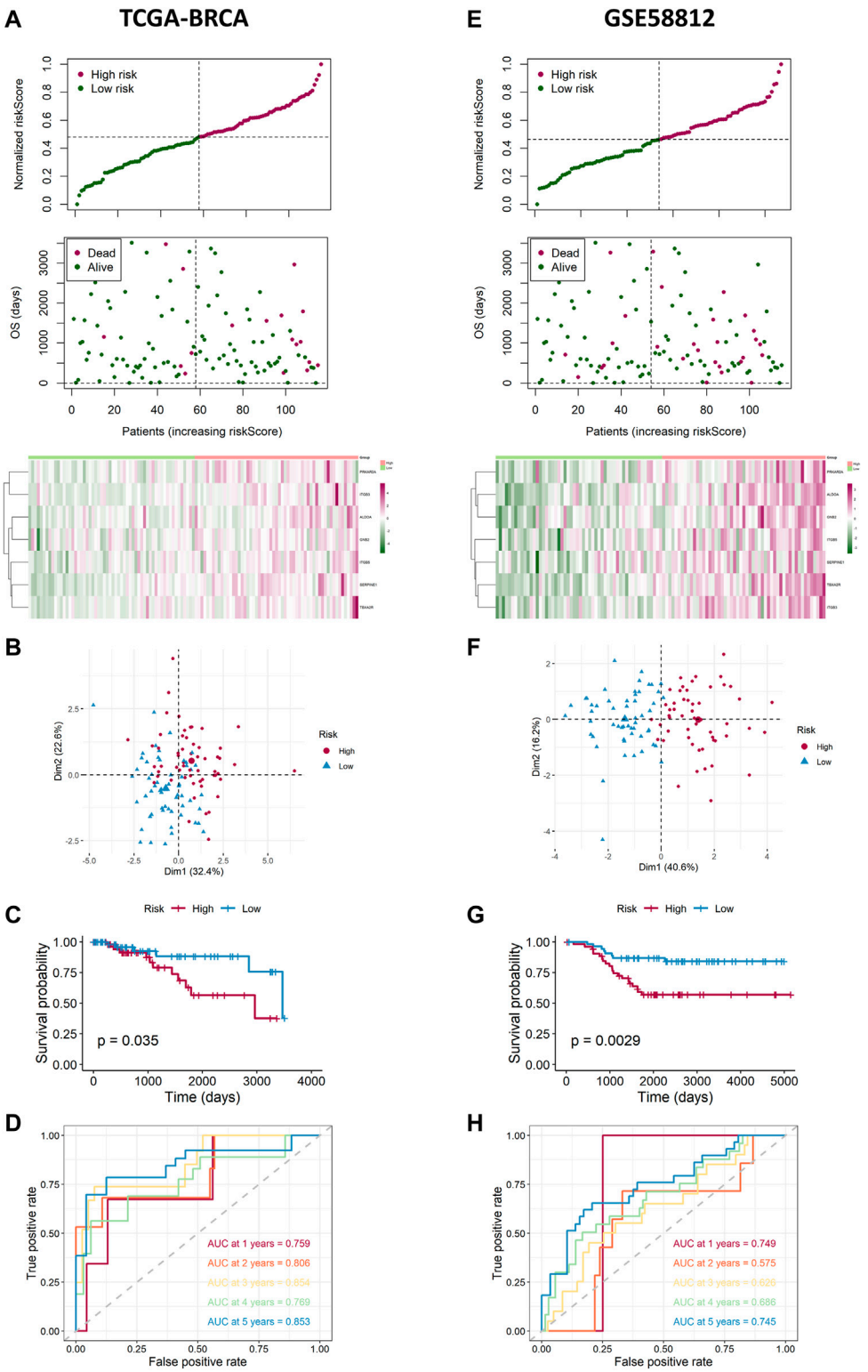


FIGURE 4 | Landscape of adjusted risk score, heatmaps, PCA, Kaplan–Meier analysis, and ROC analysis of the risk signature in the internal set and the external GEO (GSE58812) set. **(A)** Distribution of the adjusted risk score and heatmap in the training TCGA set. **(B)** PCA plot based on the risk groups in the training TCGA set. **(C)** Kaplan–Meier survival analysis of the patients in the training TCGA set. **(D)** ROC curve analysis according to the 1-, 2-, 3-, 4-, and 5-year survival of the AUC value in the training TCGA set. **(E)** Distribution of the adjusted risk score and heatmap in the validation GEO (GSE58812) set. **(F)** PCA plot based on the risk score in the validation GEO (GSE58812) set. **(G)** Kaplan–Meier survival analysis of the patients in the validation GEO (GSE58812) set. **(H)** ROC curve analysis according to the 1-, 2-, 3-, 4-, and 5-year survival of the AUC value in the validation GEO (GSE58812) set.

Construction of a Prognostic Gene Signature

Survival information was collected and matched with TNBC patients. Univariate Cox regression analysis was performed separately for general filtration. In TCGA, 46 genes met the cutoff of $p < 0.1$ as did 90 genes in GSE58812. The intersection of the two outputs had nine genes (*ALDOA*, *SERPINE1*, *COL1A2*, *GNB2*, *ITGB5*, *COL1A1*, *PRKAR2A*, *TBXA2R*, and *ITGB3*). All were risk factors with hazard ratios (HRs) > 1. By performing the LASSO Cox regression analysis, we constructed a seven-gene signature with a minimum value of lambda (λ) (**Figures 3E,F**). We investigated the correlation of each model gene (**Supplementary Figure S2**), using Kaplan–Meier (K-M) analysis to determine their respective influences on the OS time (**Supplementary Figure S3**) and the Wilcoxon test to examine the model genes' expression levels between the TNBC tissues and normal samples (**Supplementary Figure S4**). The results showed that five genes (*ALDOA*, *SERPINE1*, *GNB2*, *ITGB5*, and *ITGB3*) had a significant influence on the OS time ($p < 0.05$), and the expression of each gene was significantly different ($p < 0.05$). The risk score was calculated as follows: risk score = (0.5144062 * ALDOA exp.) + (0.1688905 * SERPINE1 exp.) + (0.3378084 * GNB2 exp.) + (0.1297068 * ITGB5 exp.) + (0.7508113 * PRKAR2A exp.) + (0.6483026 * TBXA2R exp.) + (0.5324253 * ITGB3 exp.).

Internal Training and External Validation of the Risk Signature

Based on the median value calculated by the risk score formula above, we divided the 115 patients with TNBC into low- and high-risk groups equally within the TCGA cohort, which we had selected as a training dataset. We found that patients in the high-risk group were more likely to survive as the risk score increased (**Figure 4A**). Principal component analysis (PCA) showed that this classification performed well (**Figure 4B**). K-M analysis showed that patients in the low-risk group were more likely to have lower death rates ($p < 0.05$, **Figure 4C**). In addition, we applied time-dependent ROC analysis to evaluate the efficacy of the prognostic model. The area under the ROC curve (AUC) was 0.759 for 1 year, 0.806 for 2-year, 0.854 for 3-year, 0.769 for 4-year, and 0.853 for 5-year survival (**Figure 4D**). Consequently, we used GSE58812 as the validation cohort. A higher risk score resulted in shorter survival time (**Figure 4E**). The two groups were separated using PCA (**Figure 4F**). K-M analysis also indicated that patients in the low-risk group had longer survival times ($p < 0.05$, **Figure 4G**). Moreover, ROC curve analysis of GSE58812 indicated that the established prognostic model had excellent predictive efficacy (AUC = 0.749 for 1-year, 0.575 for 2-year, 0.626 for 3-year, 0.686 for 4-year, and 0.745 for 5-year survival) (**Figure 4H**).

Independent Prognostic Value of the Risk Signature

We performed univariate and multivariable Cox regression analyses to explore whether the risk score was an independent prognostic factor. The univariate Cox regression analysis showed that, compared with other features, the risk score was regarded as a risk factor (HR = 5.287, 95% CI: 2.465–11.337, and $p < 0.05$, **Figure 5A**), predicting poor survival in TCGA cohort. The multivariate analysis further confirmed, after removing confounding factors, the risk score was still an independent risk factor (HR = 5.796, 95% CI: 2.550–13.175, and $p < 0.05$, **Figure 5B**) for TNBC patients in TCGA cohort.

Functional Enrichment Analyses Based on the Risk Signature

To further detect the distinction in the signaling pathways and biological processes between the subgroups classified by the risk signature, we applied GSEA to analyze with the criteria of $p < 0.05$ and $Q < 0.25$. The results are presented in **Supplementary Table S4**. We observed that the high-risk group was active in the excitation of the extracellular matrix (ECM) receptor interaction, the PDGF receptor signaling pathway, the VEGF signaling pathway, and the transforming growth factor-β (TGF-β) signaling pathway. As for biological processes, the low-risk group was inactive in blood vessel endothelial cell migration, focal adhesion, vascular smooth muscle contraction, regulation of blood pressure, vascular permeability, and vasculature development (**Figures 5C,D**). Moreover, we also performed an analysis to determine whether there was a relationship between the risk score and pathways derived from the reactome database. This showed that most of the patients in the high-risk group were strongly correlated with platelet-related biological processes, such as platelet activation and platelet aggregation, while negatively correlated with downstream B-cell receptor events and mitochondrial translation (**Figure 5E**). As shown in **Figure 5F**, most patients in cluster three were classified into the low-risk group, which was related to better survival outcomes. Collectively, these data suggested that the risk score had a strong correlation with the function and processes of platelets and could effectively predict prognosis.

Establishment and Assessment of the Nomogram Model

We used multivariable Cox and stepwise regression analyses to establish a nomogram model in TCGA cohort to estimate the probability of two-, three-, and 5-year OS. The stage and risk scores were selected for the model (**Figure 6A**). The C-index value of the model was 0.916 (95% CI: 0.858–0.973). Calibration curves were used to assess the consistency between the predicted and actual survival rates. The accuracy of this model in predicting the two-, three-, and 5-year survival rates was favorable (**Figures 6B–D**). Moreover, we performed DCA to confirm a range of threshold probabilities for the model and found that the nomogram model was apparently better than any other predictor applied in this study (**Figure 6E**).

Comparison of the Tumor Immune Microenvironment Between Risk Groups

We then proceeded with analysis to identify whether there was a difference in the tumor immune microenvironment between the two risk groups in TCGA and GSE58812 cohorts. We used

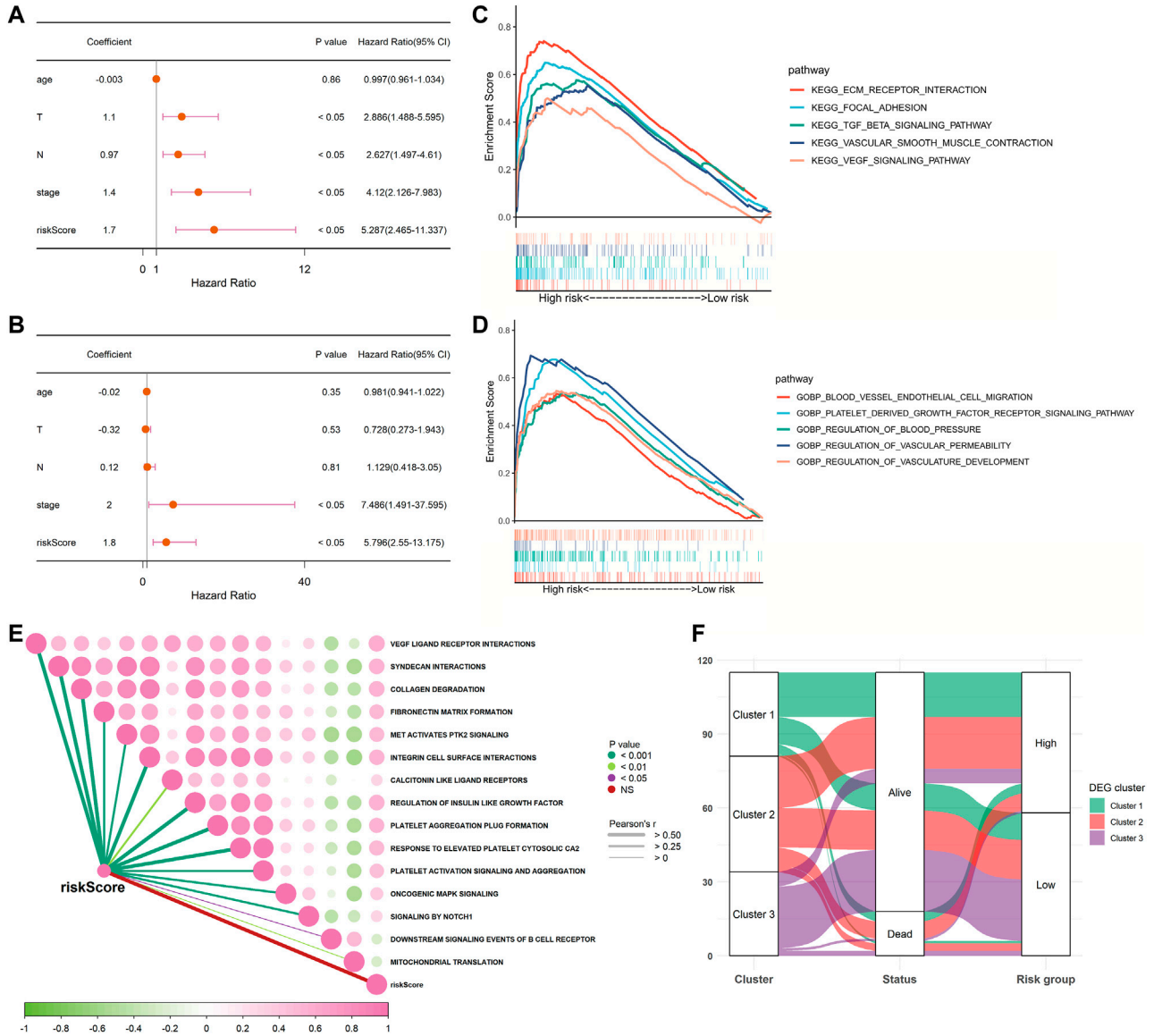


FIGURE 5 | Univariate and multivariate Cox regression analyses for the risk score and GSEA between low- and high-risk groups in TCGA cohort. **(A)** Univariate analysis for TCGA cohort. **(B)** Multivariate analysis for TCGA cohort. **(C,D)** Involved signaling pathways in two risk groups in TCGA cohort. **(E)** Correlation analysis of the risk score and pathways from the reactome gene set. **(F)** Alluvial diagram shows the changes of clusters, vital status, and risk groups in TCGA cohort.

CIBERSORTx to calculate the enrichment scores of 22 immune-related cells in each sample. In TCGA cohort, the high-risk group frequently had lower levels of immune cell infiltration, especially $CD8^+$ T cells, follicular helper T cells, active NK cells, and dendritic cells. However, the expression of macrophage 0 (M0) infiltration was significantly upregulated in the high-risk group (**Figure 7A**). A similar immune status was found in the GSE58812 cohort; the high-risk group generally had lower levels of infiltration of immune cells, especially memory B cells, memory $CD4^+$ T cells, macrophage 1 (M1) cells. However, the levels of infiltration of $CD8^+$ T cells, regulatory T cells (Tregs), M0 cells, and macrophage 2 (M2) cells were higher in the high-risk group (**Figure 7B**).

Extra Test of the Risk Signature

To ensure that the established risk signature was widely applicable, we collected 298 patients with TNBC from the METABRIC dataset, which was used as a test cohort. After calculating the risk scores, the patients were divided into two risk groups. As the risk score increased, patients were more likely to die (**Figure 8A**). PCA indicated that the classification was distinct (**Figure 8B**). The K-M analysis showed that patients in the low-risk group lived longer, with a nearly significant p value (p = 0.062, **Figure 8C**). ROC curve analysis of the METABRIC cohort revealed that the risk signature was suitable for predicting the prognosis (AUC = 0.639 for 1-year, 0.562 for 2-year, 0.563 for 3-year, 0.569 for 4-year, and 0.596 for 5-year survival) (**Figure 8D**). With respect to the tumor

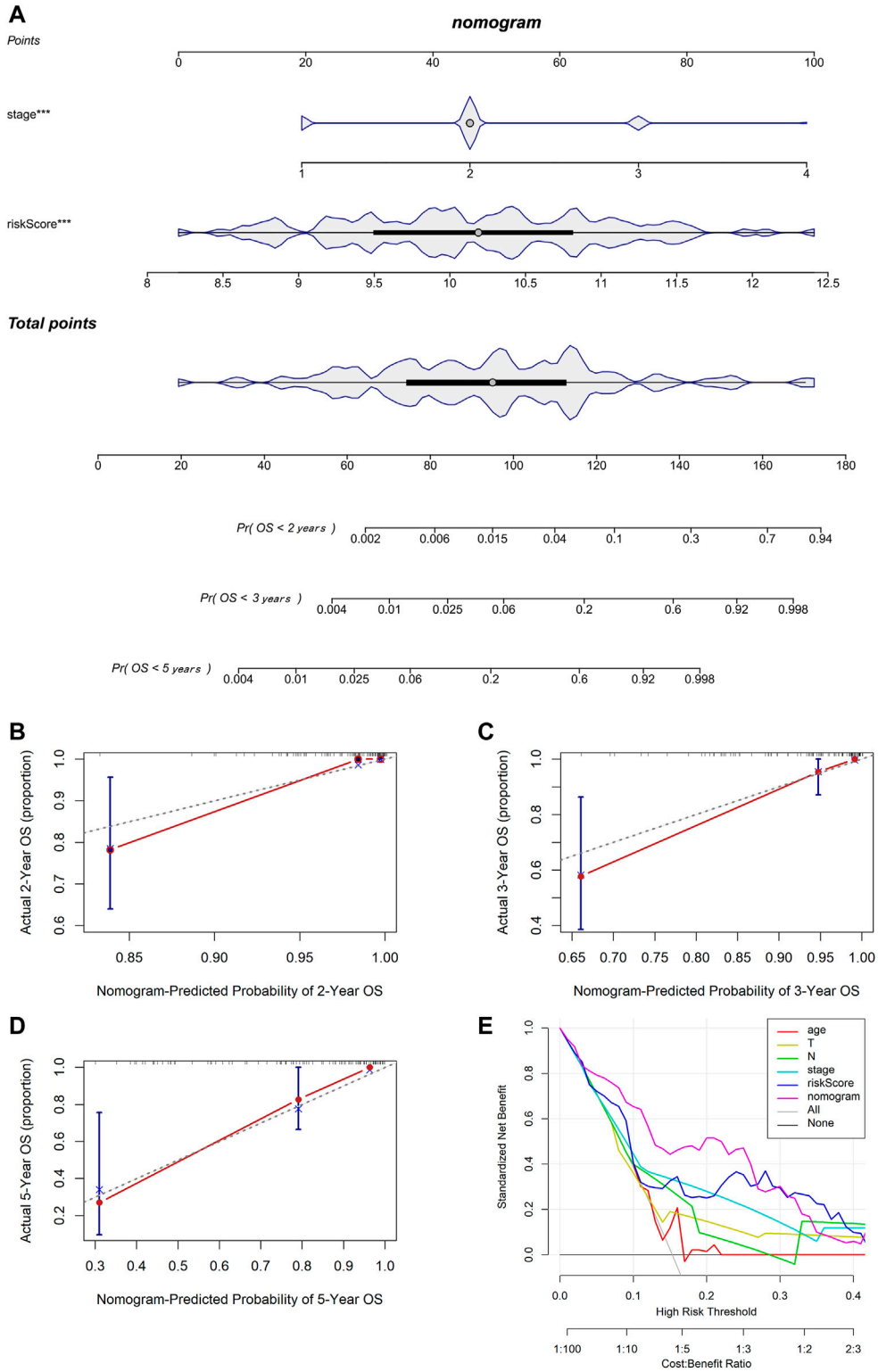


FIGURE 6 | Establishment and assessment of the nomogram based on the 7-gene signature. **(A)** Nomogram for predicting the proportion of patients with 2-, 3-, or 5-year OS (*** means $p < 0.001$). **(B–D)** Calibration plots of the nomogram-predicted probability of 2-, 3-, and 5-year survival in TCGA cohort. **(E)** DCA of the nomogram predicting 2-, 3-, and 5-year OS comparing age, pathologic T, pathologic N, stage, and risk score.

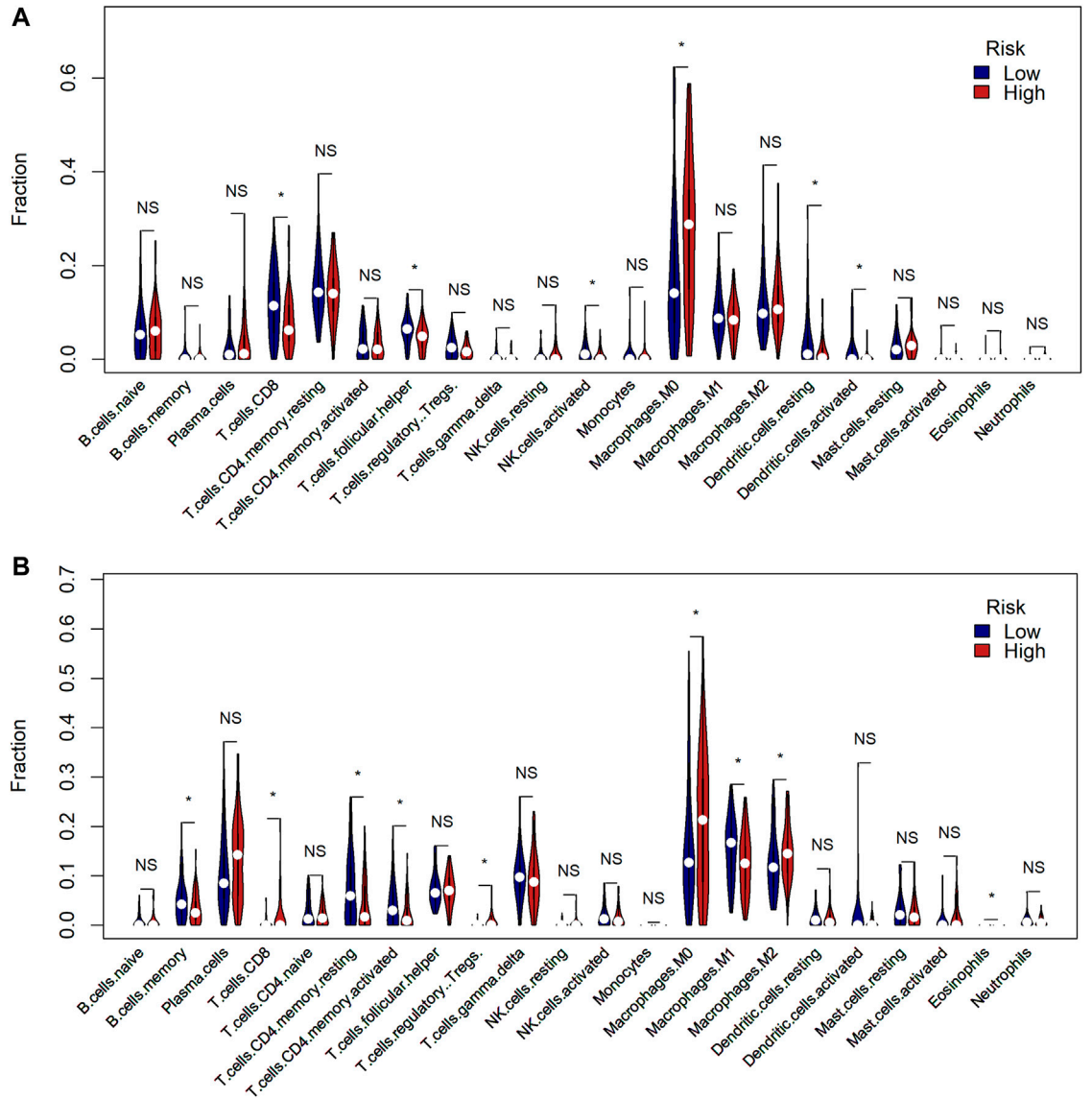


FIGURE 7 | Comparison of the tumor immune microenvironment between risk groups. **(A)** Comparison of the enrichment scores of 22 types of immune cells between low- (blue) and high-risk (red) groups in TCGA cohort (* means $p < 0.05$, NS means no significance). **(B)** Comparison of the enrichment scores of 22 types of immune cells between low- (blue) and high-risk (red) groups in the GEO cohort.

immune microenvironment, the result was the most similar to the GSE58812 cohort, in which the high-risk group had a lower level of infiltration of active memory T cells and M1 cells, while it had higher levels of infiltration of M0 and M2 cells (**Figure 8E**).

In addition, we applied the risk signature to GSE25066, which was the neoadjuvant therapy cohort. We sought to identify whether the risk signature was valuable for predicting the curative effect of neoadjuvant therapy in patients with TNBC. A total of 170 TNBC patients were classified into two groups, pathologic complete response (pCR) and non-complete response (nCR), in accordance with the response to neoadjuvant therapy. By utilizing the ROC analysis, we found that the risk scores of 113 samples in the nCR group were higher than the 57 sampled in the pCR group, and the AUC was 0.620, which was noteworthy (**Figure 8F**). In addition, we applied the chi-square test and found that there were 36 patients with pCR in the low-risk group while only 21 in the high-risk group ($p = 0.015$). The proportion of pCR in the low-risk group was much higher, indicating that these patients could be sensitive to neoadjuvant therapy (**Supplementary Figure S5**).

DISCUSSION

Regarded as the most invasive breast cancer subtype, TNBC lacks effective therapeutic targets and accurate efficacy prediction models. In this study, we first applied bioinformatics methods to explore the relationship between platelets and TNBC. We

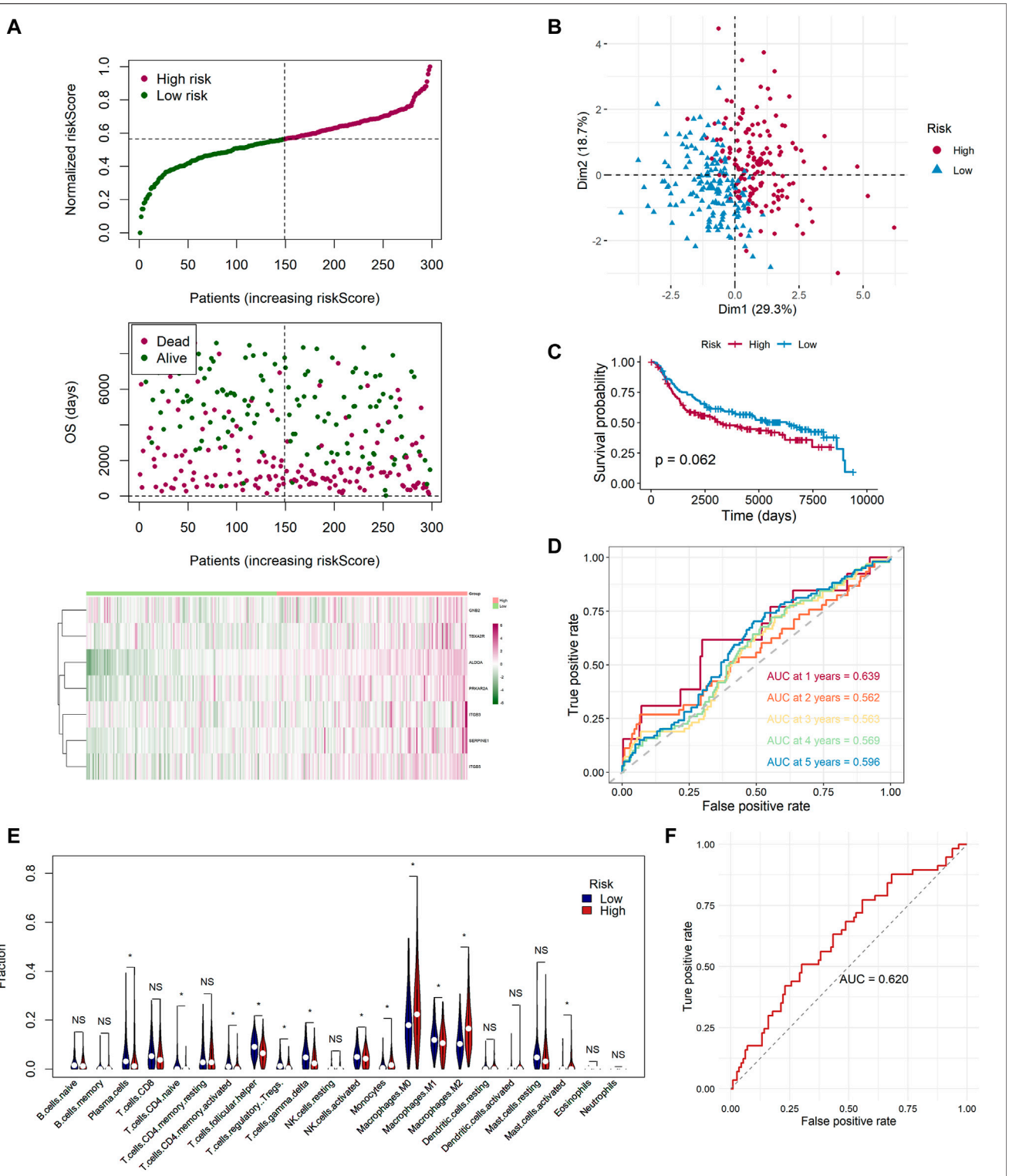


FIGURE 8 | Extra test of the risk signature in the METABRIC cohort and a GEO neoadjuvant therapy cohort (GSE25066). **(A)** Distribution of the adjusted risk score and heatmap in the METABRIC cohort. **(B)** PCA plot based on the risk score in the METABRIC cohort. **(C)** Kaplan–Meier survival analysis of the low- and high-risk group patients in the METABRIC cohort. **(D)** ROC curve analysis according to the 1-, 2-, 3-, 4-, and 5-year survival of the AUC value in the METABRIC cohort. **(E)** Comparison of the enrichment scores of 22 types of immune cells between low- (blue) and high-risk (red) groups in the METABRIC cohort. **(F)** ROC curve analysis for predicting the response of neoadjuvant therapy in the GEO neoadjuvant therapy cohort (GSE25066).

constructed a seven-gene risk signature in the TCGA cohort *via* univariate Cox analysis and LASSO Cox regression analysis, and the model was shown to perform well in the external validation dataset GSE58812. Some classical signaling pathways and platelet-related biological processes were found to be active in high-risk groups. We built a nomogram including clinical characteristics and risk scores, and the results showed that it performed well. The differences in the tumor immune microenvironment between the two risk groups were compared, and universally upregulated levels of M0 cells and downregulated levels of M1 and M2 cells were found in the high-risk group compared with the low-risk group. Finally, we added a test set downloaded from the METABRIC to ensure that the established risk signature was widely applicable, and we obtained a good outcome. Moreover, we applied the model to a neoadjuvant therapy cohort and found that the model could be useful for predicting the efficacy of neoadjuvant therapy in patients with TNBC.

Platelets are one of the three main types of blood cells in the human body and act as a double-edged sword. On the one hand, platelets are involved in the angiogenesis and metastasis of tumors (Menter et al., 2017). On the other hand, inhibition of platelet activity and quantity could be a novel therapeutic target (Yeung et al., 2018). Our study built a signature covering seven platelet-related genes (*ALDOA*, *SERPINE1*, *GNB2*, *ITGB5*, *PRKAR2A*, *TBXA2R*, and *ITGB3*) and found that it had the ability to predict OS in TNBC patients. Aldolase A (ALDOA) is an aldolase isozyme (ALDOA, ALDOB, and ALDOC). ALDOA participates in many biological processes and cellular functions, including cell morphology, motor regulation, muscle maintenance, actin filament constitution, and ATP biosynthesis (Carr and Knull, 1993; Kusakabe et al., 1997; Kajita et al., 2001). It is worth noting that ALDOA is highly expressed in various cancers, such as colorectal cancer, hepatocellular carcinomas, and pancreatic cancer (Peng et al., 2012; Ji et al., 2016). In addition, a deficiency of ALDOA is related to hemolytic anemia (Yao et al., 2004). In our study, ALDOA was highly expressed in TNBC samples, and a high expression of ALDOA was not effective for prolonging survival time, which was in agreement with previous studies. SERPINE1, also called plasminogen activator inhibitor 1 (PAI1), mediates the inhibition of fibrin degradation (Placencio and DeClerck, 2015). High concentrations of SERPINE1 are associated with thrombosis (Corduan et al., 2015). SERPINE1 is overexpressed in numerous cancers, especially breast cancer. Previous studies confirmed that upregulation of SERPINE1 predicted worse overall survival, greater possibility of metastasis, and poorer responses to chemotherapy, which was consistent with our study results (Duffy et al., 2014). G Protein Subunit-β2 (GNB2) is a protein-coding gene that is responsible for the formation of beta subunits (Blatt et al., 1988). GNB2 has been reported to be frequently mutated and upregulated in many hematological neoplasms, and a lower expression of GNB2 could reduce the proliferation potential of tumor cells (Kotani et al., 2019). Our study confirmed that GNB2 was a risk gene. Integrin-β5 (ITGB5) belongs to the integrin family, which regulates many biological functions in tumors such as proliferation, adhesion, migration, and invasion (Lin et al., 2018). The role of ITGB5 in angiogenesis has also been demonstrated (Su et al., 2007). Although GNB2 was a risk factor for survival in our study, interestingly, the expression of GNB2 was lower in TNBC samples than in normal tissues. Given the sparse data from TNBC and the contradictory results in other tumors, further studies are needed. PRKAR2A codes for protein kinase A. Previous studies have shown that PRKAR2A deficiency predisposes patients to hematopoietic malignancies (Saloustros et al., 2015). In addition, PRKAR2A has been found to regulate the response of cancer cells to chemotherapy, which might be the reason why our established model helped predict the curative effect of neoadjuvant therapy in TNBC patients (Zynda et al., 2014). The thromboxane A2 receptor (TBXA2R) is a specific coordinator of the T prostanoid receptor (TPR), which is associated with the platelet activity. The variants of the TBXA2R gene cause bleeding and increased metastasis in multiple cancers (Mundell and Mumford, 2018; Pulley et al., 2018). Our study also confirmed that TBXA2R was highly expressed in TNBC samples and was a risk factor for survival. Integrin-β3 (ITGB3), also named CD61 or GP3A, is a member of the integrin family and has been widely studied by scientists. ITGB3 acts as a receptor and participates in forming the tumor stromal and immune microenvironment, as well as maintaining tumor stemness (Zhu et al., 2019). In addition, ITGB3 can mediate extracellular vesicles to facilitate intercellular communication in BC cells (Fuentes et al., 2020). The expression level of ITGB3 was lower in the TNBC samples according to our study, which was similar to that of ITGB5.

Tumor cells can survive because the tumor microenvironment provides a haven for them to escape immune surveillance and drug interference (Zou et al., 2019). High infiltration levels of immune cells can enhance the efficacy of neoadjuvant therapy for BC patients, and the levels of antitumor-infiltrating immune cells in the high-risk group would be low, which indicates holistic damage of immune functions (Zou et al., 2020; Zou et al., 2021). Surprisingly, among the three cohorts, the expressions of M0 and M2 cells were upregulated, while the expression of M1 cells was downregulated in the high-risk group. M0 cells are resting macrophages that can be polarized into two different phenotypes, M1 and M2. Both M1 and M2 cells are closely associated with inflammatory responses; M1 cells are mainly involved in pro-inflammatory responses, while M2 cells mainly participate in anti-inflammatory responses (Mehla and Singh, 2019). Theoretically, the increase in platelets is correlated with anti-inflammation, which was similar to our results. Previous studies have shown that the abundance of M1 cells represents a better outcome, while the enrichment of M2 cells indicates a poorer prognosis in the TNBC microenvironment (Zheng et al., 2020b). We hypothesized that the phenotypic shift of the M2 subtype toward the M1 subtype might be a strategy to overcome the early phases of inflammation and immunotherapy.

The established seven-gene signature and nomogram reflected excellent performance in internal and external cohorts; however, it had the following shortcomings. First, the seven genes mentioned above were all risk factors for survival; the interaction of each other during platelet activation needs further investigation, and the roles of some of the candidates, such as PRKAR2A, that occur in TNBC

have not been revealed, which might be an important point for further research. Second, the signature lacked verification of large-scale prospective trials. Third, TCGA cohort was composed of gene sequencing data, while the GSE58812 and METABRIC cohorts were gene chip data, indicating that the results originating from external cohorts might not fully reveal actual prognostic efficacy in TNBC. Ultimately, the detailed mechanisms have not been explored at the cellular and molecular levels.

In summary, we presented a novel platelet-related gene signature as a practical tool for patients with TNBC, which can offer an independent value in the assessment of clinical outcomes.

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**.

AUTHOR CONTRIBUTIONS

WW and XX designed the study. JX, YZ, and FY collected the data. JX, YZ, FY, WZ, XX, and XO analyzed the data. JX and YZ wrote the original manuscript. XX and WW edited the manuscript. All authors agreed to be accountable for the content of this work.

FUNDING

This research was funded by the National Natural Science Foundation of China (81872152, XX).

ACKNOWLEDGMENTS

We thank all the authors who contribute their valuable methods and data and make them public.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2021.795600/full#supplementary-material

Supplementary Figure S1 | CNV status analysis of platelet-related genes in TCGA cohort (Top 20 mutation frequency).

Supplementary Figure S2 | A correlogram of each model genes. Pearson's correlation coefficients (r) for all model genes are given in the plot. The areas of the sectors are related to the r. Pink indicates a positive correlation and green indicates a negative correlation.

Supplementary Figure S3 | Kaplan-Meier survival analysis of each model genes in TCGA cohort (blue: low-expression group; red: high-expression group).

Supplementary Figure S4 | Wilcoxon test of expression levels of each model genes between TNBC tissues (red) and normal samples (green) (* means $P < 0.05$).

Supplementary Figure S5 | The correlation of risk groups with the neoadjuvant therapy efficacy in the GEO neoadjuvant therapy cohort (GSE25066, red: nCR; blue: pCR).

Supplementary Table S1 | The information of 480 genes related to platelets.

Supplementary Table S2 | 177 platelet-related DEGs between TNBC and normal tissues.

Supplementary Table S3 | Functional enrichment analyses based on the DEGs between TNBC and normal tissues. Table S4. GSEA based on the subgroups categorized by the risk signature.

Supplementary Table S4 | GSEA based on the subgroups categorized by the risk signature.

REFERENCES

Blanche, P., Dartigues, J.-F., and Jacqmin-Gadda, H. (2013). Estimating and Comparing Time-Dependent Areas under Receiver Operating Characteristic Curves for Censored Event Times with Competing Risks. *Statist. Med.* 32 (30), 5381–5397. doi:10.1002/sim.5958

Blatt, C., Eversole-Cire, P., Cohn, V. H., Zollman, S., Fournier, R. E., Mohandas, L. T., et al. (1988). Chromosomal Localization of Genes Encoding Guanine Nucleotide-Binding Protein Subunits in Mouse and Human. *Proc. Natl. Acad. Sci.* 85 (20), 7642–7646. doi:10.1073/pnas.85.20.7642

Bottsford-Miller, J., Choi, H.-J., Dalton, H. J., Stone, R. L., Cho, M. S., Haemmerle, M., et al. (2015). Differential Platelet Levels Affect Response to Taxane-Based Therapy in Ovarian Cancer. *Clin. Cancer Res.* 21 (3), 602–610. doi:10.1158/1078-0432.ccr-14-0870

Britt, K. L., Cuzick, J., and Phillips, K.-A. (2020). Key Steps for Effective Breast Cancer Prevention. *Nat. Rev. Cancer* 20 (8), 417–436. doi:10.1038/s41568-020-0266-x

Buergy, D., Wenz, F., Groden, C., and Brockmann, M. A. (2012). Tumor-Platelet Interaction in Solid Tumors. *Int. J. Cancer* 130 (12), 2747–2760. doi:10.1002/ijc.27441

Carey, L., Winer, E., Viale, G., Cameron, D., and Gianni, L. (2010). Triple-Negative Breast Cancer: Disease Entity or Title of Convenience? *Nat. Rev. Clin. Oncol.* 7 (12), 683–692. doi:10.1038/nrclinonc.2010.154

Carr, D., and Knull, H. (1993). Aldolase-Tubulin Interactions: Removal of Tubulin C Terminals Impairs Interactions. *Biochem. Biophysical Res. Commun.* 195 (1), 289–293. doi:10.1006/bbrc.1993.2043

Corduan, A., Plé, H., Laffont, B., Wallon, T., Plante, I., Landry, P., et al. (2015). Dissociation of SERPINE1 mRNA from the Translational Repressor Proteins Ago2 and TIA-1 upon Platelet Activation. *Thromb. Haemost.* 113 (5), 1046–1059. doi:10.1160/TH14-07-0622

Duffy, M. J., McGowan, P. M., Harbeck, N., Thomssen, C., and Schmitt, M. (2014). uPA and PAI-1 as Biomarkers in Breast Cancer: Validated for Clinical Use in Level-Of-Evidence-1 Studies. *Breast Cancer Res.* 16 (4), 428. doi:10.1186/s13058-014-0428-4

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* 33 (1), 1–22. doi:10.18637/jss.v033.i01

Fuentes, P., Sesé, M., Guijarro, P. J., Emperador, M., Sánchez-Redondo, S., Peinado, H., et al. (2020). ITGB3-Mediated Uptake of Small Extracellular Vesicles Facilitates Intercellular Communication in Breast Cancer Cells. *Nat. Commun.* 11 (1), 4261. doi:10.1038/s41467-020-18081-9

Haemmerle, M., Stone, R. L., Menter, D. G., Afshar-Kharghan, V., and Sood, A. K. (2018). The Platelet Lifeline to Cancer: Challenges and Opportunities. *Cancer Cell* 33 (6), 965–983. doi:10.1016/j.ccell.2018.03.002

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinformatics* 14, 7. doi:10.1186/1471-2105-14-7

Heer, E., Harper, A., Escandor, N., Sung, H., McCormack, V., and Fidler-Benaoudia, M. M. (2020). Global Burden and Trends in Premenopausal and Postmenopausal Breast Cancer: A Population-Based Study. *Lancet Glob. Health* 8 (8), e1027–e1037. doi:10.1016/s2214-109x(20)30215-1

Holinstat, M. (2017). Normal Platelet Function. *Cancer Metastasis Rev.* 36 (2), 195–198. doi:10.1007/s10555-017-9677-x

Ji, S., Zhang, B., Liu, J., Qin, Y., Liang, C., Shi, S., et al. (2016). ALDOA Functions as an Oncogene in the Highly Metastatic Pancreatic Cancer. *Cancer Lett.* 374 (1), 127–135. doi:10.1016/j.canlet.2016.01.054

Johnson, R., Sabnis, N., McConathy, W. J., and Lacko, A. G. (2013). The Potential Role of Nanotechnology in Therapeutic Approaches for Triple Negative Breast Cancer. *Pharmaceutics* 5 (2), 353–370. doi:10.3390/pharmaceutics5020353

Kajita, E., Moriwaki, J., Yatsuki, H., Hori, K., Miura, K., Hirai, M., et al. (2001). Quantitative Expression Studies of Aldolase A, B and C Genes in Developing Embryos and Adult Tissues of *Xenopus L. Mech. Dev.* 102 (1-2), 283–287. doi:10.1016/s0925-4773(01)00324-0

Kopp, H.-G., Placke, T., and Salih, H. R. (2009). Platelet-Derived Transforming Growth Factor-β Down-Regulates NKG2D Thereby Inhibiting Natural Killer Cell Antitumor Reactivity. *Cancer Res.* 69 (19), 7775–7783. doi:10.1158/0008-5472.can-09-2123

Kotani, S., Yoda, A., Kon, A., Kataoka, K., Ochi, Y., Shiozawa, Y., et al. (2019). Molecular Pathogenesis of Disease Progression in MLL-Rearranged AML. *Leukemia* 33 (3), 612–624. doi:10.1038/s41375-018-0253-3

Kusakabe, T., Motoki, K., and Hori, K. (1997). Mode of Interactions of Human Aldolase Isozymes with Cytoskeletons. *Arch. Biochem. Biophys.* 344 (1), 184–193. doi:10.1006/abbi.1997.0204

Lin, Z., He, R., Luo, H., Lu, C., Ning, Z., Wu, Y., et al. (2018). Integrin-β5, a miR-185-Targeted Gene, Promotes Hepatocellular Carcinoma Tumorigenesis by Regulating β-Catenin Stability. *J. Exp. Clin. Cancer Res.* 37 (1), 17. doi:10.1186/s13046-018-0691-9

Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118

Mehla, K., and Singh, P. K. (2019). Metabolic Regulation of Macrophage Polarization in Cancer. *Trends Cancer* 5 (12), 822–834. doi:10.1016/j.trecan.2019.10.007

Menter, D. G., Kopetz, S., Hawk, E., Sood, A. K., Loree, J. M., Gresele, P., et al. (2017). Platelet "First Responders" in Wound Response, Cancer, and Metastasis. *Cancer Metastasis Rev.* 36 (2), 199–213. doi:10.1007/s10555-017-9682-0

Mundell, S. J., and Mumford, A. (2018). TBXA2R Gene Variants Associated with Bleeding. *Platelets* 29 (7), 739–742. doi:10.1080/09537104.2018.1499888

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Nieswandt, B., Hafner, M., Echtenacher, B., and Männel, D. N. (1999). Lysis of Tumor Cells by Natural Killer Cells in Mice Is Impeded by Platelets. *Cancer Res.* 59 (6), 1295–1300.

Peng, Y., Li, X., Wu, M., Yang, J., Liu, M., Zhang, W., et al. (2012). New Prognosis Biomarkers Identified by Dynamic Proteomic Analysis of Colorectal Cancer. *Mol. Biosyst.* 8 (11), 3077–3088. doi:10.1039/c2mb25286d

Placencio, V. R., and DeClerck, Y. A. (2015). Plasminogen Activator Inhibitor-1 in Cancer: Rationale and Insight for Future Therapeutic Testing. *Cancer Res.* 75 (15), 2969–2974. doi:10.1158/0008-5472.can-15-0876

Pulley, J. M., Jerome, R. N., Ogletree, M. L., Bernard, G. R., Lavieri, R. R., Zaleski, N. M., et al. (2018). Motivation for Launching a Cancer Metastasis Inhibition (CMI) Program. *Targ Oncol.* 13 (1), 61–68. doi:10.1007/s11523-017-0542-1

Riess, L. (1872). Zur pathologischen Anatomie des Blutes. *Arch. Anat. Physiol. Wissensch Med.* 39, 237–249.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* 26 (1), 139–140. doi:10.1093/bioinformatics/btp616

Saloustros, E., Salpea, P., Qi, C.-F., Gugliotti, L. A., Tsang, K., Liu, S., et al. (2015). Hematopoietic Neoplasms in Prkar2a-Deficient Mice. *J. Exp. Clin. Cancer Res.* 34, 143. doi:10.1186/s13046-015-0257-z

Su, G., Hodnett, M., Wu, N., Atakilit, A., Kosinski, C., Godzich, M., et al. (2007). Integrin αvβ5 Regulates Lung Vascular Permeability and Pulmonary Endothelial Barrier Function. *Am. J. Respir. Cel Mol Biol* 36 (3), 377–386. doi:10.1165/rcmb.2006-0238oc

van der Meijden, P. E. J., and Heemskerk, J. W. M. (2019). Platelet Biology and Functions: New Concepts and Clinical Perspectives. *Nat. Rev. Cardiol.* 16 (3), 166–179. doi:10.1038/s41569-018-0110-0

Vinholt, P. J. (2019). The Role of Platelets in Bleeding in Patients with Thrombocytopenia and Hematological Disease. *Clin. Chem. Lab. Med.* 57 (12), 1808–1817. doi:10.1515/cclm-2019-0380

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A Class Discovery Tool with Confidence Assessments and Item Tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Wojtukiewicz, M. Z., Sierko, E., Hempel, D., Tucker, S. C., and Honn, K. V. (2017). Platelets and Cancer Angiogenesis Nexus. *Cancer Metastasis Rev.* 36 (2), 249–262. doi:10.1007/s10555-017-9673-1

Xie, X., Wang, J., Shi, D., Zou, Y., Xiong, Z., Li, X., et al. (2019). Identification of a 4-mRNA Metastasis-Related Prognostic Signature for Patients with Breast Cancer. *J. Cel Mol Med* 23 (2), 1439–1447. doi:10.1111/jcmm.14049

Yao, D. C., Tolan, D. R., Murray, M. F., Harris, D. J., Darras, B. T., Geva, A., et al. (2004). Hemolytic Anemia and Severe Rhabdomyolysis Caused by Compound Heterozygous Mutations of the Gene for Erythrocyte/Muscle Isozyme of Aldolase, ALDOA(Arg303X/Cys338Tyr). *Blood* 103 (6), 2401–2403. doi:10.1182/blood-2003-09-3160

Yeung, J., Li, W., and Holinstat, M. (2018). Platelet Signaling and Disease: Targeted Therapy for Thrombosis and Other Related Diseases. *Pharmacol. Rev.* 70 (3), 526–548. doi:10.1124/pr.117.014530

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Zhang, H., Meltzer, P., and Davis, S. (2013). RCircos: an R Package for Circos 2D Track Plots. *BMC Bioinformatics* 14, 244. doi:10.1186/1471-2105-14-244

Zheng, S., Zou, Y., Liang, J. Y., Xiao, W., Yang, A., Meng, T., et al. (2020). Identification and Validation of a Combined Hypoxia and Immune index for Triple-negative Breast Cancer. *Mol. Oncol.* 14 (11), 2814–2833. doi:10.1002/1878-0261.12747

Zheng, S., Zou, Y., Xie, X., Liang, J. y., Yang, A., Yu, K., et al. (2020). Development and Validation of a Stromal Immune Phenotype Classifier for Predicting Immune Activity and Prognosis in Triple-negative Breast Cancer. *Int. J. Cancer* 147 (2), 542–553. doi:10.1002/ijc.33009

Zhu, C., Kong, Z., Wang, B., Cheng, W., Wu, A., and Meng, X. (2019). ITGB3/CD61: A Hub Modulator and Target in the Tumor Microenvironment. *Am. J. Transl Res.* 11 (12), 7195–7208.

Zou, Y., Zheng, S., Deng, X., Yang, A., Xie, X., Tang, H., et al. (2019). The Role of Circular RNA CDR1as/ciRS-7 in Regulating Tumor Microenvironment: A Pan-Cancer Analysis. *Biomolecules* 9 (9), 429. doi:10.3390/biom9090429

Zou, Y., Zou, X., Zheng, S., Tang, H., Zhang, L., Liu, P., et al. (2020). Efficacy and Predictive Factors of Immune Checkpoint Inhibitors in Metastatic Breast Cancer: A Systematic Review and Meta-Analysis. *Ther. Adv. Med. Oncol.* 12, 1758835920940928. doi:10.1177/1758835920940928

Zou, Y., Hu, X., Zheng, S., Yang, A., Li, X., Tang, H., et al. (2021). Discordance of Immunotherapy Response Predictive Biomarkers between Primary Lesions and Paired Metastases in Tumours: A Systematic Review and Meta-Analysis. *EBioMedicine* 63, 103137. doi:10.1016/j.ebiom.2020.103137

Zynda, E. R., Matveev, V., Makhanov, M., Chenchik, A., and Kandel, E. S. (2014). Protein Kinase A Type II-α Regulatory Subunit Regulates the Response of Prostate Cancer Cells to Taxane Treatment. *Cell Cycle* 13 (20), 3292–3301. doi:10.4161/15384101.2014.949501

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: 08 March 2022
doi: 10.3389/fonc.2022.796263

Tumor Mutational Burden Associated With Response to Hyperthermic Intraperitoneal Chemotherapy

Lisi Zeng[1†], Xubo Huang[1†], Yun Tian[2†], Jinxia Huang[1], Huiyan Liu[1], Juncai Wen[2], Kaihua Liu[3], Yang Shao[3], Jiali Luo[1], Hongsheng Tang[2], Quanxing Liao[2], Ziying Lei[2], Weiwen Cui[4], Qianghua Xia[1], Tianpei Guan[2], Jin Li[1*] and Shuzhong Cui[2,5*]*

[1] Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou, China, [2] Department of Abdominal Surgery, Affiliated Cancer Hospital & Institute of Guangzhou Medical University, Guangzhou, China, [3] Medical Department, Nanjing Geneseeq Technology Inc., Nanjing, China, [4] Department of Bioengineering, University of California, Berkeley, Berkeley, CA, United States, [5] State Key Laboratory of Respiratory Disease, Guangzhou Medical University, Guangzhou, China

Background: Gastric cancer (GC) is one of the most common cancer types, especially in Asian countries. Hyperthermic intraperitoneal chemotherapy (HIPEC) has been shown to improve the progression-free survival among gastric cancer patients with peritoneal metastases; however, not all patients demonstrate response to HIPEC.

Methods: Biomarkers are needed to select patients for effective treatment of HIPEC. Here, we performed whole-exome sequencing on tumor samples from 18 gastric cancer patients who received HIPEC treatment and assessed the association between genomic mutation features and progression-free survival. Exome sequencing was further conducted on tumor samples from additional 15 gastric cancer patients as a replication study.

Results: The tumor mutational burden (TMB) was significantly higher in the group of patients with a better response to HIPEC treatment than that of the others. Kaplan–Meier survival curve showed that patients with high TMB had a significantly longer survival time than that in patients with low TMB. This discovery was validated in the replication cohort. Genes bearing mutations recurrently and selectively in patients with better response to HIPEC were found in the two cohorts.

Conclusion: We found that higher TMB is significantly associated with better response to HIPEC. Our results provide useful hints for prognostic stratification of HIPEC treatment.

Keywords: gastric cancer, hyperthermic intraperitoneal chemotherapy, survival, tumor mutational burden, biomarker

OPEN ACCESS

Edited by:
Feng He,
Shanghai University of Traditional Chinese Medicine, China

Reviewed by:
Shiwei Liu,
Harvard University, United States
Wenhua Xu,
University of Kentucky, United States
Cai Chen,
Merck, United States
Digant Nayak,
The University of Texas Health Science Center at San Antonio, United States

***Correspondence:**
Shuzhong Cui
cuishuzhong@gzhmu.edu.cn
Jin Li
jinli@gzhmu.edu.cn
Tianpei Guan
gtp120@126.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *16 October 2021*
Accepted: *04 February 2022*
Published: *08 March 2022*

Citation:
Zeng L, Huang X, Tian Y, Huang J, Liu H, Wen J, Liu K, Shao Y, Luo J, Tang H, Liao Q, Lei Z, Cui W, Xia Q, Guan T, Li J and Cui S (2022) Tumor Mutational Burden Associated With Response to Hyperthermic Intraperitoneal Chemotherapy. Front. Oncol. 12:796263. doi: 10.3389/fonc.2022.796263

INTRODUCTION

Gastric cancer is among the most common cancer types and a leading cause of cancer death in both men and women worldwide (1), especially in Asian countries, which imposes a considerable global health burden. In addition, when gastric cancer is diagnosed, patients are often already at an advanced stage. Surgery with subtotal gastrectomy or total gastrectomy is the current mainstay of

treatment (1). Surgical resection is a curative therapeutic approach for gastric cancer and confers good outcome for early-stage gastric cancer. However, early gastric cancer often remains asymptomatic unless detected by endoscopy and biopsy. The survival rate for patients with metastatic gastric cancer is very low, ranging from 4 to 12 months (1), despite the successful application of modern chemotherapy to other solid tumors.

Hyperthermic intraperitoneal chemotherapy (HIPEC) is carried out by perfusing the abdominal cavity with 43°C circulating hyperthermic saline containing anticancer drugs. HIPEC treatment has been applied to gastric cancer, colorectal cancer, ovarian cancer, pancreatic cancer, and other abdominal cancer types. Poor peritoneal penetration of chemotherapy drugs is an outstanding limitation in the treatment of gastric cancer. HIPEC has thus been administered to patients in combination with cytoreductive surgery to achieve better therapeutic effects. Cumulative evidence has shown the beneficial treatment effect of HIPEC. Studies showed that a combination of cytoreductive surgery (CRS) and HIPEC can reduce the incidence of peritoneal recurrence of advanced gastric cancer and improve the median survival (2–4). A systematic review (5) on 7 studies (6–12) has been done to compare the prophylactic HIPEC after surgery for patients with gastric cancer without clinically evident metastases or positive peritoneal cytology sign. Their analysis results suggested that, compared to surgery alone, the combination with HIPEC may decrease peritoneal recurrence and increase survival rate without affecting the morbidity and mortality of patients despite the overall risk of bias in these studies that is likely due to the non-standard of care in the studies that was carried out more than 10 years ago. A meta-analysis of randomized controlled trials in patients with advanced gastric cancer and peritoneal metastases also showed a beneficial effect of HIPEC in terms of 3-year survival rate and complete response rate (13). Continuous breakthroughs have been made in developing the theoretical basis and technical execution in HIPEC, which further improved the treatment effect (14, 15). Because of the promising treatment of HIPEC, the International Conference on Peritoneal Cancer in Amsterdam (Netherlands) and the American Anti-Cancer Association adopted CRS combined with HIPEC as the standard treatment for gastric cancer with peritoneal metastasis (16, 17).

Similar to the fact that cancer patients respond distinctly to immunotherapy, patients also have different responses to HIPEC treatment. The different results observed regarding the efficacy of HIPEC imply that not all patients can benefit from the treatment and proper stratification may be necessary before HIPEC treatment (14, 15). Few studies have been carried out to examine the biomarkers that can effectively predict the efficacy of HIPEC and stratify patients for HIPEC treatment. Through a candidate gene approach, three studies identified biomarkers, the expression level of which can be predictive for the efficacy or resistance to HIPEC treatment on patients with colorectal peritoneal metastasis, including Bloom syndrome protein (BLM) (18), Vascular endothelial growth factor (VEGF), Versican (VCAN) (19), PAX interacting protein 1 (PAXIP1), and Single stranded DNA binding protein 2 (SSBP2) (20). A recent study focusing on candidate genetic variants reported the association of NQO1∗3 allele with poor peritoneal recurrence rate and low disease-free survival (21) among colorectal cancer patients who received cytoreductive surgery plus hyperthermic intraperitoneal mitomycin C. There was no study examining the genetic association with efficacy of HIPEC on gastric cancer. Furthermore, an unbiased study *via* a genomic approach may be more fruitful in identifying biomarkers to predict HIPEC efficacy.

METHODS

Tumor Samples and Clinical Characteristics

The study was approved by the institutional review boards of the Affiliated Cancer Hospital of Guangzhou Medical University. All tissue samples were obtained with the approval of patients' consents from 2010 to 2017. The recruited patients met the following criteria. They had proven gastric cancer with histopathology and received CRS and closed HIPEC treatment after surgery. HIPEC was administered intraperitoneally with chemotherapeutic agents in 4–6 L of perfusate at a temperature of 43°C for 90 min with a flow rate of 400–600 ml/min. The resected tumor samples were examined by the pathology department to confirm the stage of tumor tissue. The tumor tissues and adjacent regions were prepared as formalin-fixed paraffin-embedded (FFPE) tissue blocks following the standard protocol of the diagnosis laboratory.

Approximately 30% of GC patients have regional spread at diagnosis; the local regional progression of gastric cancer generally results in peritoneal metastases (PMs), which have a significant negative impact on the overall survival (OS) and quality of life as a result of refractory ascites, progressive intestinal obstruction, and uncontrollable abdominal pain (22). CRS+HIPEC could obviously decrease the volumes of ascites for a long time compared to only CRS for advanced GC patients. We checked the OS as the criteria of HIPEC effect. Patients with survival time of more than 1 year after HIPEC treatment were classified as the durable clinical benefit (DCB) group, and those with less than 1 year were included in the no durable benefit (NDB) group.

Statistical Analysis

The data analysis was conducted using IBM SPSS Statistics 24.0 (SPSS, Inc., Chicago, IL, USA). Continuous data were expressed as mean ± standard deviation (SD), and inter-group comparison was conducted using an independent-samples t-test. Count data were expressed by percentage or constituent ratio, and the comparison between groups was carried out by chi-square test or exact probabilities method. The level of statistical significance was set at $p < 0.05$ and $\alpha = 0.05$. Specifically, the comparison of baseline characteristics of patients was conducted with the following statistical methods: age between groups was examined by unpaired t test; the comparison of sex, Eastern Cooperative Oncology Group (ECOG), and degree of

differentiation between groups was examined by chi-square test; and the comparison of the degree of peritoneal metastasis (P degree), ascites, and number of chemotherapy in half a year between groups was examined by Fisher exact test. The Kaplan–Meier (K-M) method and the log-rank test were used to evaluate OS.

Whole-Exome Sequencing

Tumor samples and para-cancer control samples from FFPE tissue blocks were analyzed by whole-exome sequencing (WES). The para-cancer control samples were taken from the normal tissue within the 1–5-cm distance from the visible tumor area. Library preparations were performed with KAPA Hyper Prep Kit (KAPA Biosystems, USA). Target enrichment was performed using the xGen Exome Research Panel and Hybridization and Wash Reagents Kit (Integrated DNA Technology, USA) according to manufacturer's protocol. Standard WES was performed with paired-end sequencing on the Illumina HiSeq2000 platform to generate reads of 2 × 100 bp with an average of 200× mean target coverage for tumor samples and 20× mean coverage for controls.

Exome Analysis Pipeline

Quality control and filtering steps were performed on the raw sequencing data, and the data were further checked using software FastQC, including per sequence quality scores, GC content, per base sequence quality, per base sequence content, sequence duplication levels, and overrepresented sequence. Then, the Quality Control (QC)'ed sequencing data were aligned to the Grch37 genome built using Burrows–Wheeler Aligner (BWA) (v0.7.12) (23). Picard Tool was used to mark duplicates, and Genome Analysis Tool Kit (GATK) was used to conduct indel realignment, base-quality score recalibration, and duplicate-read removal (24). Variant calls for both Single Nucleotide Variation (SNV) and Insertion Deletion (INDEL) were generated using software VarDict (25) based on the paired tumor–normal variant calling algorithm. Variant annotation was performed using ANNOVAR (26) and vcf2maf on the VCF file, and further filtering was described below.

Variant Filtering

Variants were filtered out in any of the following conditions: (allele frequency × read depth <6) and (mean number of mismatches >1.0 and mean mapping quality <55.0); or (mean number of mismatches >2.0 and mean mapping quality <60.0); or (read depth <10) and (alternate allele quality <45); or alternate allele quality <55 and allele frequency <0.2 and somatic variant p-value >0.06. In addition, variants were filtered out if the reference genotype likelihood of their para-cancer control were >3.5. To keep stringent data quality, we only kept strong somatic variants. Then, variants were further filtered based on their predicted effect on protein structure and function and variant allele frequency in large population databases. Variants that might have an impact on protein function (e.g., in_frame_del, in frame_ins, missense_mutation, nonsense_mutation, nonstop_mutation, splice_site, and translation_start_site) and allele frequency <0.01 in the following databases were kept, including ExAC_EAS, gnomAD_exom_ALL, gnomAD_exom_EAS, gnomAD_genome_ALL, gnomAD_genome_EAS, 1000g2015aug_all, and 1000g2015aug_eas.

Analysis of Tumor Mutational Burden on HIPEC Efficacy

Tumor mutational burden (TMB) was defined as the total number of strong somatic non-synonymous mutations in each tumor exome that passed our QC filtering. Mann–Whitney test was used to compare the number of somatic mutations in DCB and NDB groups. Patients were divided into the high mutation group and the low mutation group based on the median of somatic mutation number of all patients, which is 373. It is about 9.8 mutations/Mb, which is similar to the TMB threshold adopted in other studies (27, 28). Log-rank test was used to compare K-M survival curves between the two groups.

Association Analysis at Single-Variant Level

We first evaluated the association of variants with HIPEC efficacy at the single-variant level by Fisher's exact test, and the statistical significance threshold was p-value <0.05.

Association Analysis at the Gene Level

RVTESTS (29) were used to perform correlation analysis of rare variants at the gene level. We used the VCF file generated from the above pipeline as the input file for RVTESTS, followed by Fisher's exact test (CMC Fisher) to evaluate genes being positively or negatively associated with the efficacy of HIPEC.

Pathway Enrichment Analysis of Genes Associated With HIPEC Efficacy

Pathway enrichment analysis was performed *via* WEB-based GEne SeT AnaLysis Toolkit (30) (http://www.webgestalt.org/) that is based on the hypergeometric test.

From International Cancer Genome Consortium (ICGC) databases (https://dcc.icgc.org/), whole-genome sequencing data and follow-up data of Chinese population were obtained, and the TMB value of non-synonymous mutations in the exon regions was calculated for each patient. The patients were divided into high TMB group and low TMB group by the median value of TMB, and K-M survival curve was performed to examine the association between TMB and survival.

RESULTS

High Tumor Mutational Burden Was Associated With Improved Patient Survival

We conducted a retrospective study to identify genetic determinants for HIPEC response involving 18 gastric cancer patients, among whom 8 patients showed DCB (patients with survival time of more than 1 year) and 10 patients had NDB. The baseline characteristics of these patients are shown in **Table 1**. The average age of the NDB group is significantly lower than that of the DCB group (43.8 ± 9.7 vs. 54.1 ± 8.1, p = 0.030), while the other

TABLE 1 | The baseline characteristics of patients who received hyperthermic intraperitoneal chemotherapy (HIPEC).

	DCB (*n* = 8)	NDB (*n* = 10)	P
Age	54.1 ± 8.1	43.8 ± 9.7	0.030*
Sex			
Male	3 (37.5)	4 (40.0)	0.999
Female	5 (62.5)	6 (60.0)	
ECOG			
0~1 score	7 (87.5)	9 (90.0)	0.999
2~4 score	1 (12.5)	1 (10.0)	
Degree of differentiation			
G1+G2	2 (25.0)	2 (20.0)	0.999
G3+GX	6 (75.0)	8 (80.0)	
P degree			
p1x	1 (12.5)	0 (0.0)	0.999
p1	1 (12.5)	2 (20.0)	
p2	3 (37.5)	4 (40.0)	
p3	3 (37.5)	4 (40.0)	
Ascites			0.552
No	6 (75.0)	5 (50.0)	
A small amount number of chemotherapy in half a year	2 (25.0)	5 (50.0)	
1~3	7 (87.5)	7 (70.0)	0.751
4~6	1 (12.5)	3 (30.0)	

ECOG, the grade of Eastern Cooperative Oncology Group.
They were classified into durable clinical benefit (DCB) group and no durable benefit (NDB) group based on their response to treatment.
**p-value < 0.05.*

characteristics were similar between DCB and NDB groups. We analyzed the WES data of these tumor samples with their matched normal tissues for germline references. The mean target coverage in the WES data of tumor samples is 200× and, on average, >90% of the target sequence was covered to a depth of >10×. The number of non-synonymous somatic mutations in all patients ranged from 90 to 3,027, with a median of 373.

We first analyzed the overall TMB of non-synonymous strong somatic mutation with low frequency in a large population database (variant allele frequency <0.01). The TMB was defined as the total number of strong somatic non-synonymous mutations in the coding regions of the human genome. The TMB in the DCB group ranges from 113 to 3,027, with a median of 799.5; and the TMB in the NDB group ranges from 90 to 569, with a median of 278. Therefore, patients with a partial or stable response to HIPEC had a significantly higher TMB than patients with NDB (p = 0.034) (**Figure 1A**). We noticed that there is a subject in the DCB group carrying >3,000 mutations. After removing this sample, the comparison yielded a result of p = 0.07 with marginal significance but still suggests a positive correlation between TMB and response to HIPEC. The K-M survival curves showed significantly better survival in the DCB group [median OS 1,754 vs. 186 days, log-rank p = 0.0001; hazard ratio (HR) = 10.47, 95% confidence interval (CI) ranged from 3.139 to 34.93].

Patients were divided into the high mutation group and the low mutation group based on the median of somatic mutation number of all patients. To measure the effect of TMB on the OS of GC patients, we conducted K-M survival analysis. K-M survival curve showed that patients with high TMB had a significant longer survival time than that of patients with low TMB (HR = 4.82, 95% CI = 1.46–15.95, p = 0.001; **Figure 1B**). In addition, analyzing the data of Chinese gastric cancer patients in the ICGC database, we found no significant association between TMB and patient survival time (p > 0.05; **Figure 1C**). In addition, a previous research with comparable sequencing approach among 262 GC patients after excluding patients who received routine preoperative chemoradiotherapy or biological immunotherapy showed that the prognosis of GC and OS are better among patients with higher TMB than those with lower TMB (28). Thus, the association of high TMB and better response to HIPEC observed in our study is not likely due to the relationship between TMB and prognosis of gastric cancer *per se*. The above evidence suggests an association between high TMB and longer survival for patients who received HIPEC treatment.

We further conducted a replication study. The replication cohort is composed of an independent set of 15 GC samples from patients with similar HIPEC treatment, including 10 patients in the DCB group and 5 patients in the NDB group according to the same criteria. The baseline clinical characteristics of the replication cohort were similar to those of the discovery cohort (**Table 2**). The median TMB was 364.5 in tumors from patients in the DCB group compared to 233 in those in the NDB group (Mann–Whitney p = 0.02) (**Figure 1D**). With the same TMB criteria, the GC samples were classified as high TMB group (n = 7) and low TMB group (n = 8). K-M survival analysis was similarly conducted to measure the effect of TMB on the OS of GC patients in the replication cohort. The results confirmed that patients with high TMB had a significantly longer survival time than that in patients with low TMB (HR = 5.842, 95% CI = 1.179–28.94, p = 0.031; **Figure 1E**).

Single Variant Associated With HIPEC Therapy

We then sought to identify mutations associated with efficacy of HIPEC therapy. The results based on Fisher's exact test did not suggest any strong somatic mutations surpassing statistical significance, which is likely due to the limitation of a small sample size. Three mutations occurred in either the DCB or the NDB group specifically without carriers in the other group (**Table 3**).

Genes Correlated With HIPEC Treatment

We then focused on identifying genes harboring mutations recurrently and selectively associated with response or resistance to HIPEC. We found that genes *GPI, PCDH9*, and *C21ORF140* harbored mutations in three or more NDB patients but none in the DCB group, and further statistical analyses of collapsing variants to the gene level showed a significant negative correlation with HIPEC response (p < 0.05). On the other hand, a total of 25 genes bearing mutations are recurrently and selectively enriched in the DCB group but none in the NDB group (**Figures 2A, B**). Several of these genes were known to be involved in regulating cancer proliferation, metastasis, or invasion (**Supplementary Table S1**). Analyzing the association

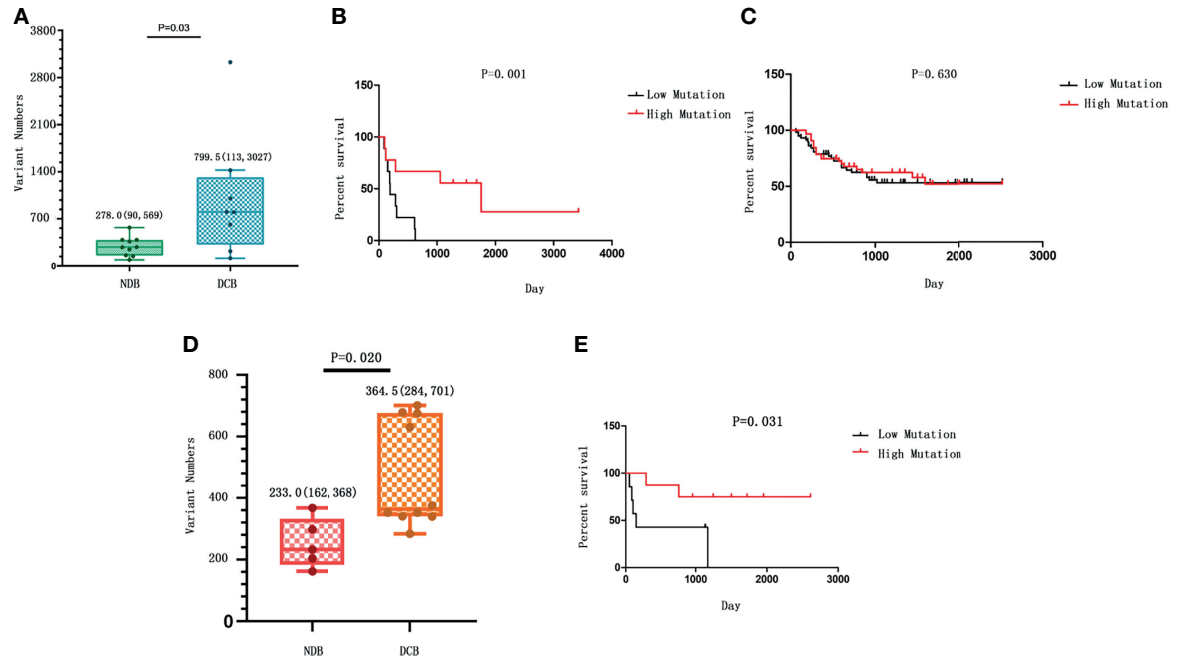


FIGURE 1 | The correlation between tumor mutational burden (TMB) and response to hyperthermic intraperitoneal chemotherapy (HIPEC). **(A)** The number of strong somatic mutations compared between the durable clinical benefit (DCB) and the no durable benefit (NDB) groups. The median value and the range of the data in each group are shown. **(B)** The comparison of survival rate between gastric cancer patients with high TMB and those with low TMB in our study who received HIPEC treatment. **(C)** The survival analysis between patients with high TMB and those with low TMB from the ICGC database without HIPEC treatment. In **(B, C)**, the overall survival time is shown on the X-axis, and the survival rate with a percentage scale is shown on the Y-axis. **(D)** The number of strong somatic mutations compared between the DCB and the NDB groups in the replication cohort. The median value of each group and the standard deviation showing data variability are shown. **(E)** The comparison of survival rate between gastric cancer patients with high TMB and those with low TMB in the replication cohort who received HIPEC treatment.

TABLE 2 | The baseline characteristics of patients who received hyperthermic intraperitoneal chemotherapy (HIPEC) in the replication cohort.

	DCB (n = 10)	NDB (n = 5)	p
Age	50.80 ± 10.63	41.24 ± 15.90	0.186
Sex			
Male	8 (80)	1 (80)	0.089
Female	2 (20)	4 (20)	
ECOG			
0–1 score	10 (100)	5 (100)	0.999
2–4 score	0 (0)	0 (0)	
Degree of differentiation			
G1+G2	2 (20)	0 (0)	0.524
G3+G_X	8 (80)	5 (100)	
P degree			
p1x	7 (80)	1 (20)	0.119
p1-p3	3 (20)	4 (80)	
Ascites			
no	8 (80)	5 (100)	0.524
A small amount	2 (20)	0 (0)	
Number of chemotherapy in half a year			
1–3	4 (40)	3 (60)	0.364
4–6	6 (60)	1 (20)	
≥7	0 (0)	1 (20)	

ECOG, the grade of Eastern Cooperative Oncology Group; DCB, group with durable clinical benefit; NDB, group without durable benefit.

between non-synonymous mutations in these genes and OS of Chinese gastric cancer patients in the ICGC database, we did not find any significant association with OS. We checked the distribution of the above enriched single variants and genes in the replication cohort. Gene MUC16 harbored mutations in three patients of the DCB group but none in the NDB group; gene DNAH3 contained mutations in four patients of the DCB group but none in the NDB group; the other genes did not exhibit such enrichment. It is not unexpected for discrepancies to appear at the variant and gene levels between the discovery and replication cohorts, considering the sample size. Thus, studies with larger sample sizes are certainly required to identify and validate predictive markers for HIPEC treatment.

Gene Set Enrichment Analysis of Significantly Correlated Genes

To identify potential signaling pathways that may contribute to HIPEC sensitivity, we conducted pathway enrichment on 348 genes that bear mutations occurring only in the DCB group but not the NDB group. Among the 134 genes that were mapped to Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, the pathways of extracellular matrix (ECM)–receptor interaction, Notch signaling pathway, focal expression, and

TABLE 3 | The three variants that are specifically enriched in the durable clinical benefit (DCB) group or the no durable benefit (NDB) group.

Chr	Pos	Variant	AA alteration	Cases	Controls
11	1092910	NM_002457:exon31:c.G4729A	MUC2:p.G1577S	3	0
19	23158703	NM_001267716:exon4:c.T1436C	ZNF728:p.L479P	3	0
13	103395322	NM_001146197:exon4:c.C7725A	CCDC168: p.N2575K	0	3

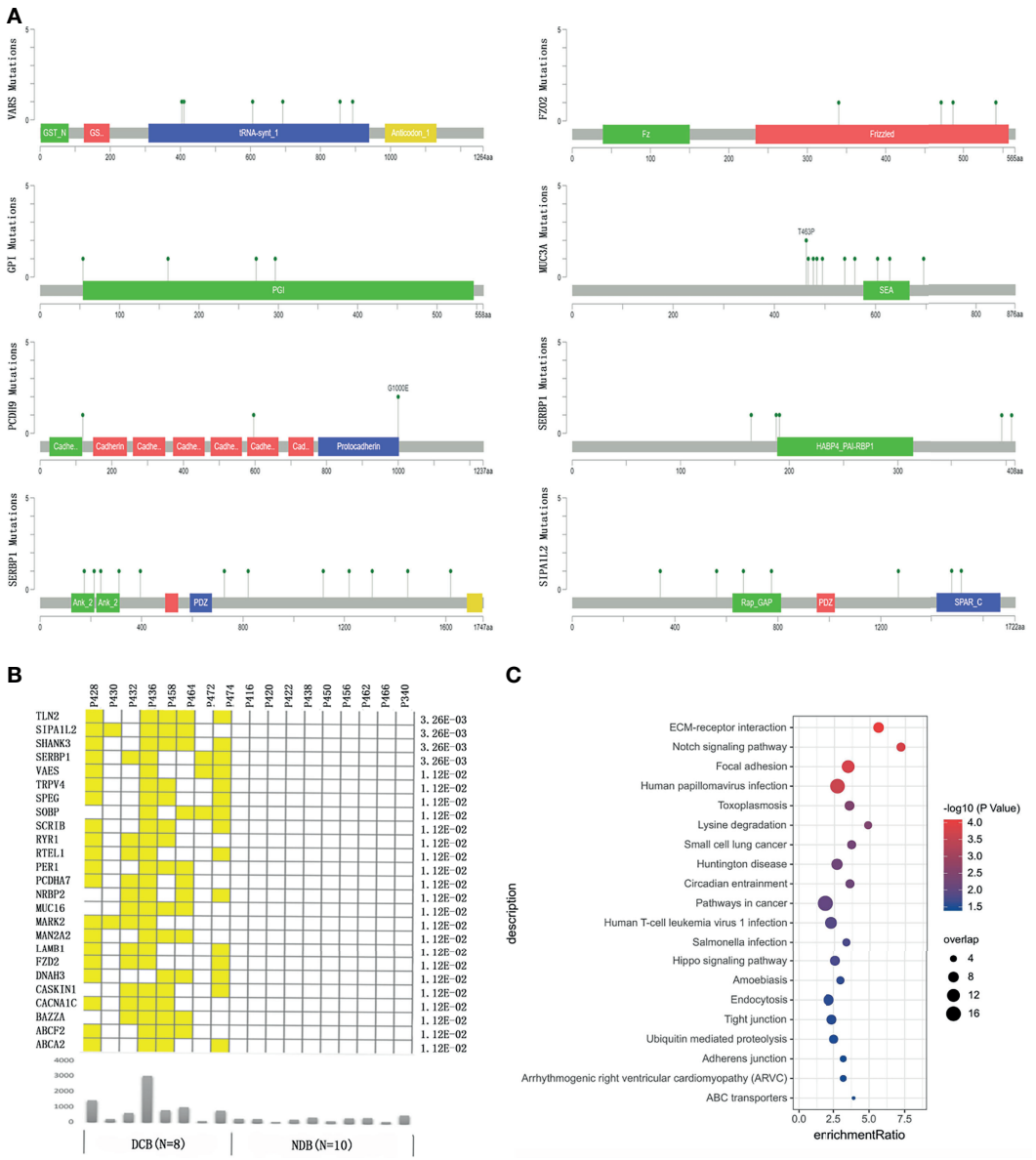


FIGURE 2 | The genes bearing mutations specifically enriched in the durable clinical benefit (DCB) group or no durable benefit (NDB) group. **(A)** The location of the mutations in the top response-associated genes with respect to the structural domains of each protein product encoded by these genes. **(B)** The genes carrying mutations that only occurred in patients in the DCB group but none in the NDB group. The yellow color indicates a patient in the DCB group carrying a mutation in the corresponding gene. The lower panel showed the total number of mutations in each patient. The number on the right shows the p-value of Fisher exact test comparing the number of GC patients carrying mutation(s) in each gene between the DCB and NDB groups. **(C)** The pathways enriched by genes with at least one mutation specifically occurring in patients in the DCB group.

human papillomavirus infection showed significant enrichment (q-value <0.05) (**Figure 2C**).

DISCUSSION

Though HIPEC has been widely administered to patients of gastric cancer, it has been controversial about its efficacy likely due to tumor heterogeneity; therefore, it is important to identify genomic determinants of response to HIPEC treatment. To our knowledge, this is the first study examining molecular features of HIPEC efficacy using an unbiased high-throughput sequencing approach. Within our discovery cohort and a replication cohort, we found a significant association between high TMB and patients' beneficial treatment effect.

HIPEC may induce immune response activation. It has been reported that focal thermal ablation of tumor may stimulate a systemic antitumor immune response (31). From the tumor cells and surrounding tissues that were damaged by heat, a variety of immune molecules were released, including cytokines and chemokines. It has been reported that the serum levels of tumor necrosis factor-α (TNFα) and interleukin-1β (IL-1β), IL-6, and IL-8 were elevated after thermal ablation (32–35). In addition, heat shock proteins (HSPs) such as HSP70, which is highly expressed by tumor cells, were detected to be increased following thermal ablation. HSPs play important roles in inhibition of apoptosis inside cells, as well as antigen-presenting and activating dendritic cells and modulating the activity of T-regulatory cells in the microenvironment of tumor cells. HSPs are also involved in other antitumor immune responses (36–39). Increased activation of dendritic cells and decreased activity of T-regulatory cells have been observed in multiple studies (32, 33). Immune cell infiltration was also observed, such as B and T lymphocytes, dendritic cells, neutrophils, natural killer (NK) cells, and macrophages. Such intense immune and inflammatory response occurring after focal thermal ablation may also be able to explain the positive correlation between TMB and response to HIPEC. We hypothesize that tumor cell damage caused by HIPEC results in the release of various intracellular molecules to the tumor microenvironment, and the higher the TMB, the more neo-antigens may be captured by immune cells, therefore leading to better response after treatment.

We identified multiple genes carrying mutations enriched in the DCB or NDB group specifically. The expression level or SNPs in these genes have been reported as being significantly associated with several cancer types. Functional studies showed that these genes have been implicated in key signaling pathways in cancer, such as the proliferation, migration, and invasion of tumor cells, as being summarized in **Supplementary Table S1**. However, due to the limited sample size, the individual variant or gene is not robust enough to be identified as a definitive predictive marker for responses to HIPEC treatment. A larger cohort with more samples needs to be sequenced to validate these findings, and experimental studies need to be carried out to investigate the role of these genes in molecular and cellular responses to HIPEC.

Limitations

The limitation of our study lies in three aspects. First, the sample size is small; therefore, it is not unexpected that none of the enriched variants at the single-variant level is statistically significant. The results from gene level need to be further examined in studies with larger sample sizes. Second, an independent replication study sampled from other populations is needed to investigate whether TMB that was identified as a predictive mutation feature for HIPEC efficacy can be generalized to other populations. Third, functional studies are warranted to further examine how such a mutational feature contributes to response or resistance to HIPEC treatment. Nevertheless, our study demonstrated that a high mutational burden is significantly associated with better response to HIPEC, which is consistent with previous reports that HIPEC induces immune responses and high TMB confers better immunotherapy efficacy.

Conclusion

Though limited by the small sample size and single data type at the DNA level, we conducted the first biomarker study of HIPEC treatment to gastric cancer based on unbiased high-throughput sequencing data. We found that TMB is significantly associated with HIPEC efficacy. The genes identified are relevant to the activation and inhibition of signaling pathways in cancer cells. Our results demonstrated that high mutational burden confers a beneficial HIPEC treatment effect, which warrants further confirmation in independent cohorts and functional examination of the underlying mechanism of HIPEC sensitivity. Our study provides insights into the development of mutational features to select patients for effective HIPEC treatment.

DATA AVAILABILITY STATEMENT

According to national legislation/guidelines, specifically the Administrative Regulations of the People's Republic of China on Human Genetic Resources (http://www.gov.cn/zhengce/content/2019-06/10/content_5398829.htm, http://english.www.gov.cn/policies/latest_releases/2019/06/10/content_281476708945462.htm), no additional raw data are available at this time. Data of this project can be accessed after an approval application to the China National Genebank (CNGB, https://db.cngb.org/cnsa/). Please refer to https://db.cngb.org/or email CNGBdb@cngb.org for detailed application guidance. The accession code CNP0002628 should be included in the application.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the institutional review boards of the Affiliated Cancer Hospital of Guangzhou Medical University. The patients/

participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SC and JLi designed and supervised the study. LZ and YT prepared patients' samples for sequencing and collected clinical data. LZ and XH performed data analysis. The other authors partially participated in the study and helped on data analysis and interpretation. LZ, XH, QX, JLi, and SC wrote the initial article. All authors have read and approved the article.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 81972918, 81771769, and 82172885), Natural Science Foundation of Guangdong Province (No. 2018A030310249, 2020A1515010818, and 2021A1515012392), and Key Clinical Technique of Guangzhou (No. 2019ZD16) and Guangzhou Key Medical Discipline Construction Project Fund.

ACKNOWLEDGMENTS

We are grateful to all the patients for their participation in the study and Nanjing Geneseeq Technology Inc. and 3D Medicines Inc. for conducting whole-exome sequencing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.796263/full#supplementary-material

REFERENCES

1. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric Cancer. *Lancet* (2016) 388:2654–64. doi: 10.1016/S0140-6736(16)30354-3
2. Zhu ZG, Tang R, Yan M, Chen J, Yang QM, Li C, et al. Efficacy and Safety of Intraoperative Peritoneal Hyperthermic Chemotherapy for Advanced Gastric Cancer Patients With Serosal Invasion. A Long-Term Follow-Up Study. *Dig Surg* (2006) 23:93–102. doi: 10.1159/000093778
3. Yang XJ, Li Y, Yonemura Y. Cytoreductive Surgery Plus Hyperthermic Intraperitoneal Chemotherapy to Treat Gastric Cancer With Ascites and/or Peritoneal Carcinomatosis: Results From a Chinese Center. *J Surg Oncol* (2010) 101:457–64. doi: 10.1002/jso.21519
4. Yang XJ, Huang CQ, Suo T, Mei LJ, Yang GL, Cheng FL, et al. Cytoreductive Surgery and Hyperthermic Intraperitoneal Chemotherapy Improves Survival of Patients With Peritoneal Carcinomatosis From Gastric Cancer: Final Results of a Phase III Randomized Clinical Trial. *Ann Surg Oncol* (2011) 18:1575–81. doi: 10.1245/s10434-011-1631-5
5. Brenkman HJF, Paeva M, Van Hillegersberg R, Ruurda JP, Haj Mohammad N. Prophylactic Hyperthermic Intraperitoneal Chemotherapy (HIPEC) for Gastric Cancer-A Systematic Review. *J Clin Med* (2019) 8(10):1685. doi: 10.3390/jcm8101685
6. Koga S, Hamazoe R, Maeta M, Shimizu N, Murakami A, Wakatsuki T. Prophylactic Therapy for Peritoneal Recurrence of Gastric Cancer by Continuous Hyperthermic Peritoneal Perfusion With Mitomycin C. *Cancer* (1988) 61:232–7. doi: 10.1002/1097-0142(19880115)61:2<232::AID-CNCR2820610205>3.0.CO;2-U
7. Yonemura Y, Ninomiya I, Kaji M, Sugiyama K, Fujimura K, Sawa T, et al. Prophylaxis With Intraoperative Chemohyperthermia Against Peritoneal Recurrence of Serosal Invasion-Positive Gastric Cancer. *World J Surg* (1995) 19:450–454; discussion 455. doi: 10.1007/BF00299188
8. Hirose K, Katayama K, Iida A, Yamaguchi A, Nakagawara G, Umeda S, et al. Efficacy of Continuous Hyperthermic Peritoneal Perfusion for the Prophylaxis and Treatment of Peritoneal Metastasis of Advanced Gastric Cancer: Evaluation by Multivariate Regression Analysis. *Oncology* (1999) 57:106–14. doi: 10.1159/000012016
9. Kim JY, Bae HS. A Controlled Clinical Study of Serosa-Invasive Gastric Carcinoma Patients Who Underwent Surgery Plus Intraperitoneal Hyperthermo-Chemo-Perfusion (IHCP). *Gastric Cancer* (2001) 4:27–33. doi: 10.1007/s101200100013
10. Kunisaki C, Shimada H, Nomura M, Akiyama H, Takahashi M, Matsuda G. Lack of Efficacy of Prophylactic Continuous Hyperthermic Peritoneal Perfusion on Subsequent Peritoneal Recurrence and Survival in Patients With Advanced Gastric Cancer. *Surgery* (2002) 131:521–8. doi: 10.1067/msy.2002.123769
11. Kunisaki C, Shimada H, Nomura M, Matsuda G, Otsuka Y, Ono H, et al. Therapeutic Strategy for Scirrhous Type Gastric Cancer. *Hepatogastroenterology* (2005) 52:314–8. doi: 10.1159/000084820
12. Coccolini F, Celotti A, Ceresoli M, Montori G, Marini M, Catena F, et al. Hyperthermic Intraperitoneal Chemotherapy (HIPEC) and Neoadjuvant Chemotherapy as Prophylaxis of Peritoneal Carcinosis From Advanced Gastric Cancer-Effects on Overall and Disease Free Survival. *J Gastrointest Oncol* (2016) 7:523–9. doi: 10.21037/jgo.2016.06.05
13. Liu YW, Du Y, Chen BA. Effect of Hyperthermic Intraperitoneal Chemotherapy for Gastric Cancer Patients: A Meta-Analysis of the Randomized Controlled Trials. *J Int Med Res* (2019) 47:5926–36. doi: 10.1177/0300060519882545
14. Granieri S, Bonomi A, Frassini S, Chierici AP, Bruno F, Paleino S, et al. Prognostic Impact of Cytoreductive Surgery (CRS) With Hyperthermic Intraperitoneal Chemotherapy (HIPEC) in Gastric Cancer Patients: A Meta-Analysis of Randomized Controlled Trials. *Eur J Surg Oncol* (2021) 47:2757–67. doi: 10.1016/j.ejso.2021.05.016
15. Yap DRY, Wong JSM, Tan QX, Tan JW, Chia CS, Ong CJ. Effect of HIPEC on Peritoneal Recurrence in Peritoneal Metastasis Treated With Cytoreductive Surgery: A Systematic Review. *Front Oncol* (2021) 11:795390. doi: 10.3389/fonc.2021.795390
16. Li Y, Yu Y, Liu Y. Report on the 9(Th) International Congress on Peritoneal Surface Malignancies. *Cancer Biol Med* (2014) 11:281–4. doi: 10.7497/j.issn.2095-3941.2014.04.008
17. Van Driel WJ, Koole SN, Sikorska K, Schagen Van Leeuwen JH, Schreuder HWR, Hermans RHM, et al. Hyperthermic Intraperitoneal Chemotherapy in Ovarian Cancer. *N Engl J Med* (2018) 378:230–40. doi: 10.1056/NEJMoa1708618
18. Kwakman R, De Cuba EM, De Winter JP, De Hingh IH, Delis-Van Diemen PM, Tijssen M, et al. Tailoring Heated Intraperitoneal Mitomycin C for Peritoneal Metastases Originating From Colorectal Carcinoma: A Translational Approach to Improve Survival. *Br J Cancer* (2015) 112:851–6. doi: 10.1038/bjc.2015.18
19. Sluiter NR, De Cuba EM, Kwakman R, Meijerink WJ, Delis-Van Diemen PM, Coupe VM, et al. Versican and Vascular Endothelial Growth Factor Expression Levels in Peritoneal Metastases From Colorectal Cancer Are Associated With Survival After Cytoreductive Surgery and Hyperthermic Intraperitoneal Chemotherapy. *Clin Exp Metastasis* (2016) 33:297–307. doi: 10.1007/s10585-016-9779-9
20. Shannon NB, Tan JW, Tan HL, Wang W, Chen Y, Lim HJ, et al. A Set of Molecular Markers Predicts Chemosensitivity to Mitomycin-C Following Cytoreductive Surgery and Hyperthermic Intraperitoneal Chemotherapy for Colorectal Peritoneal Metastasis. *Sci Rep* (2019) 9:10572. doi: 10.1038/s41598-019-46819-z

21. Hulshof EC, Lurvink RJ, Caserta N, De Hingh I, Van Wezel T, Bohringer S, et al. Identification of Pharmacogenetic Biomarkers for Efficacy of Cytoreductive Surgery Plus Hyperthermic Intraperitoneal Mitomycin C in Patients With Colorectal Peritoneal Metastases. *Eur J Surg Oncol* (2020) 46:1925–31. doi: 10.1016/j.ejso.2020.04.019
22. Bijelic L, Sugarbaker PH. The Role of Intraperitoneal Chemotherapy in the Treatment of Patients With Advanced Gastric Cancer. *Ann Ital Chir* (2012) 83:224–31.
23. Li H, Durbin R. Fast and Accurate Long-Read Alignment With Burrows-Wheeler Transform. *Bioinformatics* (2010) 26:589–95. doi: 10.1093/bioinformatics/btp698
24. Li H, Durbin R. Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform. *Bioinformatics* (2009) 25:1754–60. doi: 10.1093/bioinformatics/btp324
25. 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
26. Wang K, Li M, Hakonarson H. ANNOVAR: Functional Annotation of Genetic Variants From High-Throughput Sequencing Data. *Nucleic Acids Res* (2010) 38:e164. doi: 10.1093/nar/gkq603
27. Wang F, Wei XL, Wang FH, Xu N, Shen L, Dai GH, et al. Safety, Efficacy and Tumor Mutational Burden as a Biomarker of Overall Survival Benefit in Chemo-Refractory Gastric Cancer Treated With Toripalimab, a PD-1 Antibody in Phase Ib/II Clinical Trial NCT02915432. *Ann Oncol* (2019) 30:1479–86. doi: 10.1093/annonc/mdz197
28. He X, Yu M, Wang X, Chen J, Li X. Analysis of Threshold Change of Tumor Mutation Burden in Gastric Cancer. *J Oncol* (2021) 2021:3374939. doi: 10.1155/2021/3374939
29. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: An Efficient and Comprehensive Tool for Rare Variant Association Analysis Using Sequence Data. *Bioinformatics* (2016) 32:1423–6. doi: 10.1093/bioinformatics/btw079
30. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: Gene Set Analysis Toolkit With Revamped UIs and APIs. *Nucleic Acids Res* (2019) 47:W199–205. doi: 10.1093/nar/gkz401
31. Chu KF, Dupuy DE. Thermal Ablation of Tumours: Biological Mechanisms and Advances in Therapy. *Nat Rev Cancer* (2014) 14:199–208. doi: 10.1038/nrc3672
32. Ali MY, Grimm CF, Ritter M, Mohr L, Allgaier HP, Weth R, et al. Activation of Dendritic Cells by Local Ablation of Hepatocellular Carcinoma. *J Hepatol* (2005) 43:817–22. doi: 10.1016/j.jhep.2005.04.016
33. Fietta AM, Morosini M, Passadore I, Cascina A, Draghi P, Dore R, et al. Systemic Inflammatory Response and Downmodulation of Peripheral CD25 +Foxp3+ T-Regulatory Cells in Patients Undergoing Radiofrequency Thermal Ablation for Lung Cancer. *Hum Immunol* (2009) 70:477–86. doi: 10.1016/j.humimm.2009.03.012
34. Ahmad F, Gravante G, Bhardwaj N, Strickland A, Basit R, West K, et al. Changes in Interleukin-1beta and 6 After Hepatic Microwave Tissue Ablation Compared With Radiofrequency, Cryotherapy and Surgical Resections. *Am J Surg* (2010) 200:500–6. doi: 10.1016/j.amjsurg.2009.12.025
35. Erinjeri JP, Thomas CT, Samoilia A, Fleisher M, Gonen M, Sofocleous CT, et al. Image-Guided Thermal Ablation of Tumors Increases the Plasma Level of Interleukin-6 and Interleukin-10. *J Vasc Interv Radiol* (2013) 24:1105–12. doi: 10.1016/j.jvir.2013.02.015
36. Srivastava P. Interaction of Heat Shock Proteins With Peptides and Antigen Presenting Cells: Chaperoning of the Innate and Adaptive Immune Responses. *Annu Rev Immunol* (2002) 20:395–425. doi: 10.1146/annurev.immunol.20.100301.064801
37. Schueller G, Kettenbach J, Sedivy R, Stift A, Friedl J, Gnant M, et al. Heat Shock Protein Expression Induced by Percutaneous Radiofrequency Ablation of Hepatocellular Carcinoma In Vivo. *Int J Oncol* (2004) 24:609–13. doi: 10.3892/ijo.24.3.609
38. Rai R, Richardson C, Flecknell P, Robertson H, Burt A, Manas DM. Study of Apoptosis and Heat Shock Protein (HSP) Expression in Hepatocytes Following Radiofrequency Ablation (RFA). *J Surg Res* (2005) 129:147–51. doi: 10.1016/j.jss.2005.03.020
39. Teng LS, Jin KT, Han N, Cao J. Radiofrequency Ablation, Heat Shock Protein 70 and Potential Anti-Tumor Immunity in Hepatic and Pancreatic Cancers: A Minireview. *Hepatobil Pancreat Dis Int* (2010) 9:361–5.

Conflict of Interest: Authors KL and YS are employed by Nanjing Geneseeq Technology Inc.

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.

frontiers | Frontiers in Oncology

CASE REPORT
published: 16 May 2022
doi: 10.3389/fonc.2022.835220

A Novel Germline *SDHA* Gene Mutation and Co-Occurring Somatic *KIT* Activating Mutation in a Patient With Pediatric Central Nervous System Germ Cell Tumor: Case Report

Xizan Yue[1], *Bo Liu*[1], *Tiantian Han*[2,3,4], *Ningning Luo*[2,3,4], *Guanghua Lu*[2,3,4], *Didi Guo*[2,3,4], *Fanfeng Bu*[2,3,4] *and Guangyu Wang*[1]*

[1] *Department of Neurosurgery, Qilu Children's Hospital of Shandong University, Jinan, China,* [2] *The Medical Department, Jiangsu Simcere Diagnostics Co., Ltd., Nanjing, China,* [3] *The Medical Department, Nanjing Simcere Medical Laboratory Science Co., Ltd., Nanjing, China,* [4] *The State Key Lab of Translational Medicine and Innovative Drug Development, Jiangsu Simcere Diagnostics Co., Ltd., Nanjing, China*

OPEN ACCESS

Edited by:
Qinghua Xu,
The Canhelp Genomics Research Center, China

Reviewed by:
J. Wolter Oosterhuis,
Erasmus University Medical Center, Netherlands
Eyas M. Hattab,
University of Louisville, United States

***Correspondence:**
Guangyu Wang
wgywjc@163.com

Specialty section:
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

Received: *14 December 2021*
Accepted: *24 March 2022*
Published: *16 May 2022*

Citation:
Yue X, Liu B, Han T, Luo N, Lu G, Guo D, Bu F and Wang G (2022) A Novel Germline SDHA Gene Mutation and Co-Occurring Somatic KIT Activating Mutation in a Patient With Pediatric Central Nervous System Germ Cell Tumor: Case Report. Front. Oncol. 12:835220. doi: 10.3389/fonc.2022.835220

Central nervous system germ cell tumors (CNS GCTs) are a heterogeneous group of primary CNS tumors. GCTs are more common and mostly observed in pediatric and young adult patients. CNS GCTs are divided into germinomas and non-germinomatous germ cell tumors (NGGCTs), with different therapeutic strategies depending on diagnosis. Herein, we report a patient with pediatric central nervous system germinoma harboring a somatic *KIT* p.Y823D and a heterozygous germline *SDHA* p. T396Nfs*14 mutation detected by next generation sequencing. After surgery, the patient received chemotherapy (temozolomide + nedaplatin + etoposide). This is the first report of a Chinese pediatric patient with CNS GCT harboring concurrent germline *SDHA* and somatic *KIT* mutation, which enriches molecular profiles of CNS GCTs and provides more molecular evidence of clinical diagnosis and potential targeted therapy in CNS GCTs.

Keywords: central nervous system germ cell tumors, *SDHA* gene, *KIT* gene, pediatric, germline

INTRODUCTION

Central nervous system (CNS) germ cell tumors (GCTs) are rare tumors originating primarily from midline locations, including pineal and suprasellar regions. The 2021 world health organization (WHO) classification of tumors of the CNS identified 8 GCTs subtypes: mixed germ cell tumor; germinoma (GE); nongerminomas germ cell tumors (NGGCTs) including choriocarcinoma, yolk sac tumor, embryonal carcinoma; and teratoma classified further as mature teratoma, immature teratoma and teratoma with somatic-type malignancy (1). The incidence of CNS GCTs was once considered more prevalent in East Asia than in the Western Countries (2). But the study by McCarthy et al. revealed that the incidence of primary GCTs between Japan and the United States is

similar, and the gender-based patterns by location and high rates of survival are similar too (3). The median age at which CNS GCTs are diagnosed is 10–14 years, and there is a significant predominance in males (4, 5). GE has a relatively better prognosis than NGGCTs, with a 5-year survival rate of approximately 90%, which in NGGCTs is 75–82% (5, 6). CNS GCTs are very responsive to radiation therapy/chemotherapy. NGGCTs usually secrete elevated alpha-fetoprotein (AFP) and/or beta-human chorionic gonadotropin (beta-hCG) into serum and cerebrospinal fluid, so AFP and beta-hCG are helpful in distinguishing GEs and NGGCTs. Different tests including CT or MRI scans, tumor markers in serum and cerebrospinal fluid, and histopathology are used to diagnose GCT, but there are still difficulties (7, 8). To date, there are limited large-scale data of GCTs in China, with only 2 large studies (9, 10).

Molecular characterization studies have shown that *KIT* or *KRAS* gene mutations are present in about 25% of seminomas, yet it is rare in non-seminomatous germ cell tumor (11). Gain of chromosome 12p exists in almost all testicular GCTs (12). The gain of chromosome 12p is observed in CNS GCTs with lower frequency (13). With the application of next-generation sequencing (NGS) analysis in recent years, the combination of activated KIT signaling pathway and complex chromosomal anomalies, especially gain of chromosome 12p seems to be the key molecular drive of gonadal GCTs pathogenesis (11, 12). Recently, whole-exome sequencing of CNS GCTs reveals mutations in the KIT-RAS-MAPK or AKT-MTOR pathways, including *KIT* (26%), *KRAS/NRAS* (20%), *CBL* (11%), *MTOR* (8%), and *NF1* (3%) (14). Moreover, a rare germline variant of *JMJD1C* S880P is associated with CNS GCTs, especially in Japanese. Germline *SDHA* mutation has not been reported in GCTs (14). Nonetheless, the complete molecular characterization of CNS GCTs in Chinese has not been achieved. In the present case, for the first time, we report a Chinese CNS germinoma patient with a *KIT* hotspot somatic mutation concurrent with a novel *SDHA* germline mutation that showed distinct early-onset.

CASE PRESENTATION

A 4-year-old male toddler presented to our hospital due to the elevated serum HCG levels, produced by the tumor. On examination, the muscle strength of the lower limbs was grade 4 and the skin of the whole body was hairy, with a beard and pubic hair. The development of external genitalia was significantly larger than the boys of the same age. An irregular cafe-au-lait macule could be seen on the right forehead, which might be considered as a sign of the syndrome. Brain magnetic resonance imaging (MRI) scans revealed a space-occupying lesion in the right lateral ventricle with hydrocephalus (**Figures 1A–C**). Increasing levels of β-hCG (207.72 U/L—NV 0-5) were found in serum. The patient had epiphyseal dysplasia of the lower limbs with a slight limp in the right lower limb. The patient underwent space-occupying resection of the lesions and external ventricular drain (**Figures 1D–F**). Microscopical analysis indicated that the tumor was a malignant tumor of germ cell origin without any nongerminomas germ cell tumors components (**Figure 1J**). Immunochemistry was performed and showed AFP/HCG negative and PLAP/CD117 positive by postoperative tissue, and probably the HCG-producing trophoblastic giant cells have been missed. The Ki67 proliferation index of hot spots was high (50%-60%). After surgery, acne improved and β-hCG returned to normal levels. Combined with histological morphology and immunohistochemistry, the final diagnosis was a secreting CNS germinoma.

The white blood cell and tumor tissue specimens of the patient were collected for NGS detection based on a panel including 539 cancer-related genes (Simceredx). By comprehensive genomic profiling, somatic *KIT* p.Y823D (AF 46.29%) and *SDHA* p.T396Nfs*14 (AF 48.35%) was detected but no second hits or loss of heterozygosity (LOH) in tumor tissue (**Figures 2A, B**). Copy number variations (CNV) including the gain of chromosome 7 and 21, and loss of chromosome 13 were identified (**Figure 2C**). Due to the younger age of diagnosis with a cafe-au-lait macule phenotype, we confirmed the germline mutation in leukocytes. No cafe-au-lait macule-associated tumor syndrome mutation was identified, but *SDHA* p. T396Nfs*14 was confirmed a heterozygous germline mutation. Further analysis by Sanger sequencing identified that the *SDHA* heterozygous mutation originated from his father with a normal phenotype (**Figures 2D, E**). The patient didn't have familial cancer history.

The patient received chemotherapy (temozolomide + nedaplatin + etoposide) and radiation therapy after surgery. Computed tomography scan after chemotherapy in of the brain showed a few effusions in the right subdural. Up to the article submission, there was no tumor recurrence (**Figures 1G–I**).

DISCUSSION AND CONCLUSION

In our case, we found concurrent germline *SDHA* and somatic *KIT* mutations, which was not reported in CNS GCTs. *SDHA* gene encodes a major subunit of Succinate dehydrogenase, a complex of the mitochondrial respiratory chain whose subunits are encoded by *SDHA*, *SDHB*, *SDHC*, *SDHD*, *SDHAF1*, and *SDHAF2*. As tumor suppressor genes, aberrations in the SDH complex genes result in the SDH complex deficiency (SCD) syndrome, an autosomal dominant disease with the occurrence of multiple tumors including gastrointestinal stromal tumors (GISTs), paragangliomas (PGLs), pheochromocytomas (PCCs), renal cell carcinomas (RCCs), and others (15). According to the paper by Korpershoek et al., the penetrance of *SDHA* - associated apparently sporadic paragangliomas and pheochromocytoma is low (16). Recently, germline mutations in SDHx were also reported in germ cell tumors including *SDHB* p.L157W in mediastinal germ cell tumor and *SDHD* p.W43* in testicular seminoma (17, 18). The patient has no evidence of PGLs/PCCs and his father carries the same *SDHA* mutation without tumor phenotype, which may be related to the incomplete penetrance of SCD syndrome.

The p. T396Nfs*14 (c.1186dup) mutation is located in exon 9 of the *SDHA* gene. The variant is predicted to result in a

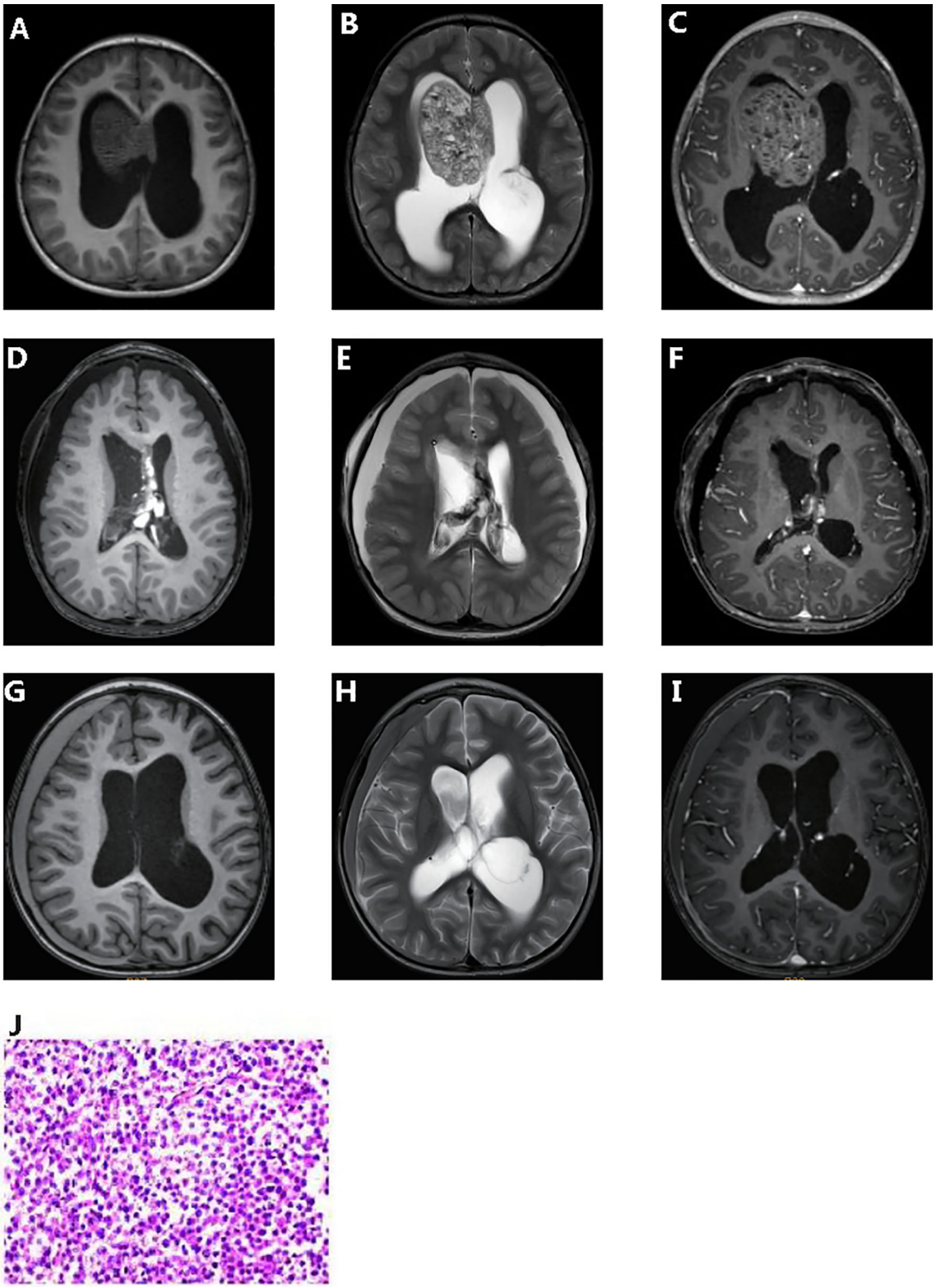


FIGURE 1 | Clinicopathologic features of patient. Pre-operative T1-weighted **(A)** and T2- weighted **(B)** dynamic contrast enhanced MRI **(C)**. Post-operative T1-weighted **(D)** and T2- weighted **(E)** dynamic contrast enhanced MRI **(F)**. 5 months post-operative T1-weighted **(G)** and T2- weighted **(H)** dynamic contrast enhanced MRI **(I)**. Hematoxylin and eosin **(J)**.

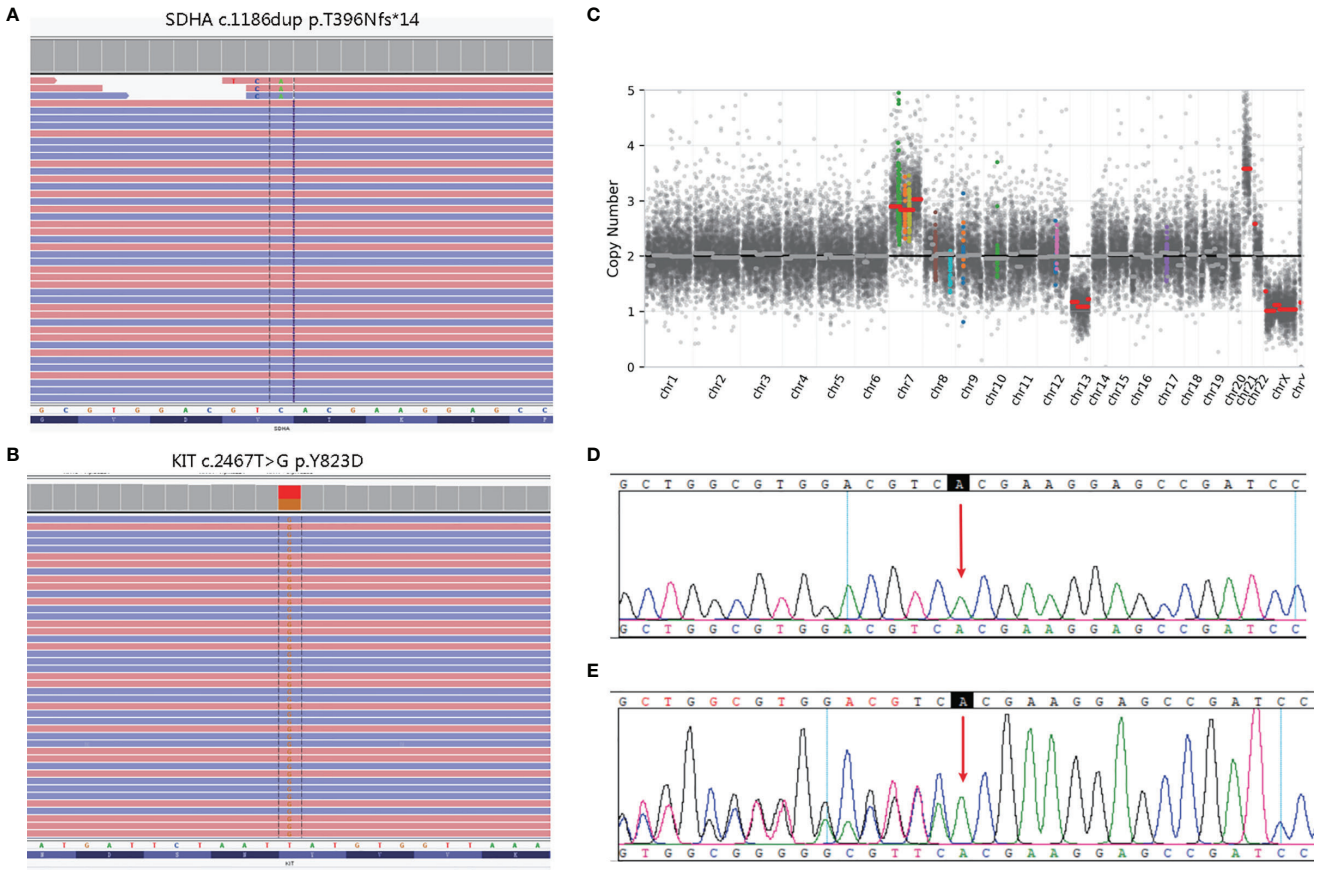


FIGURE 2 | Molecular pathology data of patient. Next generation sequencing (NGS) revealing a germline *SDHA* c.1186dup variant **(A)**, and a somatic *KIT* c.2467T>G variant **(B)**. Copy number variation (CNV) in the whole genome **(C)**. The unaffected father of the patient carried the heterozygous *SDHA* c.1186dup mutation **(E)**, whereas the healthy mother did not **(D)**.

premature stop codon at position 409 of protein likely causing an absence or loss of C-terminal domain protein product (19). According to the guideline of the American College of Medical Genetics and Genomics (ACMG), we defined the *SDHA* mutation as a pathogenic variant. *KIT* p.Y823D mutation lies in the protein kinase domain of the Kit protein conferring a gain of function on Kit and then activates the KIT-RAS-MAPK pathway (20). Although the current evidence is insufficient to establish a clinical relationship for GCTs and *SDHx* mutations. Considering the early onset of the disease in this patient, we speculate that *SDHA* germline mutation promotes a more aggressive behavior of GCTs related to the KIT-RAS-MAPK pathway potentially, by activating the pseudo-hypoxic pathway to increased ROS, angiogenesis, and cell proliferation in SDH-deficient tumors (21).

At present, surgery, radiotherapy, and chemotherapy can all improve the outcomes of CNS GCTs. The genomic analysis could provide the application of TKIs in in CNS GCTs. In CNS GCTs, *KIT* mutations were mainly concentrated in exon 17, followed by exon 11. *KIT* p.Y823D in GCT patients was reported to be sensitive to sorafenib and resistant to imatinib and Sunitinib (14). Tyrosine kinase inhibitors (TKIs) targeting KIT have been approved in other tumors. *KIT* mutations and *SDHx* germline mutations are also common in gastrointestinal stromal tumors, and distinct from our case, they are usually mutually exclusive (22). The possible synergistic or antagonistic effects of the co-mutation on the drug are unknown.

In conclusion, we for the first time report a 4-year-old Chinese male child with CNS GCT harboring concurrent germline *SDHA* and somatic *KIT* mutation, which showed the necessity to detect germline mutations in CNS GCTs, especially in children with CNS GCTs. In addition, the relationship between *SDHA* germline mutations and the clinical diagnosis, prognosis, and therapy of CNS GCTs needs to be further investigated.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

Written informed consent was obtained from the patient for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

GW designed the study. XY, BL drafted the manuscript. TH participated in the manuscript. TH, NL analyzed the literature. NL, GL revised the manuscript. DG, FB participated in the revision. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors thank Ms. Xiaomin Li, Mr. Wanglong Deng, and Mr. Ran Ding from Simceredx for the kindly assistance.

REFERENCES

1. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: A Summary. *Neuro Oncol* (2021) 23(8):1231–51. doi: 10.1093/neuonc/noab106
2. Packer RJ, Cohen BH, Cooney K. Intracranial Germ Cell Tumors. *Oncologist* (2000) 5:312–20. doi: 10.1634/theoncologist.2000-0312
3. McCarthy BJ, Shibui S, Kayama T, Miyaoka E, Narita Y, Murakami M, et al. Primary CNS Germ Cell Tumors in Japan and the United States: An Analysis of 4 Tumor Registries. *Neuro Oncol* (2012) 14:1194–200. doi: 10.1093/neuonc/nos155
4. Shibui S. Report of Brain Tumor Registry of Japan (2001–2004). *Neurol Medico-Chirurgica* (2014) 54(suppl.1):9–102. doi: 10.2176/nmc.sup.2014-0001
5. Gittleman H, Cioffi G, Vecchione-Koval T, Ostrom QT, Kruchko C, Osorio DS, et al. Descriptive Epidemiology of Germ Cell Tumors of the Central Nervous System Diagnosed in the United States From 2006 to 2015. *J Neuro-Oncol* (2019) 143(2):251–60. doi: 10.1007/s11060-019-03173-4
6. Calaminus G, Frappaz D, Kortmann RD, Krefeld B, Saran F, Pietsch T, et al. Outcome of Patients With Intracranial Non-Germinomatous Germ Cell Tumors-Lessons From the SIOP-CNS-GCT-96 Trial. *Neuro Oncol* (2017) 19(12):1661–72. doi: 10.1093/neuonc/nox122
7. Takami H, Fukuoka K, Fukushima S, Nakamura T, Mukasa A, Saito N, et al. Integrated Clinical, Histopathological, and Molecular Data Analysis of 190 Central Nervous System Germ Cell Tumors From the iGCT Consortium. *Neuro-Oncology* (2019) 21:12: 1565–77. doi: 10.1093/neuonc/noz139
8. Murray MJ, Bartels U, Nishikawa R, Fangusaro J, Matsutani M, Nicholson JC. Consensus on the Management of Intracranial Germ-Cell Tumours. *Lancet Oncol* (2015) 16:9: e470–7. doi: 10.1016/S1470-2045(15)00244-2
9. Fang LX, Zhu MH, Xu SX, Li ZY, Qi S. Diagnostic Analysis of 162 Casesof Primary Intracranial Germ Cell Tumors (in Chinese). *Chin J Neurosurg* (2014) 30:541–4. doi: 10.3760/cma.j.issn.1001-2346.2014.06.001
10. Zhou ZG, Yan HL, Chen BG, Hu ZP, Deng LT, Zhou XH, et al. 125cases of Diagnostic Evaluation of Primary Intracranial Germ Cell Tumors (in Chinese). *J Clin Surg* (2017) 25:658–60. doi: 10.3969/j.issn.1005-6483.2017.09.006
11. Sheikine Y, Genega E, Melamed J, Lee P, Reuter VE, Ye H. Molecular Genetics of Testicular Germ Cell Tumors. *Am J Cancer Res* (2012) 2(2):153–67.
12. Reuter VE. Origins and Molecular Biology of Testicular Germ Cell Tumors. *Mod Pathol* (2005) 18 Suppl 2:S51–60. doi: 10.1038/modpathol.3800309
13. Sukov WR, Cheville JC, Giannini C, Carlson AW, Shearer BM, Sinnwell JP, et al. Isochromosome 12p and Polysomy 12 in Primary Central Nervous System Germ Cell Tumors: Frequency and Association With Clinicopathologic Features. *Hum Pathol* (2010) 41(2):232–8. doi: 10.1016/j.humpath.2009.07.017
14. Wang L, Yamaguchi S, Burstein MD, Terashima K, Chang K, Ng HK, et al. Novel Somatic and Germline Mutations in Intracranial Germ Cell Tumours. *Nature* (2014) 511(7508):241–5. doi: 10.1038/nature13296
15. Hensen EF, Bayley JP. Recent Advances in the Genetics of SDH-Related Paraganglioma and Pheochromocytoma. *Familial Cancer* (2011) 10(2):355–63. doi: 10.1007/s10689-010-9402-1
16. Korpershoek E, Favier J, Gaal J, Burnichon N, van Gessel B, Oudijk L, et al. SDHA Immunohistochemistry Detects Germline SDHA Gene Mutations in Apparently Sporadic Paragangliomas and Pheochromocytomas. *J Clin Endocrinol Metab* (2011) 96(9):E1472–6. doi: 10.1210/jc.2011-1043
17. De Filpo G, Cilotti A, Rolli L, Pastorino U, Sonzogni A, Pradella S, et al. SDHx and Non-Chromaffin Tumors: A Mediastinal Germ Cell Tumor Occurring in a Young Man With Germline SDHB Mutation. *Medicina* (2020) 56(11):561. doi: 10.3390/medicina56110561
18. Galera-Ruiz H, Gonzalez-Campora R, Rey-Barrera M, Rollón-Mayordomo A, Garcia-Escudero A, Fernández-Santos JM, et al. W43X SDHD Mutation in Sporadic Head and Neck Paraganglioma. *Anal Quant Cytol Histol* (2008) 30 (2):119–23.
19. Renkema GH, Wortmann SB, Smeets RJ, Venselaar H, Antoine M, Visser G, et al. SDHA Mutations Causing a Multisystem Mitochondrial Disease: Novel Mutations and Genetic Overlap With Hereditary Tumors. *Eur J Hum Genet* (2015) 23(2):202–9. doi: 10.1038/ejhg.2014.80
20. Caenepeel S, Renshaw-Gegg L, Baher A, Bush TL, Baron W, Juan T, et al. Motesanib Inhibits Kit Mutations Associated With Gastrointestinal Stromal Tumors. *J Exp Clin Cancer Res* (2010) 29(1):1–8. doi: 10.1186/1756-9966-29-96
21. Vicha A, Taieb D, Pacak K. Current Views on Cell Metabolism in SDHx-Related Pheochromocytoma and Paraganglioma. *Endocrine-Related Cancer* (2014) 21(3):R261. doi: 10.1530/ERC-13-0398
22. Corless CL, Barnett CM, Heinrich MC. Gastrointestinal Stromal Tumours: Origin and Molecular Oncology. *Nat Rev Cancer* (2011) 11(12):865–78. doi: 10.1038/nrc3143

Conflict of Interest: Authors NL, TH, DG, GL, and FB are employed by the company Jiangsu Simcere Diagnostics Co., Ltd.

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.

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