frontiers | Research Topics

Genomic alteration landscapes of aging, metabolic disorders, and cancer: Emerging overlaps and clinical importance

Edited by
Rajkumar S. Kalra, Amrendra K. Ajay, Dhanendra Tomar and Jaspreet Kaur Dhanjal

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-83251-383-5
DOI 10.3389/978-2-83251-383-5

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

Genomic alteration landscapes of aging, metabolic disorders, and cancer: Emerging overlaps and clinical importance

Topic editors

Rajkumar S. Kalra — Okinawa Institute of Science and Technology Graduate University, Japan

Amrendra K. Ajay — Brigham and Women's Hospital, Harvard Medical School, United States

Dhanendra Tomar — Wake Forest University, United States

Jaspreet Kaur Dhanjal — Indraprastha Institute of Information Technology Delhi, India

Citation

Kalra, R. S., Ajay, A. K., Tomar, D., Dhanjal, J. K., eds. (2023). *Genomic alteration landscapes of aging, metabolic disorders, and cancer: Emerging overlaps and clinical importance*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-83251-383-5

Table of contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 06 January 2023
DOI 10.3389/fgene.2022.1102953

OPEN ACCESS

EDITED AND REVIEWED BY
Maxim B. Freidin,
Queen Mary University of London,
United Kingdom

*CORRESPONDENCE
Jaspreet Kaur Dhanjal,
jaspreet@iiitd.ac.in

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 19 November 2022
ACCEPTED 23 November 2022
PUBLISHED 06 January 2023

CITATION
Dhanjal JK, Kalra RS, Tomar D and Ajay AK (2023), Editorial: Genomic alteration landscapes of aging, metabolic disorders, and cancer: Emerging overlaps and clinical importance.
Front. Genet. 13:1102953.
doi: 10.3389/fgene.2022.1102953

Editorial: Genomic alteration landscapes of aging, metabolic disorders, and cancer: Emerging overlaps and clinical importance

Jaspreet Kaur Dhanjal[1]*, Rajkumar Singh Kalra[2], Dhanendra Tomar[3] and Amrendra K. Ajay[4,5]

[1]Department of Computational Biology, Indraprastha Institute of Information Technology Delhi, New Delhi, India, [2]Immune Signal Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa, Japan, [3]Department of Internal Medicine, Section of Cardiovascular Medicine, Section of Molecular Medicine, Wake Forest University School of Medicine, Winston-Salem, NC, United States, [4]Department of Medicine, Harvard Medical School, Boston, MA, United States, [5]Division of Renal Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA, United States

KEYWORDS

genetic alterations, cancer, aging, metabolic disorder, mutations, senescence

Editorial on the Research Topic
Genomic alteration landscapes of aging, metabolic disorders, and cancer: Emerging overlaps and clinical importance

The biology of aging, cancer, and various metabolic disorders shows a clear association with genetic and epigenetic changes. These genomic alterations arise from diverse intrinsic and extrinsic/environmental factors. The efficiency of a cell to proofread its newly synthesized DNA strand gradually decreases with age hampering its genomic integrity. An increased burden of genomic changes, therefore, gives rise to multiple health issues like metabolic disorders (Abou Ziki and Mani, 2016; Varshavi et al., 2018). However, on the other hand, recent studies provide evidence for the role of metabolic perturbations in accelerated aging (Spinelli et al., 2020). These transformations, following either way, involve diverse interactions between molecular players of aging, metabolism, and redox biology (including mitochondria fitness, Ca^{2+} signaling, and bioenergetics); all encrypted in the genomic sequence. Accumulation of irreversible genomic changes over a long time then leads to the onset and progression of cancer. Cancer cells have been shown to operate with reengineered metabolic processes to satisfy their surplus needs during uncontrolled proliferation (Liu et al., 2022). Therefore, aging, metabolic changes, and cancer exist as a network of crossroads (Tidwell et al., 2017; Golubev and Anisimov, 2019; Poljsak et al., 2019). These broadly categorized pathologies share common genomic signatures that further strengthen the link between aging, metabolic disorders, and cancer (Aunan et al., 2017; Lacroix et al., 2020). Along with metabolic alterations, occurrence of aberrant mutations in the mitochondrial genome is also a common characteristic of aging and cancer (Smith

et al., 2022). Therefore, it becomes important to uncover the contribution of genomic changes in the context of these cellular health states and the sequential order that defines these states, if any. In line with this, the original research articles and reviews published in the present Research Topic focus on genomic alteration landscapes of aging, metabolic disorders and cancer, the existing and emerging overlaps, and its clinical importance for therapeutic interventions (Figure 1).

The role of autophagy has been recently established in the pathogenesis of aortic dissection; however, the complete molecular mechanism has not been uncovered yet. In an original report, Huang et al. has studied the correlation between the family of ULK (UNC51-like enzymes) genes and the age of first onset of type B aortic dissection (TBAD). The authors analyzed the genome of 159 TBAD patients from Chinese population. A pool of 1,180,097 SNPs was included. Among the different ULK genes, only ULK4 was found to be significantly associated with the first onset age. They concluded that high level of ULK4 gene expression was related to delayed onset of TBAD among these patients. Further experimental validation of these findings can suggest ULK4 to be a diagnostic target for TBAD.

Telomere shortening is one of the important hallmarks of cellular senescence, however, telomere length related cellular senescence has been shown to have varying effects in different cancers. To delineate this paradoxical relationship, Son et al. made use of 42 telomere length associated SNPs, and performed Mendelian randomization analysis to explore the causal relationship between telomere length, skin aging and the susceptibility risk of different skin cancer types. The authors found that telomere shortening can promote aging of the skin and reduce the risk of cutaneous melanoma and non-melanoma skin cancer.

As mentioned before, cellular senescence has often been correlated to oncogenic activation and tumor suppression leading to cancer development. In line with this, Dai et al. has constructed a prognostic risk score signature using the senescence related genes differentially expressed in gastric cancer samples. A total of 135 such genes were identified with significant dys-regulation. Integration of survival data associated 24 of these genes with gastric cancer prognosis. Patients with high expression of SMARCA4 (gene with highest mutational frequency) were associated with higher overall survival and progression-free survival. A total of 11 genes were then identified using LASSO Cox regression analysis to develop the prognostic risk score signature. Testing using an independent data showed that this signature could accurately distinguish low-risk and high-risk samples. The authors further showed that the low-risk score group was also more susceptible to chemotherapy and immunotherapy, and hence can be used for better decision making for treatment to be given. Another study by Dai et al. on similar grounds looked into the cellular senescence related genes that can be used for prediction of prognosis and immunotherapy response in colon cancer patients. Dong et al. and Sun et al. also reported a senescence-related prognostic model that has been shown to predict the prognosis, immunotherapeutic response, and identify potential drug targets for colorectal and hepatocellular carcinoma patients, respectively. These studies show the potential of using huge amount of publicly available clinical data for learning and developing predictive models to design personalized treatment regimen for cancer patients.

Intracellular calcium levels play an important role in homeostasis and various cell signalling processes. Dysregulated levels of calcium have been shown to be remarkably associated with cancer growth, angiogenesis, and metastasis. Elevated serum calcium level is a proposed diagnostic marker for head and neck malignancy. In association with this, Hegde et al. carried out *in silico* analysis to demonstrate the role of store-operated calcium channels in regular mitochondrial function, and further suggest that alteration in these calcium channels might be a predictive and prognostic marker for head and neck squamous cell cancer patients.

Though cancer in general is thought to arise from accumulation of somatic mutations, they do have a substantial hereditary component. To look at the contribution of the pathogenic germline variants in the development of bowel cancer in Chinese population, Xie et al. analyzed the mutation profile of 573 patients accounting for various stages of bowel/colorectal cancer. The profiled germline mutations were categorized as pathogenic, likely-pathogenic and non-pathogenic. Some rare germline alterations in genes like ANCD2, CDH1, and FLCN were also observed. The other germline mutations were enriched in genes involved in DNA-damage repair and homologous recombination. Patients carrying germline mutations also showed a distinctive somatic mutation profile and tumor mutation burden, which also affected the overall survival of these patients. This study provides an assessment of a wider range of susceptibility genes in Chinese bowel/colorectal patients.

Along with accumulating genetic mutations, cancer cells also reprogram the other biochemical processes to generate conditions favorable for sustenance and continuous proliferation. Metabolic reprogramming to switch from oxidative phosphorylation to aerobic glycolysis is one of the major hallmarks of cancer. In this Research Topic of articles, Sharma et al. has presented a comprehensive review highlighting the role of hypoxia-inducible factor-1 (HIF-1) in imparting aggressive behavior in cancer cells through hypoxic glycolysis, and novel therapeutic strategies currently available for targeting HIF-1 in cancer.

Not only gene expression, but its regulation by non-coding RNAs like miRNA also plays a crucial role in onset and progression of various diseases. A review article from Chhichholiya et al. gives information about the reported single nucleotide polymorphisms in miRNA(s) and their target sequences known to be involved in cancer and diabetic pathologies.

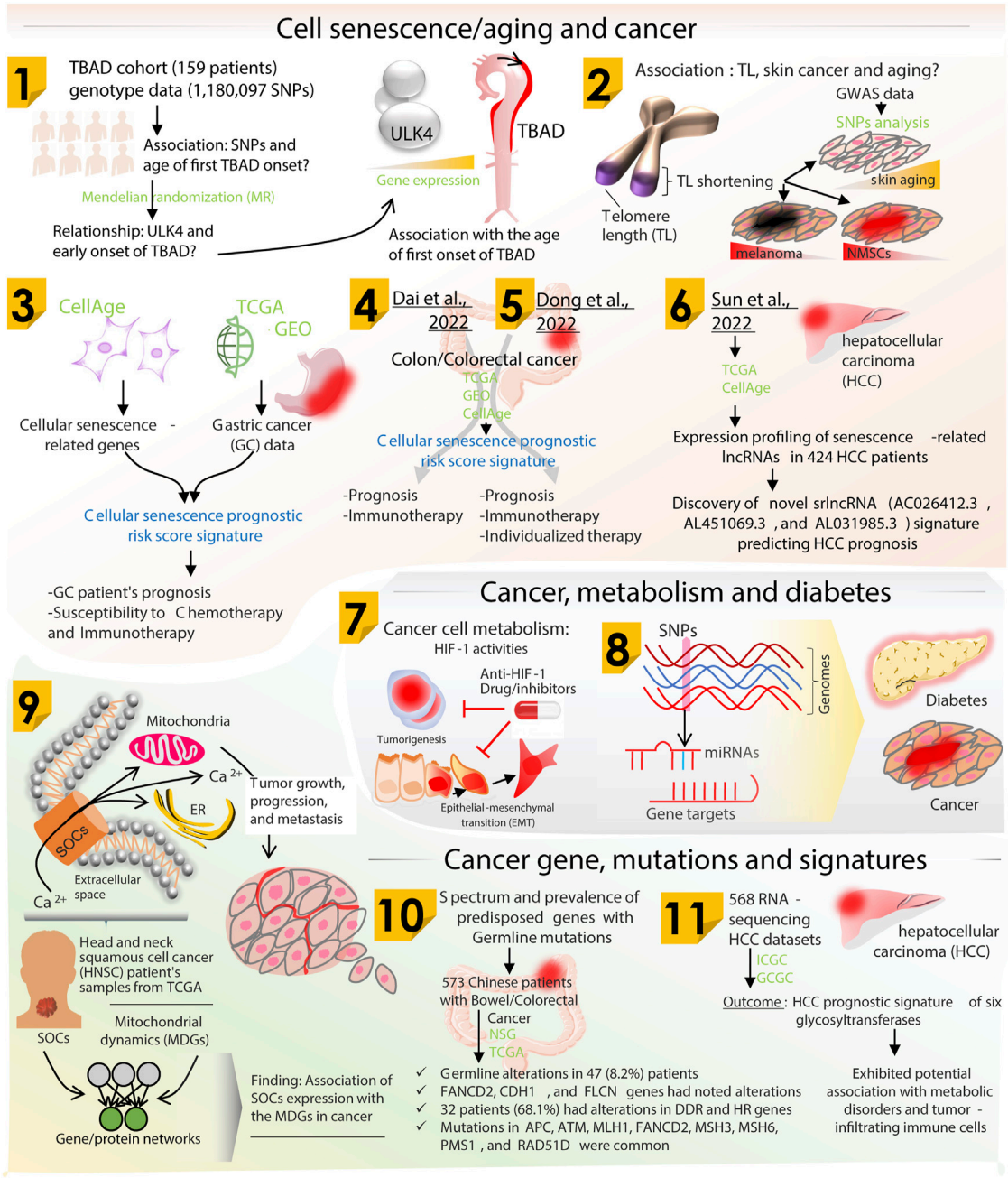


FIGURE 1
Schematic diagram showing research summarized majorly on three themes including cell senescence/aging and cancer, metabolism and cancer gene mutations/signatures. Research Topic articles are marked by serial number within the related schematic section.

In summary, the present Research Topic gathers original research and comprehensive reviews highlighting the genomic factors behind aging, metabolic perturbations and cancer. These studies confirm the interconnecting links between these pathologies, and the need to understand these to identify the cross-points that can be further explored for diagnosis, prognosis and other therapeutic interventions.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

DT is supported in part by National Institutes of Health (NIH) grant R00DK120876. AKA is supported by a Career Development Grant from the American Heart Association award number 19CDA34780005.

Acknowledgments

We thank all the contributing authors and reviewers for their valuable contributions to this Research Topic.

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.

References

Abou Ziki, M. D., and Mani, A. (2016). Metabolic syndrome: Genetic insights into disease pathogenesis. *Curr. Opin. Lipidol.* 27, 162–171. doi:10.1097/MOL.0000000000000276

Aunan, J. R., Cho, W. C., and Soreide, K. (2017). The biology of aging and cancer: A brief overview of shared and divergent molecular hallmarks. *Aging Dis.* 8, 628–642. doi:10.14336/AD.2017.0103

Golubev, A. G., and Anisimov, V. N. (2019). Aging and cancer: Is glucose a mediator between them? *Oncotarget* 10, 6758–6767. doi:10.18632/oncotarget.27344

Lacroix, M., Riscal, R., Arena, G., Linares, L. K., and Le Cam, L. (2020). Metabolic functions of the tumor suppressor p53: Implications in normal physiology, metabolic disorders, and cancer. *Mol. Metab.* 33, 2–22. doi:10.1016/j.molmet.2019.10.002

Liu, Y. H., Hu, C. M., Hsu, Y. S., and Lee, W. H. (2022). Interplays of glucose metabolism and KRAS mutation in pancreatic ductal adenocarcinoma. *Cell Death Dis.* 13, 817. doi:10.1038/s41419-022-05259-w

Poljsak, B., Kovac, V., Dahmane, R., Levec, T., and Starc, A. (2019). Cancer etiology: A metabolic disease originating from life's major evolutionary transition? *Oxid. Med. Cell. Longev.* 2019, 7831952. doi:10.1155/2019/7831952

Smith, A. L. M., Whitehall, J. C., and Greaves, L. C. (2022). Mitochondrial DNA mutations in ageing and cancer. *Mol. Oncol.* 16, 3276–3294. doi:10.1002/1878-0261.13291

Spinelli, R., Parrillo, L., Longo, M., Florese, P., Desiderio, A., Zatterale, F., et al. (2020). Molecular basis of ageing in chronic metabolic diseases. *J. Endocrinol. Invest.*. 43, 1373–1389. doi:10.1007/s40618-020-01255-z

Tidwell, T. R., Soreide, K., and Hagland, H. R. (2017). Aging, metabolism, and cancer development: From peto's paradox to the warburg effect. *Aging Dis.* 8, 662–676. doi:10.14336/AD.2017.0713

Varshavi, D., Scott, F. H., Varshavi, D., Veeravalli, S., Phillips, I. R., Veselkov, K., et al. (2018). Metabolic biomarkers of ageing in C57bl/6J wild-type and flavin-containing monooxygenase 5 (FMO5)-Knockout mice. *Front. Mol. Biosci.* 5, 28. doi:10.3389/fmolb.2018.00028

frontiers in Genetics

REVIEW
published: 01 December 2021
doi: 10.3389/fgene.2021.793523

SNPs in miRNAs and Target Sequences: Role in Cancer and Diabetes

Yogita Chhichholiya, Aman Kumar Suryan, Prabhat Suman, Anjana Munshi and Sandeep Singh**

Department of Human Genetics and Molecular Medicine, Central University of Punjab, Bathinda, India

miRNAs are fascinating molecular players for gene regulation as individual miRNA can control multiple targets and a single target can be regulated by multiple miRNAs. Loss of miRNA regulated gene expression is often reported to be implicated in various human diseases like diabetes and cancer. Recently, geneticists across the world started reporting single nucleotide polymorphism (SNPs) in seed sequences of miRNAs. Similarly, SNPs are also reported in various target sequences of these miRNAs. Both the scenarios lead to dysregulated gene expression which may result in the progression of diseases. In the present paper, we explore SNPs in various miRNAs and their target sequences reported in various human cancers as well as diabetes. Similarly, we also present evidence of these mutations in various other human diseases.

Keywords: miRNA, microRNA, target genes, seed sequences, SNPs, cancer, diabetes mellitus

OPEN ACCESS

Edited by:
Rajkumar S Kalra,
Okinawa Institute of Science and Technology Graduate University, Japan

Reviewed by:
Choudhary Harsha,
Indian Institute of Technology Guwahati, India
Alok Kumar,
Kyoto University, Japan

****Correspondence:***
Anjana Munshi
anjana.munshi@cup.edu.in
Sandeep Singh
sandeepsingh82@cup.edu.in

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *12 October 2021*
Accepted: *28 October 2021*
Published: *01 December 2021*

Citation:
Chhichholiya Y, Suryan AK, Suman P, Munshi A and Singh S (2021) SNPs in miRNAs and Target Sequences: Role in Cancer and Diabetes. Front. Genet. 12:793523. doi: 10.3389/fgene.2021.793523

INTRODUCTION

MicroRNAs (miRNAs) are endogenous single stranded, non-coding, 20–22 nucleotides long molecules that are processed from pre-miRNA. miRNAs have been demonstrated to be tremendously versatile in their function. miRNAs have significant roles in the nucleus as well as cytoplasm in terms of controlling gene expression. They play a significant role in post-transcriptional regulation of gene expression either via translational repression or mRNA degradation (Iorio and Croce, 2012; Peng and Croce, 2016). miRNAs recognize targets by specific base-pairing complementarity between their seed sequence of miRNA (5′ end) and untranslated region (3′UTR) of target gene/mRNA (Ling et al., 2011; Si et al., 2019).

However, in some exceptional cases, base pairing is also reported between 5′ UTR region of the specific mRNA and coding regions (O'Brien et al., 2018; Valinezhad Orang et al., 2014). The standard size of 3′UTR in the human gene is about 950 nucleotides whereas the seed sequence of miRNA is around 6 to 8 nucleotides. The 3′UTR region of a particular mRNA may be recognized by a specific miRNA or by multiple miRNAs. Sequence complementarity is shared by miRNAs with respect to their mRNA targets, resulting in the interaction of a single miRNA with many genes whereas a single gene can probably be regulated by multiple miRNAs (Hashimoto et al., 2013; Mariella et al., 2019).

Around 10 million SNPs are known to be present in both coding as well as non-coding regions of the human genome at a frequency of one in every 300 bp (Moszyńska et al., 2017). Since SNPs have also been reported to be present in seed sequence, it is most likely that the presence of these alterations might disrupt or create new interaction of miRNA with its target site (Palmero et al., 2011; Bhattacharya and Cui, 2017). Furthermore, the SNPs in the 3′UTRs of gene/mRNA can also modulate miRNA-mRNA interactions, protein-mRNA interactions, polyadenylation, all of which might have a serious impact on translation efficiency and mRNA stability (Malhotra et al., 2019a).

This in turn might result in the development of various diseases including neurodevelopment disorders, cardiovascular diseases, cancer, autoimmune diseases, and many more (Bruno et al., 2012; Moszyńska et al., 2017).

Cancer and diabetes are multifactorial life threatening human diseases in which various miRNAs have been reported in the pathogenesis as well as the severity of these diseases (Ayaz Durrani et al., 2021). Tens of millions of people are diagnosed with cancer each year around the world, with more than half of those diagnosed dying from it. miRNA profiling and high throughput sequencing in the recent past revealed that miRNA expression is dysregulated in cancer and that its fingerprints might be utilized to classify, diagnose, and prognosis of tumors. miRNAs have been reported to act as oncogenes or tumor suppressors under certain biological conditions. Cancer hallmarks such as sustaining proliferative signals, evading growth suppressors, resisting cell death, activating invasion and metastasis, and initiating angiogenesis have been linked to dysregulated miRNAs (Peng and Croce, 2016).

Diabetes mellitus (DM) affects 347 million people worldwide. Diabetes-related fatalities are expected to double between 2005 and 2030, according to the World Health Organization (Chen et al., 2014a). High blood glucose levels are a defining feature of DM. Diabetes is classified into two types. A deficiency of insulin synthesis in pancreatic cells causes type 1 diabetes (T1D), whereas type 2 diabetes (T2D) is caused by insulin resistance, which causes the body to utilize insulin inefficiently. In both T1D and T2D, long-term hyperglycemia can cause macrovascular (coronary artery disease, peripheral arterial disease, and stroke) as well as microvascular complications (diabetic nephropathy, neuropathy, and retinopathy) (Fowler, 2008). miRNAs are implicated in the etiology and pathogenesis of diabetes and associated complications (Chen et al., 2014a). However, the role of miRNAs in diabetes and its complications are comparatively explored less.

Very few reports are available on SNPs reported in the seed sequence of miRNAs and the 3′UTR region of their target genes. The current review has been compiled with an aim to evaluate the role of genetic variation in the seed sequence of the miRNA and the 3′UTR of their specific target genes in association with the development of the two most common prevalent diseases—cancer and diabetes.

SNPS IN THE SEED SEQUENCE OF MIRNA AND THE 3′UTR OF SPECIFIC TARGET GENE IN CANCER

Many miRNAs have been discovered to play a role in the genesis of cancer, either as tumor suppressor genes or as oncogenes. The study of tumor-specific miRNA expression profiles in a variety of malignancies revealed extensive dysregulation of these molecules, some of the overexpression and underexpression of various miRNA (Song and Chen, 2011). As evidenced by multiple findings demonstrating the importance of miRNAs in carcinogenesis, miRNA dysregulation leads to modulation of tumor cell signaling, changes in DNA repair or stress response, and function of the effector protein (Moszyńska et al., 2017; Malhotra et al., 2019a; Galka-Marciniak et al., 2019).

SNPs in the Seed Sequence of miRNA

SNPs in the mature or primary miRNA or seed sequence might affect miRNA processing or binding. Several SNPs present in miRNA main sequences or upstream regulatory regions have been linked to increased cancer risk as well as its prognosis (Duan et al., 2007). In cancer, SNPs reported in seed sequence of various miRNA include rs2910164 in miR-146a; rs3746444 in miR-499; rs12975333 in miR-125a; rs34059726 in miR-124; and rs11614913 in miR-196-a2. Information about SNPs in the seed sequence of miRNAs is summed up in **Table 1** and elaborated functional role of these miRNAs in the development of cancer has been depicted in **Figure 1**.

miRNA-146a

miR-146a is a widely expressed miRNA in mammalian cells. Multiple studies have shown that miR-146a is involved in inflammation, differentiation, and function of adaptive and innate immune cells. miR-146a has been found to be a regulator of cell function and differentiation in innate and adaptive immunity. A subset of human T cells exhibits different expression level of niR-146a. Memory T cells and naive T lymphocytes have different levels of miR-146a expression (Nahand et al., 2020). This miRNA is produced by T-cell receptor activation, and the binding sites of c-ETS and transcription factor nuclear factor-κB (NF-kB) are required for miR-146a transcription in T lymphocyte cells (Curtale et al., 2010; Lu et al., 2010). Some studies observed an association between the NF-kB signaling pathway and miR-146a expression (Rusca and Monticelli, 2011). Taganov et al. (2006) found that LPS stimulation enhanced miR-146a expression in an NF-kB-dependent manner and that miR-146a targeted the IRAK1 and TRAF6 genes (Taganov et al., 2006; Rusca and Monticelli, 2011). After a cell surface receptor (such as TLR4) is activated, a biochemical cascade involving IRAK1 and TRAF6 causes IkBa to be phosphorylated and degraded, resulting in the activation of NF-kB and its nuclear translocation. Furthermore, NF-kB activation causes some genes, such as pri-miR-146a, to be expressed. miR-146a matures on the RISC and contributes to the attenuation of receptor signaling by downregulating TRAF6 and IRAK1. As a result, miR-146a inhibits the signaling pathway leading to NF-B activation (Taganov et al., 2006; Taganov et al., 2007; Labbaye and Testa, 2012).

miR-146a polymorphism, rs2910164, involves a G>C nucleotide alteration on the seed region of miR146a-3p, resulting in G:U pair to a C:U mismatch pairing in the stem of the miR-146a affecting the specificity of mature miR-146a binding to its targets and results in elevated expression of miR-146a (Brincas et al., 2020). Previous studies have established the association of rs2910164 in pre-miR-146a with strong association with breast cancer (BC), hepatocellular carcinoma (HCC), papillary thyroid carcinoma (PTC), esophageal squamous cell carcinoma (ESCC), primary liver cancer, and colorectal cancer (Hu et al., 2009; Vitale et al., 2011; Xiang et al., 2012; Zhou et al., 2012).

TABLE 1 | SNPs reported in seed sequence of miRNA in various cancers.

S. No	miRNA	Gene	SNP reported	Tumor type	Reference
1	miR-379	SEMA3F	rs61991156 A>G	Gastric cancer	Cao et al. (2016)
2	miR-627	SEMA3F	rs2620381 A>C	ESCC- esophageal squamous cell carcinoma	Cao et al. (2016)
3	miR-499-3p	PBX1, FOXO1A	rs3746444 A>G	BC, ALL, colorectal, liver, SCC of head and neck, gallbladder cancer	Xiang et al. (2012); Cao et al. (2016); Ahmad et al. (2019)
4	miR-124	VAMP3, CD164, PTPN12, ITGB1	rs34059726 G>T	Lung cancer	Yang et al. (2016); Fawzy et al. (2017)
5	miR-642a	ATP6VOE1	rs78902025 T>G	Leiomyoma	Pelletier et al. (2011); Cao et al. (2016)
6	miR-4293	SLC43A2	rs12220909 A>C	NSCLC	Malhotra et al. (2019b)
7	miR-146a-3p	BRCA1, TRAF6, IRAK1, NUMB	rs2910164 C>G	BC, HCC, PTC, ESCC, colorectal and primary liver cancer	Yekta et al. (2004); Vitale et al. (2011); Xiang et al. (2012); Zhou et al. (2012); Iguchi et al. (2016)
8	MiR-4707	CARD10	rs2273626 C>A	Rectal cancer	Shvarts et al. (1996); Erturk et al. (2014)
9	miR-4707	HAUS4	Rs2273626 C>A	Rectal cancer	Marine et al. (2006); Erturk et al. (2014)
10	miR-125a	Lin-28, lin-41, ERBB2, ERBB3	rs12975333 G>A	Breast cancer	Jin et al. (2008); Tian et al. (2009); Ristori et al. (2017)
11	miR-662	ATPVOE1	rs9745376 G>A/C/T	Ovarian cancer	Jin et al. (2008)
12	miR-196-a2	HOX, ANXA1	rs11614913 C>T	HCC, BC, NSCLC, bladder, renal, gastric, lung, glioma, head and neck, squamous cancer	Solito et al. (2001); Dou et al. (2010); Nicoloso et al. (2010); Ferracin et al. (2011); Peng and Croce (2016); Malhotra et al. (2019a); Zhao, (2020)
13	miR-585	SLIT3	rs62376935 C>T	Pharynx squamous cell carcinoma	Pelletier et al. (2011); Fawzy et al. (2017)
14	miR-605	PRKGI	rs113212828 A>G	Ovarian cancer	Pelletier et al. (2011)
15	miR-367-5p	LARP7	rs150161032A > G	Testicular germ cell tumor, leukemia	Strachan et al. (2003); Matijasevic et al. (2008)
16	miR-627	ATP6VOE1	rs2620381 A>C	Ovarian cancer, osteoblastoma	Jin et al. (2008); Cao et al. (2016)
17	miR-3161	PTPRJ	rs11382316 -/A	Colorectal cancer	Fawzy et al. (2017)

A microarray-based expression study carried out in a Chinese population found that miR-146a was significantly upregulated in breast carcinoma tissues compared to normal tissues (Omrani et al., 2014). Its expression level was three times higher in triple negative tumors in comparison with other tumor subgroups. However, this association has been reported to vary in different ethnic groups. The allele C is associated with BC risk in the European population but did not show any association with BC in the Asian population. This discrepancy might be on account of ethnicity, different exposure to carcinogens, or linkage disequilibrium with different causal variants (Brincas et al., 2020).

Molecular targets of miR-146a include BRCA1, TRAF6, IRAK1, and NUMB genes (Brincas et al., 2020). The variant allele of rs2910164 leads to increased levels of mature miR-146a that binds with greater affinity to the BRCA1 gene. Alternatively, rs2910146 might disrupt the well-documented role of miR-146a as a mediator of the pro-apoptotic transcriptional factor NF-κB. Also, the expression levels of miR146a-5p were observed three times higher in triple negative tumors compared to other subgroups of mammary tumors (Brincas et al., 2020). Two other significant targets of miR-146a, TRAF6, and IRAK1 are important adapter molecules downstream of the toll-like and cytokine receptors that have a vital role in signaling cell growth and immune recognition (Omrani et al., 2014). Both the genes have been associated with progression and metastasis. The reduced TRAF6 and IRAK1 levels reduce the activity of NF-kB, a potential inducer of proliferation, survival, angiogenesis, and metastasis (Brincas et al., 2020).

rs2910164 of miR-146a has been reported to induce liver metastasis in colorectal cancer (CRC) via Notch and JAK/STAT signaling pathways. Migratory response of NUMB has been observed in CRC cell lines (RKO, HT29, LoVo). NUMB protein is a negative regulator of Notch signaling, miR-146a activates Notch and JAK/STAT3 signaling through suppression of NUMB protein thereby enhancing the metastatic risk. Further, patients of gastric cancer bearing the altered genotype showed a higher expression of miR-146a than the ones bearing the normal genotype (Iguchi et al., 2016). In addition, this polymorphism has also been reported to increase the risk of PTC in a heterozygous condition (Vitale et al., 2011). The reduction in expression level of miR-146a led to less efficient inhibition of target genes TRAF6 and IRAF1 involved in the Toll-like receptor and

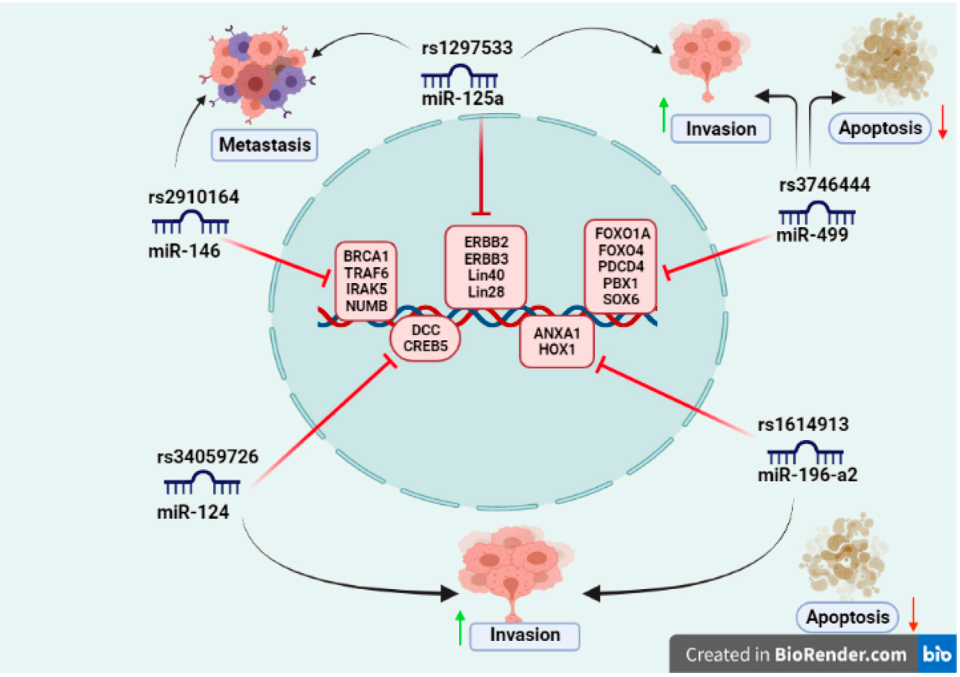


FIGURE 1 | SNPs reported in seed sequence of miRNA involved in cancer: miR-146 (rs2910164) targets include BRCA1, TRAF6, IRAKS, and NUMB gene and associated with cancer metastasis; miR-125a (rs1297533) found to be involved in invasion and metastasis and known targets include ERBB2, ERBB3, lin-40, and lin-28 gene; miR-499 (rs3746444) promotes invasion and inhibits cell apoptosis by targeting FOXO1A, FOXO4, PDCD4, PBX1, and SOX6 gene; miR-124 (rs34059726) has target genes DCC, CREB gene, and enhances invasion; miR-196-a2 (rs1614913) inhibits apoptosis and promotes invasion having target genes ANXA1 and HOX gene.

cytokine signaling pathways and thereby increase risk of PTC (Slaby et al., 2012).

miRNA-499

miR-499 is a microRNA that regulates multiple genes and signaling pathways post-transcriptionally, especially in hypoxic-ischemic situations like cancer and myocardial infarction (Wilson et al., 2010; Ando et al., 2014). Wang et al. (2015) observed reduced expression of miR-499-5p disrupted the PI3K/AKT/GSK signaling pathway (Liu et al., 2011). miR-499 functions as a tumor suppressor by decreasing cell proliferation causing apoptosis, which inhibits cancer progression. In addition, it also prevents metastases. FOXO4 and programmed cell death 4 (PDCD4) genes have been reported to be the targets of miR-499 (Liu et al., 2011). PDCD4 is an RNA-binding protein that stops particular mRNAs from being translated (Ohnheiser et al., 2015). PDCD4 modulates several signal transduction pathways and impacts the translation and transcription of many genes as a tumor suppressor (Wang et al., 2013a). It may play a key role in halting cell cycle progression and preventing tumor metastasis by inhibiting cell proliferation (Wei et al., 2012). Wei et al. (2012) demonstrate that PDCD4 may be important in stopping cell cycle progression at a critical checkpoint, limiting cell proliferation and suppressing tumor spread. In ovarian cancer cells, the PI3K-Akt pathway was thought to be involved in the regulation of PDCD4 degradation (Wei et al., 2012).

An SNP rs3746444 (T>C) has been reported in the seed region of mature miR-499 (Chen et al., 2014b). This SNP has been associated with increased susceptibility to various cancers like BC, cervical squamous cell carcinoma, acute lymphoblastic leukemia (ALL), colorectal cancer, liver cancer, gallbladder cancer, lung cancer, gastric cancer, squamous cell carcinoma of the head and neck, and prostate cancer. rs3746444 has been reported to be associated with an elevated risk of BC in the Chinese, German, and Italian populations; gastric cancer in the Japanese population; prostate cancer in the Indian population; cervical squamous cell carcinoma and lung cancer in the Chinese population, and ALL in the Iranian population (Tian et al., 2009; Catucci et al., 2010; Liu et al., 2010; Srivastava et al., 2010; Zhou et al., 2010; George et al., 2011; Okubo et al., 2011; Hasani et al., 2014). In contrast, Asian populations with the T allele of the miR-SNP are thought to have a lower risk of BC whereas Caucasians bearing the same variant allele have been reported at a higher risk of BC (Chen et al., 2014c).

rs3746444 leads to overexpression of miR-499 resulting in its enhanced binding to its target genes including FOXO4, PDCD4, and SOX6 gene (Dai et al., 2016). FOXO transcription factors regulate a variety of physiological activities, including fuel metabolism, oxidative stress response, and redox signaling, cell cycle progression, and apoptosis (Urbánek and Klotz, 2017). FOXO4 is a tumor suppressor protein that has associated with metastasis (Yang et al., 2006; Lee et al., 2009; Zhang et al., 2009). PDCD4 is a well-known tumor suppressor regulating the

growing, invading, or metastasis of the tumors. The study reported that prometastatatic action of miR-499 is on account of the suppression of FOXO4 and PDCD4 expression (Liu et al., 2011). PDCD4 inhibits the expression of mitogen-activated protein kinase (MAP4K1) via Jun N-terminal kinase (JNK) pathway. This was established by cDNA microarray analysis of PDCD4-overexpressing in RKO human colon cancer cells (Yang et al., 2006).

rs3746444 has also been reported to regulate the expression level of SOX gene. The anti-apoptosis action of miR-499 (rs3746444 T>C) can be reversed by up-regulating the SOX6 gene (Li et al., 2013a). Deregulation of the SOX gene activates the Wnt/-catenin signaling pathway, which has been linked to cancer development (Yan et al., 2017).

miRNA-125a

miR-125 plays a role in disease prevention and promotion, especially in cancer and host immunological responses. miR-125 inhibits a variety of genes, including transcription factors, matrix metalloproteinases, Bcl-2 family members, and others, causing aberrant cell proliferation, metastasis, and invasion, as well as carcinomas (Sun et al., 2013). BC, stomach cancer, and medulloblastoma all have lower levels of miR-125a, which promotes disease development. In human medulloblastoma cells, overexpression of miR-125a resulted in cell growth arrest and apoptosis. Furthermore, in stomach cancer cells, identical ectopic expressions inhibited growth. In BC cells, overexpression of miR-125a resulted in decreased anchorage-dependent proliferation. It was discovered that miR-125a modulates these cellular processes through Erbb2 in the context of gastric cancer and BC investigations (Scott et al., 2007; Ferretti et al., 2009).

The polymorphism rs12975333 (G>T) in miR-125a is in the seed sequence at the 8th nucleotide of mature miRNA. The T allele has been shown to inhibit the conversion of pri-miRNA to pre-miRNA precursor and is extremely rare, having been found only once in a panel of 1200 people from various ethnic origins and correlated with an elevated risk of BC in the Belgium population (Peterlongo et al., 2011). The reduced expression of mature miR-125a leads to the overexpression of the target genes.

The known targets of miR-125a like ERBB2 and ERBB3 have previously been reported to be associated with BC tumorigenesis (Morales et al., 2018). ERBB2 encodes the BC marker HER2 and alterations of ERBB2 and ERBB3 have been reported to promote malignancy. For example, ERBB2 overexpression is associated with approximately 25% of all human BC which drives the key aggressive features including cell proliferation, motility, and invasion (Lehmann et al., 2013). Malignant transformation can be induced by deregulation of ERBB2 and ERBB3 alone or in combination. Amplification and overexpression of ERBB2 have been associated with 25% of all human breast tumors. Overexpression of ERBB2 in particular promotes cell survival, proliferation, motility, and invasion, all of which are hallmarks of this aggressive form of human BC (Scott et al., 2007).

miRNA- 124

miR-124 is one of the most abundant miRNAs in the adult brain and is expressed primarily in the CNS. Mature miR-124 family includes three members, namely miR-124-1, miR-124-2, and miR-124-3. miR-124 has been demonstrated to induce cell differentiation while inhibiting cell proliferation in general (Ristori et al., 2017). Several cancers, including colon, breast, and lung cancers, as well as leukemia and lymphoma, are linked to miR-124 (Pal et al., 2015).

A G>T (rs3405972) has been reported in the seed sequence of miR-124-3 (Gong et al., 2012; Beretta et al., 2017). The major miR-124-3 targets include vesicle-associated membrane protein 3 (VAMP3), sialomucin core protein 24 (CD164), tyrosine-protein phosphatase non-receptor type 12 (PTPN12), neuronal growth regulator 1 (NEGR1), cyclin-dependent kinase 6 (CDK6), integrin Beta 1 (ITGB1), and insulin-like growth factor-binding protein 7 (IGFBP7) (Leong, 2013).

The 3′UTR of oncogene CDK6 is the target of mature miR-124. miR-124 epigenetic masking causes CDK6 activation and subsequent phosphorylation of retinoblastoma (Rb), resulting in cell growth acceleration which is directly involved in brain cancer (Pal et al., 2015). miR-124 expression leads to the downregulation of PTPN12 protein which regulates tyrosine phosphorylation and is implicated in cancer and cellular physiology. As PTPN12 reduces mammary cell proliferation and transformation, the targeting of PTPN12 by miR-124 suppresses its tumor suppressor behavior which promotes the oncogenic shift in breast and lung cancer (Sun et al., 2011). Leong Pei predicted that miR-124-3 with a variant allele targets novel genes DCC (deleted in colorectal cancer) and CREB5 (cyclic AMP-responsive element-binding protein 5) rather than PTPN12 (as predicted by public databases). The variant miR-124-3 is unable to suppress PTPN12 tumor suppressor and may alternatively behave as a tumor suppressor instead of an oncogene in breast or lung cancer (Leong, 2013). Hunt et al. (2011) reported miR-124 reduces oral squamous cell carcinoma (OSCC) invasion by targeting ITGB1, which is responsible for regulating intracellular signaling cascades and tissue homeostasis. Thus, miR-124 has a strong potential to be used as a prognostic marker in OSCC (Sun et al., 2011).

miRNA-196-a2

miR-196 family of molecules can operate as tumor suppressors. miR-196a, for example, inhibits metastasis in melanoma and BC cells, while miR-196b is downregulated in many types of leukemia cells. Bioinformatics research revealed that miR-196a2 could target multiple genes involved in cell cycle regulation, survival, and apoptosis, all of which could be relevant in GI malignancies. Cell proliferation, migration, invasion, and radio resistance are all functions of miR-196 family molecules' carcinogenic impacts (Fawzy et al., 2017).

The polymorphic C>T (rs11614913) is in the mature sequence of miR-196a-3p that negatively affects the processing of precursor miRNA to mature and subsequently its capability to regulate its target genes. Variant T allele influences the stability of the secondary structure of miR-196a2 (Wojcicka et al., 2014). This variant has been associated with an increased risk of bladder cancer; renal cancer; gastric cancer; lung cancer, HCC; glioma; head and neck squamous carcinoma; NSCLC and familial BC (Hu et al., 2008; Tian et al., 2009; Dou et al., 2010; Stenholm et al., 2013; Dai et al., 2015; Liu et al., 2018).

C allele impairs mature miRNA expression, resulting in lower levels of mature miR-196a2 (Yekta et al., 2004; Hoffman et al., 2009). In the Chinese population thus miR-SNP-induced decrease in miRNA expression could be used as a predictive biomarker for assessing BC risk (Qi et al., 2015). Other studies, on the other hand, have suggested that in some groups, the rs11614913 polymorphism predicts a lower risk of BC (Dai et al., 2016). A meta-analysis of 16 studies was carried out and observed that Caucasian patients had a lower risk of BC, with no significant influence on total risk (Zhang et al., 2017). It has been found that people with the CC genotype of this SNP were more likely to develop BC (Ferracin et al., 2011; Wang et al., 2013b).

rs11614913 associated with decreased risk of glioma and BC in Chinese populations and it has been found to be associated with reduced risk of cervical cancer in the Indian population (Hu et al., 2009; Dou et al., 2010; Thakur et al., 2019). According to recent studies, miR-196a2 TT genotype was associated with decreased risk for cervical cancer whereas miR-196a2 and CC/CT genotype was associated with higher risk. In another study, it was shown that C allele exhibited association with HCC in the Asian population but not in Caucasians whereas it increased the risk of colorectal, glioma, and prostate cancer in the non-African population compared to the African population (Zhao, 2020).

The potential molecular targets of miR-196a2 include HOX and annexin-A1(ANXA1) genes. ANXA1 was shown to be a key modulator of apoptosis and has since been linked to glucocorticoid activities such as cell proliferation suppression and cell migration control. ANXA1 plays a significant role in membrane trafficking, exocytosis, signal transduction, cell differentiation, and apoptosis, among other biological roles (Luthra et al., 2008). Overexpression of miR-196a2 due to variant rs11614913 leads to the suppression of ANXA1 thereby promoting cell proliferation and suppressing apoptosis (Solito et al., 2001; Luthra et al., 2008; Rahim et al., 2019). ANXA1 exhibits varied expression in different cancers. It is upregulated in glioma and oropharyngeal cancer and downregulated in prostate cancer, esophageal squamous cell, and head and neck squamous carcinoma (Rahim et al., 2019). As an upstream regulator, miR-196a has been found to partially direct the cleavage of the mRNA of the HOX gene clusters. HOX genes were shown to be abnormally expressed in BC, and HOXD10 was found to initiate tumor invasion and metastasis (Yekta et al., 2004).

SNPs in 3′UTR of miRNA Target Sequence

Several SNPs in miRNA binding sites or miRNA target gene (3′UTR) disrupt the capability to recognize the specific target. This results in dysregulation of target genes due to changes in miRNA and mRNA interactions (Nicoloso et al., 2010; Zheng et al., 2011). The presence of SNPs in the 3′UTR in an oncogene or a tumor regulatory gene might cause changes in gene regulation can shift the balance of cellular homeostasis toward cancer (Zheng et al., 2011). Variations in the 3′UTR of target genes involved in the stress response or DNA repair modify the activity of effector proteins, resulting in changes in the ability to repair damaged DNA and raising the risk of cancer (Cao et al., 2016). SNPs reported in 3′UTR of target genes of specific miRNA includes rs3092995, rs12516, rs8176318 in BRCA1 gene; rs4245739 in MDM4; rs1042538 in IQGAP1; rs7963551 in RAD52; rs9341070 in ESRI; and rs1071738 in PALLD gene. Information about genetic variation (SNPs) reported in 3′UTR of miRNA target gene along with functional outcome implicated in the pathogenesis of cancer have been summed up in **Table 2** and **Figure 2**.

BRCA 1

BRCA is one of the well-studied genes associated with BC. This gene has a highly conserved 3′UTR of 1381 nucleotides encodes a 1863 amino acid protein and functions as a tumor suppressor gene that regulates various cellular processes including cell cycle checkpoint control, chromatin remodeling, DNA repair, regulation of transcription, protein ubiquitination, and apoptosis (Yang et al., 2016; Ahmad et al., 2019). The 3′UTR region of the BRCA1 gene plays a pivotal role in the localization, stability, and mRNA transport. SNPs in 3′UTR of this gene might alter genes expression and therefore increase the risk of BC. The prevalence of SNPs in 3′UTR of BRCA1 modulating the miRNA binding site that can emerge as a significant biomarker of the disease (Ahmad et al., 2019).

SNPs including rs3092995, rs12516, and rs8176318 are in the 3′UTR of this gene. rs3092995 (G>A, C) is in the sequence of BRCA1 gene that interacts with the seed sequence of miR-103. rs3092995 is strongly associated with increased risk of BC in African American women. The variant alleles of rs8176318 (G>T) and rs12516 (G>A,T) are associated with ovarian and familial BC in Thai women (Pelletier et al., 2011). rs8176318 is located at the region where miR-639 is assumed to bind. Enhanced risk of BC was reported in individuals with GT or GG compared to TT genotype in the Pakistani population (Ahmad et al., 2019). This SNP is also associated with the risk of TNBC and ovarian cancer in an Irish population. The bearer of the rs8176318 variant allele has been linked to an increased risk of BC in menopausal women, as estrogen levels in these women drop after menopause (Malhotra et al., 2019b).

SNP rs12516 in the 3′UTR of BRCA1 gene alters the binding site of miR-1264, affecting its binding affinity and at the same time creating a binding site for other miRNAs including miR-4278, miR-4704-5p, and miR-637. Binding of these miRNAs to the 3′UTR of BRCA1 has been associated with higher NC risk. This rs12516 present in the 3′UTR of BRCA1 has been reported to be a genetic marker in the Turkish population associated with an increased risk of BC development (Erturk et al., 2014; Malhotra et al., 2019a).

MDM4

The murine double minute 4 (MDM4) protein was found as a p53-binding protein and has a fundamental amino acid sequence that is very similar to MDM2 (Shvarts et al., 1996). Over half of all human malignancies have a mutation in p53, the most frequently inactivated gene in cancer. p53 functions as a transcriptional factor that transactivates a set of genes involved in multiple cellular processes such as cell cycle arrest, cellular senescence, energy metabolism, and apoptosis in response to various extra- and intracellular stresses such as oncogene activation, DNA damage, and hypoxia. High

TABLE 2 | SNPs reported in target site of mi-RNA (mRNA-3′UTR region) in cancer.

S. No	Gene	miRNA	SNP reported	Tumor type	Reference
1	BRCA1	miR-639 miR-1264 miR-103 miR-637	rs8176318 G>T rs12516 C>T rs3092995 G>A,C	TNBC	Markey (2011); Malhotra et al. (2019a)
2	CD86	miR-337/200a-5p/184/212	rs17281995 G>C	Colorectal cancer	Moszyńska et al. (2017)
3	CDKN2B	miR-323-5p	rs1063192 G>A,T	Osteosarcoma	Cao et al. (2016); Elfaki et al. (2019)
4	DROSHA	miR-27b	rs10719 T>C	Bladder cancer	Moszyńska et al. (2017)
5	ESRI	miR-206	rs9341070 C>T	Breast cancer	Moszyńska et al. (2017)
6	ESRI	miR-453	rs2747648 C>T	Breast cancer	Wynendaele et al. (2010); Shankaran et al. (2020)
7	HIF1A	miR-199a	rs2057482 T>C	Pancreatic ductal adenocarcinoma	Moszyńska et al. (2017)
8	INSR	miR-612	rs1051690 G/A	Colorectal cancer	Moszyńska et al. (2017)
9	IQGAP1	miR-124	rs1042538 T>A,G,C	BC, ovarian, colorectal, glioblastoma, lung, gastric cancer	Stanford et al. (1986); Olefsky (2001); Sommer and Fuqua (2001); Wunderlich et al. (2004); Brekman et al. (2011); Erturk et al. (2014); Malhotra et al. (2019a)
10	ITGB4	miR-34a	rs743554 G>A	Breast cancer	Wynendaele et al. (2010)
11	MDM4	miR-191-5p, miR-887	rs4245739A > C	SCLC, prostate, ovarian, breast, lung, non-Hodgkin's lymphoma	Balenci et al. (2006); Dong et al. (2006); Moszyńska et al. (2017); Galka-Marciniak et al. (2019)
12	PALLD	miR-96 miR-182	rs1071738 C>G	Breast cancer	Moszyńska et al. (2017); Alipoor et al. (2018)
13	PRKD3	miR-145-5p miR-27b-3p miR-27a-3p	rs2160395 C>A,T	CRC- colorectal cancer	(Zhuang and Wang, 2017; Abo-Elmatty and Mehanna, 2019)
14	PSCA	miR-342-5p	rs10216533 G>A,C	GAC-gastric adenocarcinoma	Cao et al. (2016); Wu et al. (2019)
15	RAD52	Let-7	rs7963551 (C allele)	Breast cancer increase	Adams et al. (2007); Malhotra et al. (2019a)
16	TGFBR1	miR-628-5p	rs334348 A>G,T	Breast cancer	Malhotra et al. (2019a); Shankaran et al. (2020)
17	TP63	miR-140-5p	rs35592567 C>T	Bladder cancer	Moszyńska et al. (2017)

concentrations of p53 inhibitors can also inactivate p53 signaling. MDM2 and MDM4, two main negative regulators of p53, are substantially responsible for p53 activity suppression (Marine et al., 2006). MDM4 has also been reported to bind with p21 and direct it to proteasomal destruction without ubiquitination. Mdm4-mediated p21 degradation is independent of MDM2, yet it works together with MDM2 to break the G1 cell cycle arrest (Jin et al., 2008). Deletion of MDM4 causes multipolar spindle formation, increased chromosomal loss, higher proliferative potentials, and increased spontaneous tumor transformation in p53-null cells (Matijasevic et al., 2008). Because MDM4 and MDM2 form heterodimers, Mdm4 depletion may stimulate MDM2 interaction with other proteins such as Rb and p21, enhancing carcinogenesis. MDM4 also inhibits E2F1 transactivation by disrupting E2F1-DNA binding or changing the location of the E2F1 transcription complex (Strachan et al., 2003). E2F1 overexpression enhances the G1/S transition in cells and has been linked to cancer development. E2F1 overexpression, on the other hand, causes both p53-dependent and p53-independent apoptosis (Crosby and Almasan, 2004; Wunderlich et al., 2004; Stanelle and Pützer, 2006). These findings point to MDM4 upregulation, which is found in many malignant malignancies (Markey, 2011).

An SNP rs4245739 (A>C) is located within the 3′UTR region of MDM4. The C allele creates a new binding site for three miRNAs, miRNA-191-5p, miR-887, and miR-3669 (Wynendaele et al., 2010; Stegeman et al., 2015; Anwar et al., 2017). The variant allele of this polymorphism has been reported as a risk factor for many cancers including ovarian cancer, lung cancer, prostate cancer, BC, non-Hodgkin's lymphoma, esophageal cancer, and retinoblastoma (Wynendaele et al., 2010; Zhou et al., 2013; Gao et al., 2015a; Stegeman et al., 2015; Xu et al., 2016; Anwar et al., 2017).

miR-191-5p showed a greater binding affinity with C allele of rs4245739. In a genotype-based mRNA expression analysis, it was found that C allele was associated with decreased risk of ovarian cancer and retinoblastoma in an Asian population (Xu et al., 2016). MDM4 is overexpressed in A allele genotype and enhances the risk of BC and ESCC in a Chinese population (Zhou et al., 2013; Stegeman et al., 2015). The patients bearing homozygous

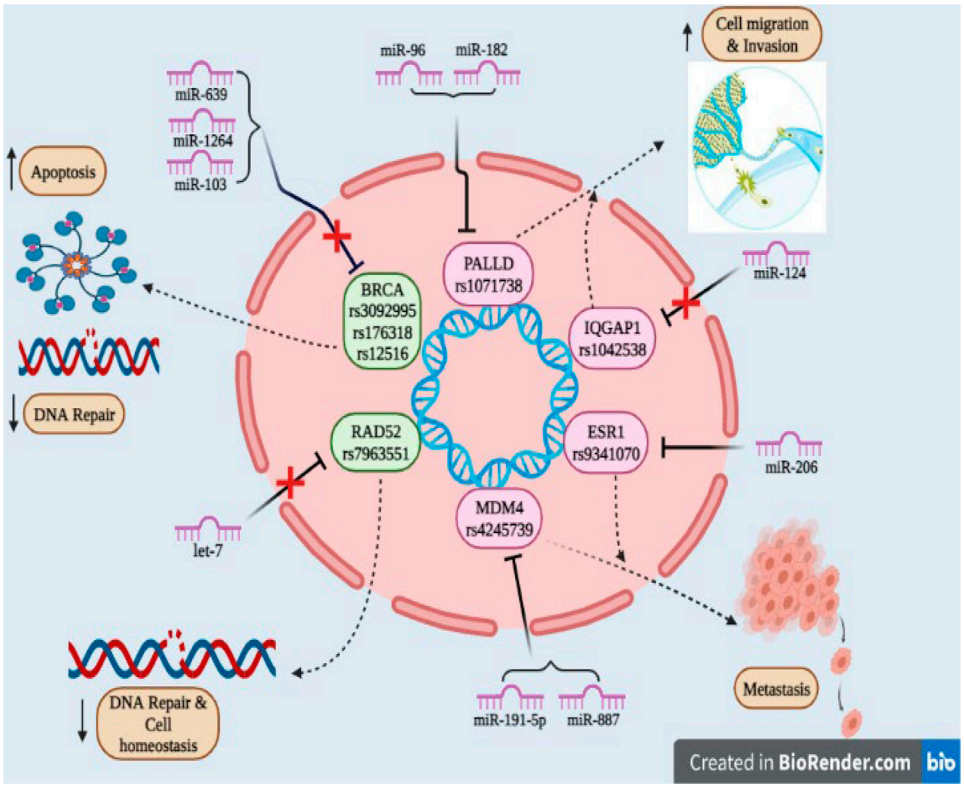


FIGURE 2 | SNPs reported in 3′UTR region of the target gene involved in cancer: Due to SNPs in the BRCA gene (rs3092995, rs176318, rs12516) and RAD52 gene (rs7963551) binding site for miR-639, miR-1264, and miR-103 within BRCA and for let-7 within the RAD52 gene found to be disrupted and as tumor suppressor genes they promote cell apoptosis, decrease the DNA repair mechanism and maintain cell homeostasis; PALLD (rs1071738), IQGAP1 (rs1042538), ESR1 (rs9341070), and MDM4 (rs4245739) genes act as oncogenes and they are responsible for metastasis, cell migration, and invasion.

AA allele were found to have 5.5-fold increased risk of tumor-associated mortality and 4.2-fold increased risk of recurrence (Guo et al., 2016). In an *in vitro* study carried out in PC3 cells in the case of C allele bearers, since it binds with miR-191-5p and miR-887. On other hand, A allele is un-targeted, that is, it directly enhances the risk of prostate cancer (Stegeman et al., 2015). Individuals carrying the rs4245739 C allele express low levels of MDM4 resulting in high DNA repair ability mediated by p53 and thus decreased cancer risk (Zhou et al., 2013).

IQGAP1

IQ-domain GTPase-activating proteins (IQGAPs) are a multi-domain protein family that regulate a variety of cellular processes such as cell adhesion, migration, extracellular signaling, and cytokinesis (Brown and Sacks, 2006). The first of three human IQGAP homologues, IQGAP1 is expressed throughout the body, whereas IQGAP2 and IQGAP3 are mostly found in the liver and intestine, the brain, and the testis (Weissbach et al., 1994; Nojima et al., 2008; Schmidt et al., 2008). IQGAP1 is hypothesized to contribute to the changed cancer cell phenotype by modulating signaling pathways involved in cell proliferation and transformation, cell-cell adhesion weakening cell motility and invasion stimulation (Johnson et al., 2009). Calmodulin, a ubiquitous calcium-binding protein, regulates IQGAP1 function via the IQ motifs, which are common calmodulin-interacting domains present in many proteins. Calmodulin is thought to affect IQGAP1 function by producing a conformational shift that affects IQGAP1-protein interactions and/or IQGAP1 subcellular localization (Briggs and Sacks, 2003). Both ERK and b-catenin-dependent signaling are aided by IQGAP1. IQGAP1 binds to B-Raf, MEK, and ERK leading to the activation of MAPK cascade. Constitutive MAPK pathway activation is a common oncogenic trigger in a variety of cancers, particularly those caused by Ras and B-Raf activating mutations. H-Ras and R-Ras were shown to have no detectable binding whereas active M-Ras had a favorable association (Roy et al., 2005; Nussinov et al., 2018). IQGAP1 is overexpressed in several cancers including ovarian cancer, colorectal cancer, glioblastoma, lung cancer, and gastric cancer (Nabeshima et al., 2002; Nakamura et al., 2005; Balenci et al., 2006; Dong et al., 2006).

miR-124 regulates IQGAP1 through a binding site in its 3′UTR. This target site sequence is disrupted by rs1042538 (T>A,C,G) in the core binding region. The presence of this variation at the miR-124 binding region has been suggested as a possible predictor of BC risk and prognosis. Based on a case-control study carried out in the Chinese population, the TT genotype was associated with a lower BC in comparison AA genotype, depicting that the T allele protects against BC (Zheng et al., 2011).

RAD52

Radiation sensitive 52 (RAD52) is a DNA strand exchange protein that mediates the DNA-DNA interaction required for complementary DNA strands to anneal during homologous recombination in DNA damage repair in order to maintain cell viability and homeostasis. Recent research has found that RAD52 plays a key function in mammalian cell genomic stability and cancer suppression (Feng et al., 2011; Lok et al., 2013). RAD52 stimulates the creation of nuclear foci, which appears to correspond to DNA repair sites, in response to DNA damage. RAD52 activity increases progressively as cells enter phase S, peaking in the S phase, and then disappearing at the start of G2. Phosphorylation and sumoylation are two post-translational changes that RAD52 can undergo. All these processes appear to work together to control the timing of RAD52 recruitment, as well as its stability and function (Liu et al., 1999; Barlow and Rothstein, 2010; Nogueira et al., 2019). RAD52 has also been shown to have a role in the response to oncogene-induced DNA replication stress (Sotiriou et al., 2016). High levels of RAD52 expression have been reported in tumor cells, particularly in lung squamous cell carcinomas and nasopharyngeal carcinoma tissues (Lieberman et al., 2016).

An SNP (C>A) rs7963551 is in the 3′UTR of RAD52 that is the binding site of let-7 miRNA (Jiang et al., 2013). In Chinese women, this SNP has been linked to an increased risk of BC. The presence of this variation reduces the binding capacity of let-7 to its target regions in the RAD52 3′UTR that has been suggested to boost its expression (Jiang et al., 2013). rs7963551 polymorphism with A allele was found to be strongly related with a lower incidence of SCLC in the Chinese population, according to a study. The functional genetic variant was only substantially associated with SCLC susceptibility among smokers but not with nonsmokers (Han et al., 2015).

ESR1

The nuclear hormone receptor and oncoprotein estrogen receptor alpha/estrogen receptor 1 (ER/ESR1) is overexpressed in around 70% of breast tumors (Stanford et al., 1986; Sommer and Fuqua, 2001). The ESR1 gene encodes estrogen receptor (ER), which is primarily a nuclear protein that operates as a ligand-dependent transcription factor (ER's genomic activity) (Olefsky, 2001). In primary human BC and human BC cell lines, MDM2 expression has been reported to be correlated with ER expression. The ER has been postulated to upregulate MDM2 expression (Baunoch et al., 1996; Hori et al., 2002; Brekman et al., 2011). MDM2 also forms a protein complex with ER, making it easier for ER to be ubiquitinated and degraded resulting in a negative feedback loop. However, the ability of ER and MDM4 (another member of the MDM family) to interact and regulate each other's expression in a comparable way has yet to be established (Liu et al., 2000).

The SNP rs9341070 (C>T) is one of the known polymorphisms located at 3′UTR of ESR1 gene at the binding site of miR-206. This variant influences the binding between miRNA and 3′UTR of ESR1 results in lower expression of ESR1 gene (Brucker et al., 2017). The T allele at 3′UTR allows binding of miR-206 to ESR1 and it is significantly downregulated (Adams et al., 2007). This SNP has been associated with an increased risk of BC (Anwar et al., 2017; Brucker et al., 2017).

PALLD

Palladin (PALLD), an actin-associated protein whose expression is intimately linked to the pathogenic cell motility properties of aggressive cancer cells, is encoded by the PALLD gene. Palladin expression is higher in invasive and malignant BC cell types than in noninvasive and normal cell lines. Palladin stimulates podosome formation, modulates the actin cytoskeleton via numerous routes, participates in matrix breakdown, and hence it aids in BC spread (Goicoechea et al., 2009; von Nandelstadh et al., 2014).

miR-96 and miR-182 inhibit BC cell migration and invasion by downregulating Palladin protein levels. This mechanism is disturbed by an SNP rs1071738 in the 3′UTR of the PALLD gene. The variation rs1071738 (C>G) is a very normal functional variant of the PALLD gene. The alternate G allele is substantially more prevalent than the ancestral minor C allele. The mRNA target sequence at the 3′UTR of PALLD is entirely complementary to the miR-96 and miR-182 seed areas. The presence of C allele favors the interaction of these miRNAs with the 3′UTR of PALLD. However, the presence of the variant G allele results in a mismatch between the two. miR-96 and miR-182 regulate the expression of PALLD reducing its expression by about 30 and 70%, respectively, in the presence of the normal CC genotype. However, the presence of the G allele leads to abolition of interaction between the two regulatory miRNAs and PALLD, on account of the mismatch between these two miRNAs and seed sequence of PALLD. In normal conditions miR-96 and miR-182 are involved in the prevention of BC metastasis. However, the G allele counteracts this impact (Gilam et al., 2016). The functional significance of rs1071738 has been proved by *in vitro* study carried out by MCF-7 (non-invasive BC cell lines) and Hs578 (highly invasive BC cell line) (Gilam et al., 2016; Moszyńska et al., 2017).

SNPS IN SEED SEQUENCE OF MIRNA AND THE 3′UTR OF SPECIFIC TARGET GENES IN DIABETES

SNPs in Seed Sequence of miRNA

SNPs reported in the seed sequence of miRNA associated with diabetes or its complications including rs3746444 in miR-499a; rs2910164 in miR-146a; rs7247237 in miR-3188; and rs34059726 in miR-124-3p. A detailed information involving SNPs in seed sequence of miRNA associated with diabetes and their functional implications has been given in **Table 3** and **Figure 3**.

miRNA-499a

The diabetic neuropathy including cardiovascular autonomic neuropathy (CAN) and diabetic neuropathy (DPN) have been reported to impact the quality of life in diabetics since these complications have been reported in a large percentage of diabetics. The miR-499a is an antiapoptotic and

TABLE 3 | SNPs reported in seed sequence of miRNA involved in diabetes.

S. No	miRNA	Target gene	SNP reported	Association with DM	Reference
1	miR-124a	Mtpn1, FOXA3, Sirt1, AKT1	rs531564 G>C	Protective role, facilitates glucose metabolism and insulin exocytosis	Wang et al. (2019); Zhao et al. (2013); Hribal et al. (2011)
2	miR146a	NF-kB, TNF associated factor 6, IL1R associated kinase 1	rs2910164 (C>G)	β-cell apoptosis	Zhao et al. (2013); Klöting et al. (2009)
3	miR3188	GSTM1, TRIB3	rs7247237 (C>T)	Impaired insulin signaling and apoptosis of human endothelial cells	Fawcett et al. (2010); Zhao et al. (2013)
4	miR126	PI3K Subunit-2 and SPRED-1	rs4636297 (A>G)	Protective factor against diabetic retinopathy, maintains vascular system integrity	Zhao et al. (2013)
5	miR125a	ENPP1 IL6R	rs12976445 (T>C)	Regulates IL6R leading to diabetic nephropathy	Fonseca et al. (2005); Bantubungi et al. (2014)
6	miR375	ADIPOR2	rs6715345 (G>C)	T1DM, T2DM, Insulin resistance syndrome	Mead et al. (2002); Zhao et al. (2013)
7	miR-499	PTEN	rs3746444 A>G	Diabetic neuropathy, impaired insulin signaling	Wang et al. (2019)

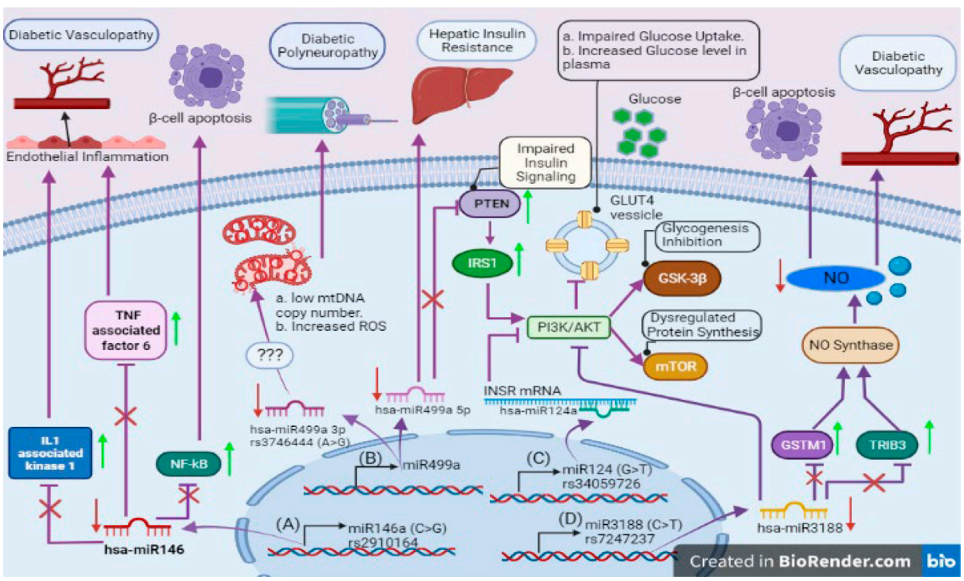


FIGURE 3 | SNPs reported in seed sequence of miRNA involved in diabetes: **(A)** miR146a with SNP rs2910164 (C>G) raises vascular complications caused by upregulation of inflammatory factors (TNF associated factor 6 and IL1 associated kinase 1) in endothelial cells and induces apoptosis in pancreatic β-cells via NF-kB mediated pathway. **(B)** Reduced expression of miR499a due to SNP rs3746444 (A>G) provokes mitochondrial stress, impairs insulin signaling via PTEN mediated pathway, and promotes hepatic insulin resistance. **(C)** The miR124 rs34059726 (G>T) creates complimentary sequence for INSR; causing failure to transport GLUT4 transporter vesicle to outer membrane. This SNP also inhibits glycogenesis process via GSK-3β activation. **(D)** miR-3188 with rs7247237 (C>T) inhibits PI3K/AKT pathway dysregulating protein synthesis. Its target genes overexpression (GSTM1 and TRIB3) curtails nitric oxide pathway which increases vascular complications and induces β-cell apoptosis.

cardioprotective miRNA (Wang et al., 2011; de Carvalho et al., 2019). It has been reported that polymorphisms in miR499 are involved in perturbed insulin secretion, CAN, and peripheral neuropathy (Ciccacci et al., 2018).

The genetic variation in miR-499a has been associated with the development of diabetic neuropathies. Especially the patients carrying the GG genotype of rs3746444 (A>G) present in the seed region of this miRNA are at higher risk of developing the CAN (Ciccacci et al., 2018). A study carried out in an Italian population investigate the association between mitochondrial DNA (mtDNA) copy number and rs3746444 in DPN patients. A decline in the mtDNA copy number in T2DM patients affecting DPN was observed in comparison with healthy controls (Latini et al., 2020). The increase in the copy number of mtDNA in association with the variant allele has been hypothesized to be on account of mitochondrial fission due to oxidative stress (Ghaedi et al., 2016). Increased ROS and mitochondrial injury might be contributing to nervous system dysfunction (Wang et al., 2011).

Apart from mitochondrial dysfunction, dyslipidemia was also observed in patients. Dyslipidemia plays a significant role in the pathogenesis of DN, synergistically with hyperglycemia (Vincent et al., 2009). An excess of long-chain fatty acids in T2D can lead to an accumulation of acetyl-CoA, as a product of mitochondrial beta-oxidation (Fan et al., 2020). miR499a-5p over-expression can enhance the glycogen level and improve insulin signaling by PTEN inhibition. Reduced miR-499-5p level is observed in hepatic insulin resistance (Wang et al., 2015a). miR-499-5p is involved in the signaling pathway of IRS1/PI3K/AKT and in particular miR-499-5p targets PTEN, which is an important regulator of the insulin signaling pathway (Peyrou et al., 2015). Therefore, unstable secondary structure with GG genotype reduces miR-499-5p levels, as a consequence an increase in PTEN impairs the insulin signaling.

miRNA-146a

The key biological role of miR146a is as immunosuppressive modulator which regulates inflammatory response. It downregulates innate immune response by suppressing expression of IRAK1 and TRAF2, decreasing NFkB activity (Gholami et al., 2020). Therefore, it functions as a negative regulator of NFkB and its inflammatory cascade and promotes apoptosis and inhibits migratory capacity by negative regulation of EGFR signaling pathway (Chen et al., 2013; Park et al., 2015).

SNP rs2910164 C>G is present within the seed sequence of miR146a which reduces its expression (Alipoor et al., 2018). This SNP plays a significant role in the pathogenesis of diabetes by participating in β-cell metabolism, proliferation, and death. The suppressed expression of miR146a enhances the activity of NFkB inflammatory events and induction of β-cell apoptosis responsible for diabetes and related complications (Elfaki et al., 2019). The potential targets of miR146a include TNF associated factor 6 and IL-1 receptor associated kinase 1 which regulate endothelial inflammation. The C>G transition causes overexpression of these target mRNAs resulting in T2DM (Shankaran et al., 2020). This polymorphism also increases the incidence of preeclampsia in gestational diabetes mellitus (GDM) (Abo-Elmatty and Mehanna, 2019).

In a study involving the Chinese population, rs2910164 was associated with an increased risk of T2DM. In some other studies, this SNP is also responsible for risk like diabetic nephropathy in T1DM patients and diabetic macular edema in T2DM patients of Caucasian population. Further it is also associated with diabetic polyneuropathy and GDM in the Italian and Egyptian populations (Zhuang and Wang, 2017).

miRNA-3188

miR-3188 is involved in regulation of the mTOR-P-PI3k/AkT pathway and has been reported to affect the pathogenesis of diabetic complications. It is one of the miRNAs discovered quite early.

rs7247237 (C>T) is considered to be located in the seed sequence of miR-3188 and has been associated with T2DM in the Chinese population. *In vitro* studies on HUVEC cell lines showed that the C allele expression was five times higher than T allele suggesting that C>T transition reduces its level which results in the overexpression of its targets; GSTM1 (glutathione S-transferase M1) and Trib3 (Tribbles pseudokinase3). This in turn reduces nitric oxide (NO) production in the endothelial cells through inhibition of endothelial NO synthase. There is also evidence that the overexpression of TRIB3 is associated with apoptosis in human endothelial cells, which could probably have an important role in the progression and pathogenesis of vascular complications in diabetes. As according to a study, miR3188 regulating mTOR and PI3K/AKT pathway involved in insulin signaling in endothelial cell; its reduced expression on account of the presence of rs7247237 resulting in T2DM (Wang et al., 2017; Wu et al., 2019).

RhoA/ROCK is another downregulated pathway by miR-3188 is RhoA/ROCK pathway via targeting ETS transcription factor ELK4. Elk4 is involved in various cancers and atherosclerosis. Its potential role via RhoA/ROCK pathway needs to be further elucidated (Li et al., 2017). The potential therapeutic value of miR-3188 could be further explored to mitigate the effect of pathogenic SNP (Wang et al., 2019)

miR-124-3p

The miR-124 is highly expressed in the brain and involved in epigenetic regulation of neurogenesis (Coolen et al., 2015). However, the *in vitro* studies miR124 overexpression of miR-124 in MIN6 pseudoislet cells caused impaired glucose induced secretion of insulin. Its silencing in MIN6 pseudoislet cells resulted in upregulation of its target genes FOXA2, Mtpn, Flot2, AKT3, Sirt1, and NeuroD1. All these targets are involved in normal Beta-cell functioning (Sebastiani et al., 2015). A study carried out in a mouse model demonstrated that miR-124 mediates triglyceride accumulation in the liver induced by high fat diet by directly targeting tribbles pseudokinase 3 (TRB3) and enhancing AKT signaling (Liu et al., 2016). An SNP rs34059726 located in the seed region of miR-124-3p, curated in PolymiRTS database is predicted to target insulin receptor transcript (INSR) (Gong et al., 2012). INSR belongs to tyrosine kinase receptor family which mediates insulin signaling via PI3K/AKT pathway. This pathway is responsible for maintaining glucose homeostasis, proliferation, and differentiation and inhibition of apoptosis (Chen et al., 2019). Down regulation of INSR via miR-124-3p leads to dysregulation of glucose uptake due to inhibition of GLUT4 vesicle transport to membrane and glucose transfer into cells (Jaakson et al., 2018). Lower PI3K/AKT signaling stimulates GSK-3β and inhibits glycogenesis. Failure of AKT to inhibit proapoptotic protein expression leads to apoptosis of the cells.

SNPs in 3′UTR of miRNA Target Sequence

The SNPs in 3′UTR of miRNA target genes reported in diabetes include rs11724758 in FABP-2; rs1046322 in WFS-1; rs2229295 in HNF1B; rs1063192 in CDKN2B; and rs13702 in LPL genes. Further details of other SNPs in 3′UTR of miRNA target genes implicated in pathogenesis of diabetes have been summed up in **Table 4** and the functional role in **Figure 4**.

TABLE 4 | SNPs reported in 3′UTR of miRNA target genes in diabetes.

S. No	Target gene	miRNA	SNP reported	DM association	Ref
1	HNF1B	miR-214 5p, miR-550a-5p	rs2229295 C>A	Susceptibility to T2DM	Xiang et al. (2012)
2	SLC30A8	miR- 183	rs3802177 G>A	T2DM	Grillari and Grillari-Voglauer (2010)
3	WFS1	miR-668	rs1046322 G>A	T2DM	Martin et al. (2007)
4	NLRP3	miR-223	rs10754558 C>G	Protective against T1DM, increased risk of T2DM via insulin resistance	Scutt et al. (2020)
5	WFS1	miR-185	rs9457 G>C	T2DM	Liu et al. (2014)
6	CDKN2A/B	miR- 323b-5p	rs1063192 CC	Gestational DM	Jeon et al. (2013)
7	ENPP1	miR-9 miR- 125 a/b	rs7754586 A>C	T2DM, end-stage renal disease	Marzec et al. (2007); Marczak et al. (2012)
8	ENPP1	miR-9, miR- 125 a/b	rs7754561 A>G	Insulin resistance and hypertriglyceridemia. Diabetic retinopathy in T2DM.	Thumser et al. (2014)
9	PYY	miR-663	rs162431 G>A	T2DM	von Otter et al. (2014)
10	LPL	miR-410	rs13702 T>C	T2DM	Schmidt et al. (2000)
11	PIK3RI	miR-29 a/b/c	rs3756668 G/G	Insulin resistance, T2DM	Sussan et al. (2008)
12	INSR1	Let-7a, miR27a	rs3745551 G/G	T2DM	Sussan et al. (2008)
13	INSR1	miR-106	rs1366600 T>C	GDM, T2DM	Tan et al. (2007); Shimoyama et al. (2014)
14	SLC30A8	miR-181a-2-3p	rs2466293 T>C	Impaired glucose regulation, reduced β cell function, GDM, T2DM	von Otter et al. (2014)
15	RNLS	miR-96	rs1048956 A>G	Affects insulin exocytosis, T2DM	Tan et al. (2007)
16	GSTA4	miR-200a	rs405729 G>A	Glucose stimulated insulin secretion	Tan et al. (2007)
17	TRIB3	miR-132	rs2295491 G>A	Obesity related impaired insulin secretion	Tan et al. (2007)
18	PRKCE	miR-410	rs41281467 C>T	Regulation of insulin secretion	Tan et al. (2007)
19	ACSL1	miR-34a	rs2292899 G>A	T2DM	Tan et al. (2007)
20	PDP2	miR-9	rs17767794 G>C	Negative control on insulin release	Tan et al. (2007)
21	SLC37A2	miR-9	rs3824926 T>C	Negative control on insulin release	Tan et al. (2007)
22	INSR1	miR-20b	rs1366600 T>C	T2DM	Tan et al. (2007)
23	FABP2	miR-132	rs11724758 G>A	Impaired insulin sensitivity	Tan et al. (2007)
24	IL10	miR-523	rs6687786 G>A	T1DM	Scutt et al. (2020)
25	CTLA4	miR-302a	rs13384548 G>A	T1DM	Scutt et al. (2020)
26	VDR	miR125b	rs3847987 C>A	T1DM	Michan and Sinclair (2007) miRdSNP
27	ESR1	miR-122	rs9341070	T2DM and breast cancer	North et al. (2003) miRdSNP
28	SLC15A4	miR-124	rs3765108	T2DM	Vaquero et al. (2006)
29	PPAR-δ	miR-1827	rs3734254	T2DM	Inoue et al. (2007)
30	ADIPOR2	miR-124, miR-375	rs1044471	Insulin resistance, T2DM	de Oliveira et al. (2012), miRdSNP
31	ADIPOR2	miR-1197 miR-375	rs12342	Insulin resistance syndrome, T2DM	de Oliveira et al. (2012), miRdSNP
32	LPIN2	miR-1227, miR-27a	rs3745012 G>A,C	T2DM	Crocco et al. (2015) miRdSNP
33	IL7R	miR-135b-5p, miR-135a-5p	rs6897932	T1DM	Sebastiani et al. (2015)
34	VPS26A	miR-381-3p	rs1802295	T2DM	Sebastiani et al. (2015)
35	HMG20A	miR-134-5p and miR-494-3p	rs7119 C>T	Impaired β-cell function	Sebastiani et al. (2015)
36	MAPK1	miR7-5p	rs12158121	β–cell proliferation by regulating mTOR pathway	Sebastiani et al. (2015)

(Continued on following page)

TABLE 4 | (*Continued*) SNPs reported in 3′UTR of miRNA target genes in diabetes.

S. No	Target gene	miRNA	SNP reported	DM association	Ref
			A>C		
37	LIN28A	Let-7a	rs3811463 T>C	GDM	Hiratsuka et al. (2003)
38	LIN28A	miR-125a	rs3811463 T>C	T2DM	Hiratsuka et al. (2003)
39	PAX4	miR-125a, miR-223	rs712699 G>A	Influences β-cell differentiation and survival	Li et al. (2013b)
40	KCNB1	miR- 448, miR-214, miR-153	rs1051295 T>C	B-cell compensatory secretory function and reduced insulin sensitivity	(Li et al., 2013b) miRdSNP

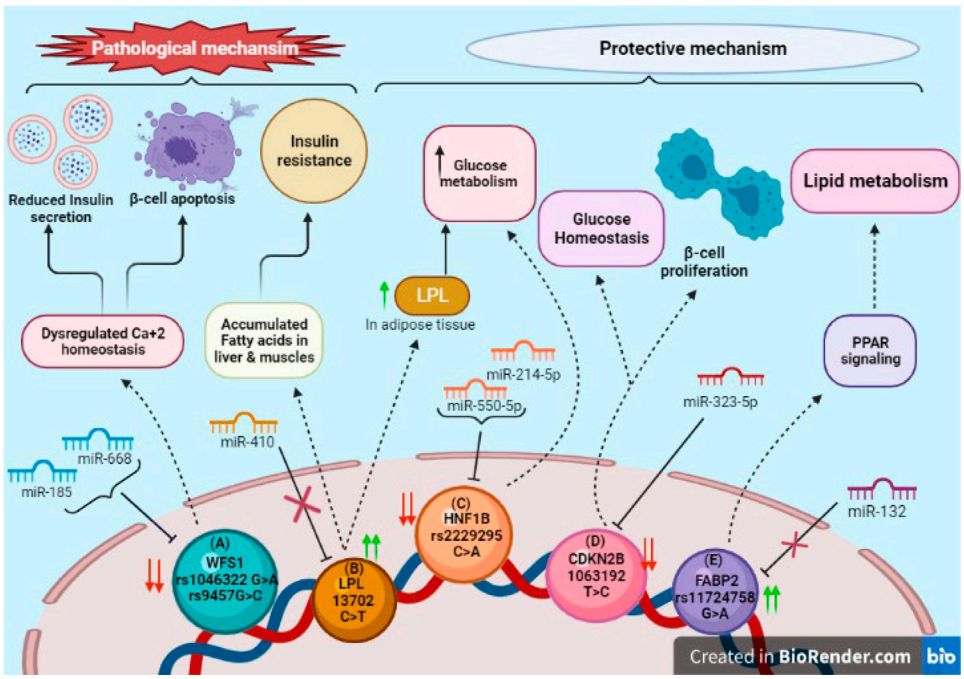


FIGURE 4 | SNPs in 3′UTR of target genes associated with diabetes: **(A)** Downregulation of WFS1 due to SNPs rs1046322 (G>A) and rs9457 (G>C) via miRNAs miR185 and miR668 induces β-cell apoptosis and declines insulin secretion due to ER Ca^{+2} stress. **(B)** SNP rs13702 (C>T) in LPL gene disrupts binding site for miR410. Its overexpression has pathological impact on liver and muscles leading to insulin resistance; whereas in adipose tissue it increases glucose metabolism. **(C)** SNP rs2229295 (C>A) of HNF1B creates new binding site for miR214-5p and miR550-5p, increasing glucose metabolism. **(D)** CDKN2B's downregulation due to SNP rs1063192 (T>C) by miR323-5p maintains glucose homeostasis and promotes proliferation of β-cells. **(E)** Over-expression of FABP2 due to SNP rs11724758 (G>A) activates PPAR signaling pathway which in-turn leads to lipid metabolism.

FABP-2

Fatty acid binding protein-2 belongs to ubiquitous lipid chaperones family is expressed in intestines which regulate fat absorption by intracellular trafficking of long chain fatty acids, eicosanoids, and other lipids (Haunerland and Spener, 2004). Its dysregulation has been associated with non-alcoholic hepatic liver disease and obesity (Thumser et al., 2014). Around 20–30% T2DM patients have renal impairment and FABP 2 is a novel biomarker for diabetic nephropathy (Tsai et al., 2020). FABP2 has been associated with insulin resistance mechanisms, indicating its essential role in protection against T2DM (Baier et al., 1995).

The 3′UTR polymorphism of rs11724758 (G>A) in FABP2 gene causes loss of binding site for miR-132. The miR-132 plays a significant role in adipose tissue dysfunction and obesity associated diabetes (Klöting et al., 2009). The AA genotype of FABP2 has been associated with decreased risk of T2DM compared to GG genotype (Zhao et al., 2013). Therefore, G>A transition functions as a protective factor against T2DM. FABP2 is involved in intracellular fatty acid transportation and fat absorption via PPAR signaling (Zhao et al., 2013).

WFS-1

Wolframin or WFS1 gene encodes for endoplasmic reticulum trans-membrane protein highly expressed in brain, pancreas, and heart (Hofmann et al., 2003). Mutation in this gene leads to a metabolic condition known as Wolfram Syndrome inherited in

an autosomal recessive manner (Harel et al., 2015). Two SNPs within 3′UTR of WFS1—rs1046322 and rs9457—have been reported to be as risk factors for T1DM and T2DM, respectively (Fawcett et al., 2010; Kovacs-Nagy et al., 2013; Elfaki et al., 2019). WFS1 is an ER transmembrane protein highly expressed in pancreas and insulinoma β-cell lines. It plays a significant role in maintaining Ca+2 ER homeostasis (Hofmann et al., 2003). In T2DM, it indicated that glucose induced insulin secretion has been found to increase WFS1 expression. The increased insulin production caused by insulin resistance in T2DM leads to chronic ER stress contributing to the death of β-cells by apoptosis. It was demonstrated that glucose induced insulin secretion leads to increased WFS1 expression in wild-type mice, whereas ER stress and β-cell dysfunction can be observed in WFS1 knock-out animals (Fonseca et al., 2005). *In vitro* interaction between WFS1 3′ UTR and miR-668 signified the rs1046322 influenced the affinity of miR-668 to WFS1 mRNA. In an *in vitro* luciferase assay it was observed that variation in 3′utr of WFS1 gene rs1046322 "A" and rs9457 "C" is sensitive to both miR-185 and miR-668, although the effect of miR-185 seemed to be stronger (Elek et al., 2015). miR-185 was reported to be strongly associated with diabetes mellitus via its targets SOCS3 and WFS1 (Bao et al., 2015; Elek et al., 2015). They showed that these different pathways can be in the background of the same phenotype, as miR-185 is suggested to be related to diabetes mellitus via WFS1 target (Bao et al., 2015).

HNF1B

HNF1B encodes for HNF1β (hepatocyte nuclear factor 1-β) homeodomain containing transcription factor expressed in pancreas, liver, and kidney (Coffinier et al., 1999). It regulates the critical function of pancreatic development, glucose metabolism, and hepatic insulin activity (Goda et al., 2015). It is the most common transcription factor associated with monogenic diabetes leading to young adult onset of T1DM with dominant inheritance patterns in familial cases. An SNP rs2229295 (C>A) within 3′-UTR of HNF1β acts as a protective factor against T2DM (Moszynska et al., 2017). In silico analysis revealed that rs2229295 in 3′UTR of HNF1β creates the binding site for miR-214-5p and miR-550-5p. The A allele of HNF1β is responsible for post-transcriptional regulation by miR214-5p and miR-550-5p. Therefore, due to this variation, expression of HNF1β is downregulated and thereby it acts as a protective factor against T2DM (Goda et al., 2015).

CDKN2B

CDKN2A/B highly expressed in pancreas is considered as a strong determinant of diabetes mellitus. The SNP rs1063192 T>C located within 3′UTR region of CDKN2A/B is associated with increased risk of GDM in pregnant Chinese Han women population (Wang et al., 2015b). The tumor-suppressor products of CDKN2A/B, p15INK4b, and p16INK4a inhibit important CDKs, i.e., CDK4 and CDK6, essential for β-cell proliferation and regeneration (Krishnamurthy et al., 2006). The T>C transition creates complimentary sequence of miR-323-5p which reduces the expression of CDKN2A/B (Hribal et al., 2011). The decreased expression of CDKN2A/B due to rs1063192 T>C results in reduced inhibition of CDK6 by p15INK4b and facilitates β-cell proliferation, lowering DM risk. On the other hand, increased expression of p15INK4b regulates glucose homeostasis. It can be speculated that miR-323-5p may also regulate other crucial genes responsible for β-cell hyperplasia and insulin signaling. Moreover, duality of p15INK4b in glucose homeostasis and deficiency of p16INK4a inducing *in vivo* hepatic glucose production via PKA-CREBPGC1a pathway possibly explains its role in GDM (Bantubungi et al., 2014).

LPL

Lipoprotein lipase enzyme is involved in hydrolysis of low-density lipoproteins and circulating chylomicrons into non-esterified fats which can be absorbed by the tissues. Disturbance in this conversion could lead to various abnormalities such as Alzheimer's, dyslipidemia, and diabetes (Mead et al., 2002). In liver and muscle tissues, the free fatty acid generated by LPL activity gets accumulated, leading to insulin resistance (Kim et al., 2001). The SNP rs13702 C>T in 3′UTR of LPL mRNA is located within the seed recognizing region of miR-410 (Hatefi et al., 2018). This C>T transition disrupts the binding site of miR-410 which leads to increased expression of LPL (Richardson et al., 2013). However, in adipose tissue, LPL increases glucose metabolism and insulin tolerance (Walton et al., 2015). Knockdown studies in MIN6 cells indicated decreased ability of glucose induced insulin secretion. In the Iranian population, the T allele of rs13702 showed protective association whereas C allele was found to be a risk factor against T2DM (Hatefi et al., 2018).

SNPS IN THE SEED SEQUENCE OF MIRNA AND TARGET GENES ASSOCIATED WITH VARIOUS DISEASES

SNPs within seed sequence of miRNA and their target genes have been also implicated in the development of various human diseases like Parkinson disease, asthma, periodontal disease, neurodegenerative disease, cardiovascular disease, and kidney and liver diseases (Martin et al., 2007; Tan et al., 2007; Rademakers et al., 2008; Wang et al., 2008; Schaefer et al., 2010; Bruno et al., 2012). The presence of SNPs in miRNA seed regions has a major impact on miRNA target loss and gain (generates a new repertoire of target genes), resulting in a considerable change in miRNA biological function (Xu et al., 2013; Zhang et al., 2019).

Among human diseases, ischemic stroke is one of the complicated diseases that consist of a variety of conditions with different hereditary and environmental risk factors. miRNAs played a role in a variety of physiopathological processes, and frequent SNPs in pre-miRNAs have been linked to disease vulnerability in humans (Liu et al., 2014). According to a case-control study, SNP (A>G) rs3746444 located within seed sequence of miR-499 may be significantly linked with higher risk of ischemic stroke in the Chinese community (Liu et al., 2014). miR-499 has been

TABLE 5 | SNPs reported in seed sequence of miRNA involve in various disease.

S.No	miRNA	Gene	SNP reported	Disease	Reference
1	miR-1304	SEMA3F	rs79759099 A>G	Pulmonary valve disease, pulmonary valve stenosis	Cao et al. (2016)
2	miR-662	ATP6VOE1	rs9745376 G>A/C/T	SLE	Cao et al. (2016)
3	miR-96-3p	-	rs546098287 A>G	Non-syndromic hearing loss	Malhotra et al. (2019b)
4	miR-548	NS1ABP, MAPK, CDK13	rs515924 A>G	Influenza virus infection	Kim et al. (2011); Erturk et al. (2014)
5	miR-122		rs41292412 C>T	AMD-age related degeneration	Erturk et al. (2014); Fawzy et al. (2017)
6	miR-431	RTL1	rs12884005 A>G	Autism	Hulf et al. (2011); Fawzy et al. (2017)
7	miR-3161	PTPRJ	rs11382316 -/A	Androgen insensitivity syndrome	Fawzy et al. (2017)
8	miR-499-3p	BCL2	rs3746444 A>G	Ischemic stroke	Cao et al. (2016)
9	miR-3618	DGCR8	rs12159555 C>G	Digeorge syndrome	Pelletier et al. (2011); Fawzy et al. (2017)
10	miR-4284	STX1A	rs11973069 C>T	Arteriosclerosis obliterons and pediatric ulcerative colitis	Pelletier et al. (2011)
11	miR-221		rs113054794 A>C	Crohn's disease	Gao et al. (2015b); Fawzy et al. (2017)

associated with ischemia condition, apoptosis, and cell death in anoxia via knockdown or calcineurin over-expression, inhibiting Drp1 dephosphorylation and mitochondrial fragmentation caused by anoxia (Wang et al., 2011). The rs3746444 polymorphism changed the stem structure of the miR-499 precursor from an A:U pair to a G:U mismatch, which changed the function or expression of mature miR-499, as well as the regulation of target mRNAs, influencing the risk of ischemic stroke (Hu et al., 2009; Xiang et al., 2012). Its targets includes peptyl arginine deiminase type 4, regulatory factor X4, IL-2, IL-2 receptor B (IL-2R), IL-6, IL-17 receptor B (IL-17RB), IL-18 receptor (IL-18R), IL-21, IL-23a, and B and T lymphocyte attenuator (Yang et al., 2012). miR-499/rs3746444 bound to its mentioned targets and can influence inflammation, fibrinogen, and CRP formation. Higher plasma CRP levels can raise blood pressure, BMI, insulin resistance, and lipids making CRP one of the common causes of cerebral ischemia (Yang et al., 2012; Jeon et al., 2013). Increased CRP, inflammation, and fibrinogen in the allele G and carried G genotypes of rs3746444 A/G may play a predisposing role in the development of ischemic stroke (Liu et al., 2014).

SNP in the 3′UTR of mRNA/target gene might disrupt or create the binding sites for miRNA. The renin-angiotensin system (RAS) is a key player in blood pressure regulation and is thought to be a contributing element in the development of hypertension (Laragh et al., 1991). Angiotensin II is a key player in the RAS pathway, inducing vasoconstriction, salt retention, and water retention, and is closely linked to the inflammatory, thrombotic, and fibrotic factors. Angiotensin II receptor type 1 (AGTR1) and type 2 (AGTR2) mediate these effects both directly and indirectly (AGTR2). AGTR1 is mostly found in vascular smooth muscle cells, as well as the heart, adrenal gland, and kidney (Oparil and Weber, 2000). A study on miR-155 and SNPs in the angiotensin II receptor, type 1 gene has been conducted by Sethupathy et al. (2007). They discovered that miRNA miR-155 could bind to the A allele of the SNP rs5186 (A>C) in the 3′UTR of the AGTR1 mRNA more efficiently than the C allele (which is more common in essential hypertension) (Sethupathy et al., 2007). In persons with the A allele, the binding of miR-155 has the capacity to reduce the level of AGTR1 mRNA and hence cause the pressor effect of Angiotensin II. Protein levels of AGTR1 in untreated essential hypertension patients homozygous for the C allele of rs5186 were also favorably linked with systolic and diastolic blood pressure. The expression levels of miR-155 were also negatively linked with AGTR1 protein levels, and miRNA levels were lower in those with the CC genotype that is directly associated with hypertensions (Ceolotto et al., 2011). **Tables 5** and **6** give a brief glimpse of SNPs reported in miRNAs or target gene sequences in various diseases.

SNPs Role in Aging

Human aging is a complicated process that has been related to dysregulation of a variety of cellular and molecular processes, including telomere shortening, altered DNA damage response, protein homeostasis loss, cellular senescence, and mitochondrial failure. These cellular and molecular processes can result in a wide range of illnesses, including cancer, cardiovascular disease, and neurological disease, as well as an increased chance of death (Huan et al., 2018). The study of the mechanics of the aging process could also benefit from the determination of an individual's SNPs. A comparison of the DNA sequences of healthy young individuals with the DNA sequences of healthy, extremely elderly people could reveal genes that play a big role in determining how long people live. Animal models had already been researched, and specific genes, such as DNA repair genes, had also been studied because of the role of repair processes in aging (Schmidt et al., 2000; Ruttan and Glickman, 2002). In

TABLE 6 | SNPs in 3′UTR in target gene of miRNA involved in various disease.

S. No	Gene	miRNA	SNP reported	Disease	Reference
1	APOC3	miR-4271	rs4225 G>T	Coronary heart disease	Moszyńska et al. (2017)
2	APOA5	miR-485-5p	rs2266788 G>A	Hyper-triglyceridemia	Moszyńska et al. (2017)
3	PLIN4	miR-522	rs8887 T>C,G	Antropometrics (Obesity related)	Moszyńska et al. (2017)
4	FXN	miR-124-3p	rs1145043 G>T	Friedrich's atria FRDA	Moszyńska et al. (2017)
5	SNCA	miR-34b	rs10024743 T>G	Parkinson's disease	Moszyńska et al. (2017)
6	EFNB2	miR-137	rs550067317 A>C	Schizophrenia	Moszyńska et al. (2017)
7	FGF20	miR-433	rs12720208 C/T	Parkinson's disease	Cao et al. (2016); Moszyńska et al. (2017)
8	AGTR1	miR-155	rs5186 A>C	Hypertension Renal disease	Jin et al. (2008); Cao et al. (2016); Moszyńska et al. (2017)
9	DHFR	miR-24	rs34764978 C>T	Methotrexate resistance	Moszyńska et al. (2017)
10	HLA-G	miR-48a miR-152	rs1063320 C>G	Childhood asthma	Moszyńska et al. (2017)
11	HTR1B	miR-96	rs13212041 A/G	Arson/property damage	Moszyńska et al. (2017)
12	HTR3E	miR-510	rs56109847 G>A	Diarrhea irritable bowel syndrome	Moszyńska et al. (2017)
13	TGFB1	miR-187	rs1982073 (rs1800470) G>A,C	Frozen shoulder development	Wynendaele et al. (2010); Hoss et al. (2014)
14	MMP9	miR-491-5p	rs1056628 A>C	SLE, Early neurologic deterioration	Wynendaele et al. (2010); Sondermeijer et al. (2011)
15	HTR3E	miR-510	rs62625044 G>A rs56109847	Irritable bowel syndrome	Usiello et al. (2000); Cao et al. (2016)
16	PFAS	miR-149-3p	rs1132554 C>T	Alcohol related neurodevelopment disorder	Cao et al. (2016)
17	TRIB2	miR-877-5p	rs1057001 T>A	Obesity	Olefsky (2001); Cao et al. (2016)
18	TMCO1	miR-296-3p	rs6660601 C>T	Skeletal anomalies, mental retardation syndrome	Cao et al. (2016)
19	SRSF3	miR-7f-2-3p	rs7344 T>C	POAG- primary open angle glaucoma	Crocco et al. (2015); Cao et al. (2016)
20	CAV2	miR-244-5p	rs1052990 T>C,G	POAG	Cao et al. (2016)
21	AAGAB	miR-329-5p	rs1050285 T>C	Punctuate palmoplantar keratoderma type1	Cao et al. (2016)
22	ABO	miR-855-3p	rs8176751 C>A,T	Hematological phenotype	Cao et al. (2016)
23	MTPN	Let-7/miR-98	rs17168525 G>A	Cardiac hypertrophy	Jin et al. (2008); Wang et al. (2013a)
24	MS4A6A	miR-382-3p	rs610932 T>G	LOAD- Late onset Alzheimer disease	Bäckman et al. (2006); Erturk et al. (2014)
25	TCF-21	miR-224-5p	rs12190287 C>G,T	Acute Coronary syndrome	Noble (2003); Erturk et al. (2014)
26	POCRlD	miR-425-3p miR-5444a/ miR-507	rs7097 C>T	DLBCL Diffuse large B-cell lymphoma	Abo-Elmatty and Mehanna (2019)
27	PRKD3	miR-329-3p mir495-3p	rs8243 C>A	Polycystic kidney and liver disease	Abo-Elmatty and Mehanna (2019)
28	ZNF155	miR-708-5p miR-28-5p	rs442220 G>A,C	Herpes simplex virus 1	Abo-Elmatty and Mehanna (2019)

recent years, it has also been suggested that post-transcriptional control by miRNAs may play a role in the phenotypic changes seen throughout aging by epigenetically modifying the expression of important regulatory proteins (Grillari and Grillari-Voglauer, 2010; Liu et al., 2012).

Human aging is linked to increased susceptibility to adverse drug reactions (ADRs), multimorbidity, and frailty, however, the intensity and age at which people become ill varies greatly. Identifying genetic indicators for this phenotype's higher risk might aid in the stratification of individuals who would benefit from specialized intervention. Nuclear factor (erythroid-derived 2)-like 2 (Nrf2) controls the expression of enzymes involved in drug metabolism, as well as the cell's response to stresses. In animal aging models, its expression has been demonstrated to

diminish (Scutt et al., 2020). In the promoter region of the human Nrf2 gene (NFE2L2), there are many single-nucleotide polymorphisms (SNPs) that influence Nrf2 expression *in vivo*. Specific age-related disorders, such as acute lung damage, reduced forearm vasodilator response, and Parkinson's disease, have been linked to these SNPs. Individuals with a variant allele may be more susceptible to the negative effects of medications, have a higher number of comorbidities, and be frailer in the setting of an age-related reduction in Nrf2 (Marzec et al., 2007; Marczak et al., 2012; von Otter et al., 2014). According to a recent study, polymorphism rs35652124 (T>A,C,G) in NFE2L2/Nrf2 gene was found to be associated with aging. Because the G allele is linked to lower NFE2L2/Nrf2 expression, another possible reason for the AA genotype's higher risk of multimorbidity and frailty is that high NFE2L2 levels are harmful in some disorders. When compared to control mice, Nrf2 knockout animals had a smaller atherosclerotic plaque (Sussan et al., 2008). Furthermore, the rs35652124 AA genotype is linked to a higher risk of high blood pressure and cardiovascular death in adults (Shimoyama et al., 2014). As a result, it is possible that cardiovascular pathology is to blame for the increased levels of multimorbidity and frailty (Scutt et al., 2020).

SIRT2 is one of seven mammalian sirtuins (Sir2-like proteins) that play critical roles in cellular activities such as metabolism and differentiation (Michan and Sinclair, 2007). It is mostly found in the cytoplasm, where it deacetylates -tubulin, but it also migrates to the nucleus during the G2/M phase, where it deacetylates histones, influencing cell cycle progression (North et al., 2003; Vaquero et al., 2006; Inoue et al., 2007). SIRT2 also deacetylates numerous additional substrates (PEPCK1, FOXO1, FOXO3a, p65, and p53) that are involved in key cellular processes linked to organism health, such as homeostasis, oxidative stress management, inflammation, and cell growth and death regulation (de Oliveira et al., 2012). SIRT2 variation rs45592833 (G>T) is located inside a binding region identified by three distinct miRNAs (miR-3170, miR-92a-1-5p, and miR-615-5p), all of which were expected to bind more firmly to the T allele, causing SIRT2 production to be reduced (Crocco et al., 2015). SIRT2 levels have been discovered to be low in various human malignancies, and SIRT2-deficient animals have been reported to develop tumors as they age (Hiratsuka et al., 2003; Kim et al., 2011; Li et al., 2013b). miR-615-5p, which has the highest binding energy change caused by rs45592833, has been found to be deregulated in cancer cell lines, patients with aging-related conditions such as Huntington's and cardiovascular diseases, and in the muscles of old mice, implying that miR-615-5p downstream targets may be involved in signaling pathways that are important in the aging process (Hulf et al., 2011; Sondermeijer et al., 2011; Hoss et al., 2014; Gao et al., 2015b).

An SNP in the DRD2 gene, rs6276 (A>G), which encodes a G protein-coupled receptor found on postsynaptic dopaminergic neurons, was also found to have a substantial connection with the longevity phenotype. DRD2 signaling is required for the appropriate control of a variety of physiological activities, including locomotion, behavior, and hormone synthesis (Usiello et al., 2000). Six distinct miRNAs were projected to bind to the region containing the polymorphism rs6276 using in silico analysis, with miR-485-5p having the greatest energy binding level to the 3′UTR with the minor G allele (Crocco et al., 2015). As a result, G allele binding is likely to be linked to enhanced miRNA–mRNA binding, resulting in more severe DRD2 expression regulation. DRD2 expression has been found to be downregulated in both striatal and extrastriatal areas of the brain in elderly adults, and that changes in DRD2 receptor density or activity have been linked to age-related declines in motor and cognitive abilities (Noble, 2003; Bäckman et al., 2006; Crocco et al., 2015).

CONCLUSION

Dysregulation of miRNAs and their targets is often reported to be involved in cancer progression. Multiple mechanisms for regulation of target gene expression by miRNAs have been proposed. However, recent evidence suggested another layer of complexity in terms of SNPs in either the miRNA seed or their target sequences. These mutations may cause dysregulated gene expression leading to cancer progression. Similar evidence is emerging in diabetes mellitus as well. There is limited scientific literature reporting SNPs in miRNA seed sequences highlighting scope of further exploration.

Overall, the current evidence suggests the need for the in-depth sequence analysis of miRNAs and target genes as well as to correlate the genetic evidence with functional studies. Since single miRNA can target multiple genes and similarly single genes can be targeted by multiple miRNAs, understanding the functional implications of these SNPs can provide new information regarding mechanisms of disease progression.

AUTHOR CONTRIBUTIONS

S.S. and A.M. conceptualized the idea. Y.C., A.S., and P.S. wrote the manuscript and prepared the figures and tables. S.S. and A.M. edited the manuscript. The manuscript has been read and approved by all the authors.

FUNDING

Y.C. is thankful to CSIR for providing SRF, P.S. is thankful to ICMR for providing RA fellowship. A.S. was Master's student in the department. The authors are thankful to Central University of Punjab for support.

REFERENCES

Abo-Elmatty, D. M., and Mehanna, E. T. (2019). MIR146A Rs2910164 (G/C) Polymorphism Is Associated with Incidence of Preeclampsia in Gestational Diabetes Patients. *Biochem. Genet.* 57 (2), 222–233. doi:10.1007/s10528-018-9886-1

Adams, B. D., Furneaux, H., and White, B. A. (2007). The Micro-ribonucleic Acid (miRNA) miR-206 Targets the Human Estrogen Receptor-α (ERα) and Represses ERα Messenger RNA and Protein Expression in Breast Cancer Cell Lines. *Mol. Endocrinol.* 21 (5), 1132–1147. doi:10.1210/me.2007-0022

Ahmad, M., Jalil, F., Haq, M., and Shah, A. A. (2019). Effect of Variation in miRNA-Binding Site (Rs8176318) of the BRCA1 Gene in Breast Cancer Patients. *Turk J. Med. Sci.* 49 (5), 1433–1438. doi:10.3906/sag-1905-17

Alipoor, B., Ghaedi, H., Meshkani, R., Omrani, M. D., Sharifi, Z., and Golmohammadi, T. (2018). The Rs2910164 Variant Is Associated with Reduced miR-146a Expression but Not Cytokine Levels in Patients with Type 2 Diabetes. *J. Endocrinol. Invest.* 41 (5), 557–566. doi:10.1007/s40618-017-0766-z

Ando, H., Asai, T., Koide, H., Okamoto, A., Maeda, N., Tomita, K., et al. (2014). Advanced Cancer Therapy by Integrative Antitumor Actions via Systemic Administration of miR-499. *J. Controlled Release* 181, 32–39. doi:10.1016/j.jconrel.2014.02.019

Anwar, S. L., Wulaningsih, W., and Watkins, J. (2017). Profile of the Breast Cancer Susceptibility Marker Rs4245739 Identifies a Role for miRNAs. *Cancer Biol. Med.* 14 (4), 387–395. doi:10.20892/j.issn.2095-3941.2017.0050

Ayaz Durrani, I., Bhatti, A., and John, P. (2021). Regulatory MicroRNAs in T2DM and Breast Cancer. *Processes* 9 (5), 819. doi:10.3390/pr9050819

Bäckman, L., Nyberg, L., Lindenberger, U., Li, S.-C., and Farde, L. (2006). The Correlative Triad Among Aging, Dopamine, and Cognition: Current Status and Future Prospects. *Neurosci. Biobehavioral Rev.* 30 (6), 791–807. doi:10.1016/j.neubiorev.2006.06.005

Baier, L. J., Sacchettini, J. C., Knowler, W. C., Eads, J., Paolisso, G., Tataranni, P. A., et al. (1995). An Amino Acid Substitution in the Human Intestinal Fatty Acid Binding Protein Is Associated with Increased Fatty Acid Binding, Increased Fat Oxidation, and Insulin Resistance. *J. Clin. Invest.* 95 (3), 1281–1287. doi:10.1172/jci117778

Balenci, L., Clarke, I. D., Dirks, P. B., Assard, N., Ducray, F., Jouvet, A., et al. (2006). IQGAP1 Protein Specifies Amplifying Cancer Cells in Glioblastoma Multiforme. *Cancer Res.* 66 (18), 9074–9082. doi:10.1158/0008-5472.can-06-0761

Bantubungi, K., Hannou, S.-A., Caron-Houde, S., Vallez, E., Baron, M., Lucas, A., et al. (2014). Cdkn2a/p16Ink4a Regulates Fasting-Induced Hepatic Gluconeogenesis Through the PKA-CREB-PGC1 Pathway. *Diabetes* 63 (10), 3199–3209. doi:10.2337/db13-1921

Bao, L., Fu, X., Si, M., Wang, Y., Ma, R., Ren, X., et al. (2015). MicroRNA-185 Targets SOCS3 to Inhibit Beta-Cell Dysfunction in Diabetes. *PLoS One* 10 (2), e0116067. doi:10.1371/journal.pone.0116067

Barlow, J. H., and Rothstein, R. (2010). Timing Is Everything: Cell Cycle Control of Rad52. *Cell Div* 5, 7. doi:10.1186/1747-1028-5-7

Baunoch, D., Watkins, L., Tewari, A., Reece, M., Adams, L., Stack, R., et al. (1996). MDM2 Overexpression in Benign and Malignant Lesions of the Human Breast. *Int. J. Oncol.* 8 (5), 895–899. doi:10.3892/ijo.8.5.895

Beretta, S., Maj, C., and Merelli, I. (2017). Rank miRNA: A Web Tool for Identifying Polymorphisms Altering miRNA Target Sites. *Proced. Comp. Sci.* 108, 1125–1134. doi:10.1016/j.procs.2017.05.189

Bhattacharya, A., and Cui, Y. (2017). Systematic Prediction of the Impacts of Mutations in MicroRNA Seed Sequences. *J. Integr. Bioinform* 14 (1), 20170001. doi:10.1515/jib-2017-0001

Brekman, A., Singh, K. E., Polotskaia, A., Kundu, N., and Bargonetti, J. (2011). A P53-independent Role of Mdm2 in Estrogen-Mediated Activation of Breast Cancer Cell Proliferation. *Breast Cancer Res.* 13 (1), R3. doi:10.1186/bcr2804

Briggs, M. W., and Sacks, D. B. (2003). IQGAP1 as Signal Integrator: Ca2+, Calmodulin, Cdc42 and the Cytoskeleton. *FEBS Lett.* 542 (1-3), 7–11. doi:10.1016/s0014-5793(03)00333-8

Brincas, H. M., Augusto, D. G., Mathias, C., Cavalli, I. J., Lima, R. S., Kuroda, F., et al. (2020). A Genetic Variant in microRNA-146a Is Associated with Sporadic Breast Cancer in a Southern Brazilian Population. *Genet. Mol. Biol.* 42 (4), e20190278. doi:10.1590/1678-4685-GMB-2019-0278

Brown, M. D., and Sacks, D. B. (2006). IQGAP1 in Cellular Signaling: Bridging the GAP. *Trends Cel Biol.* 16 (5), 242–249. doi:10.1016/j.tcb.2006.03.002

Brucker, S. Y., Frank, L., Eisenbeis, S., Henes, M., Wallwiener, D., Riess, O., et al. (2017). Sequence Variants inESR1andOXTRare Associated with Mayer-Rokitansky-Küster-Hauser Syndrome. *Acta Obstet. Gynecol. Scand.* 96 (11), 1338–1346. doi:10.1111/aogs.13202

Bruno, A. E., Li, L., Kalabus, J. L., Pan, Y., Yu, A., and Hu, Z. (2012). miRdSNP: A Database of Disease-Associated SNPs and microRNA Target Sites on 3'UTRs of Human Genes. *BMC Genomics* 13 (1), 44. doi:10.1186/1471-2164-13-44

Cao, J., Luo, C., Peng, R., Guo, Q., Wang, K., Ye, H., et al. (2016). MiRNA-binding Site Functional Polymorphisms in DNA Repair Genes RAD51, RAD52, and XRCC2 and Breast Cancer Risk in Chinese Population. *Tumour Biol.* [Epub ahead of print]. doi:10.1007/s13277-016-5459-2

Catucci, I., Yang, R., Verderio, P., Pizzamiglio, S., Heesen, L., Hemminki, K., et al. (2010). Evaluation of SNPs inmiR-146a,miR196a2andmiR-499as Low-Penetrance Alleles in German and Italian Familial Breast Cancer Cases. *Hum. Mutat.* 31 (1), E1052–E1057. doi:10.1002/humu.21141

Ceolotto, G., Papparella, I., Bortoluzzi, A., Strapazzon, G., Ragazzo, F., Bratti, P., et al. (2011). Interplay Between miR-155, AT1R A1166C Polymorphism, and AT1R Expression in Young Untreated Hypertensives. *Am. J. Hypertens.* 24 (2), 241–246. doi:10.1038/ajh.2010.211

Chen, C., Yang, S., Chaugai, S., Wang, Y., and Wang, D. W. (2014). Meta-analysis of Hsa-Mir-499 Polymorphism (Rs3746444) for Cancer Risk: Evidence from 31 Case-Control Studies. *BMC Med. Genet.* 15 (1), 126. doi:10.1186/s12881-014-0126-1

Chen, G., Umelo, I. A., Lv, S., Teugels, E., Fostier, K., Kronenberger, P., et al. (2013). miR-146a Inhibits Cell Growth, Cell Migration and Induces Apoptosis in Non-small Cell Lung Cancer Cells. *PLoS One* 8 (3), e60317. doi:10.1371/journal.pone.0060317

Chen, H., Lan, H.-Y., Roukos, D. H., and Cho, W. C. (2014). Application of microRNAs in Diabetes Mellitus. *J. Endocrinol.* 222 (1), R1–R10. doi:10.1530/joe-13-0544

Chen, Q.-H., Wang, Q.-B., and Zhang, B. (2014). Ethnicity Modifies the Association Between Functional microRNA Polymorphisms and Breast Cancer Risk: A HuGE Meta-Analysis. *Tumor Biol.* 35 (1), 529–543. doi:10.1007/s13277-013-1074-7

Chen, Y., Huang, L., Qi, X., and Chen, C. (2019). Insulin Receptor Trafficking: Consequences for Insulin Sensitivity and Diabetes. *Int. J. Mol. Sci.* 20 (20), 5007. doi:10.3390/ijms20205007

Ciccacci, C., Latini, A., Greco, C., Politi, C., D'Amato, C., Lauro, D., et al. (2018). Association Between a MIR499A Polymorphism and Diabetic Neuropathy in Type 2 Diabetes. *J. Diabetes its Complications* 32 (1), 11–17. doi:10.1016/j.jdiacomp.2017.10.011

Coffinier, C., Thepot, D., Babinet, C., Yaniv, M., and Barra, J. (1999). Essential Role for the Homeoprotein vHNF1/HNF1beta in Visceral Endoderm Differentiation. *Development* 126 (21), 4785–4794. doi:10.1242/dev.126.21.4785

Coolen, M., and Bally-Cuif, L. (2015). "MicroRNAs in Brain Development," in *MicroRNA in Regenerative Medicine*. Editor C. K. Sen (Oxford, UK: Academic Press), 447–488. doi:10.1016/b978-0-12-405544-5.00018-6

Crocco, P., Montesanto, A., Passarino, G., and Rose, G. (2015). Polymorphisms Falling within Putative miRNA Target Sites in the 3′UTR Region ofSIRT2andDRD2Genes Are Correlated with Human Longevity. *Gerona* 71 (5), 586–592. doi:10.1093/gerona/glv058

Crosby, M. E., and Almasan, A. (2004). Opposing Roles of E2Fs in Cell Proliferation and Death. *Cancer Biol. Ther.* 3 (12), 1208–1211. doi:10.4161/cbt.3.12.1494

Curtale, G., Citarella, F., Carissimi, C., Goldoni, M., Carucci, N., Fulci, V., et al. (2010). An Emerging Player in the Adaptive Immune Response: microRNA-146a Is a Modulator of IL-2 Expression and Activation-Induced Cell Death in T Lymphocytes. *Blood* 115 (2), 265–273. doi:10.1182/blood-2009-06-225987

Dai, Z.-J., Shao, Y.-P., Wang, X.-J., Xu, D., Kang, H.-F., Ren, H.-T., et al. (2015). Five Common Functional Polymorphisms in microRNAs (Rs2910164, Rs2292832, Rs11614913, Rs3746444, Rs895819) and the Susceptibility to Breast Cancer: Evidence from 8361 Cancer Cases and 8504 Controls. *Cpd* 21 (11), 1455–1463. doi:10.2174/1381612821666141208143533

Dai, Z.-M., Kang, H.-F., Zhang, W.-G., Li, H.-B., Zhang, S.-Q., Ma, X.-B., et al. (2016). The Associations of Single Nucleotide Polymorphisms in miR196a2, miR-499, and miR-608 with Breast Cancer Susceptibility. *Medicine (Baltimore)* 95 (7), e2826. doi:10.1097/md.0000000000002826

de Carvalho, J. B., de Morais, G. L., Vieira, T. C. D. S., Rabelo, N. C., Llerena, J. C., Gonzalez, S. M. D. C., et al. (2019). miRNA Genetic Variants Alter Their Secondary Structure and Expression in Patients with RASopathies Syndromes. *Front. Genet.* 10, 1144. doi:10.3389/fgene.2019.01144

de Oliveira, R. M., Sarkander, J., Kazantsev, A. G., and Outeiro, T. F. (2012). SIRT2 as a Therapeutic Target for Age-Related Disorders. *Front. Pharmacol.* 3, 82. doi:10.3389/fphar.2012.00082

Dong, P., Nabeshima, K., Nishimura, N., Kawakami, T., Hachisuga, T., Kawarabayashi, T., et al. (2006). Overexpression and Diffuse Expression Pattern of IQGAP1 at Invasion Fronts Are Independent Prognostic Parameters in Ovarian Carcinomas. *Cancer Lett.* 243 (1), 120–127. doi:10.1016/j.canlet.2005.11.024

Dou, T., Wu, Q., Chen, X., Ribas, J., Ni, X., Tang, C., et al. (2010). A Polymorphism of microRNA196a Genome Region Was Associated with Decreased Risk of Glioma in Chinese Population. *J. Cancer Res. Clin. Oncol.* 136 (12), 1853–1859. doi:10.1007/s00432-010-0844-5

Duan, R., Pak, C., and Jin, P. (2007). Single Nucleotide Polymorphism Associated with Mature miR-125a Alters the Processing of Pri-miRNA. *Hum. Mol. Genet.* 16 (9), 1124–1131. doi:10.1093/hmg/ddm062

Elek, Z., Németh, N., Nagy, G., Németh, H., Somogyi, A., Hosszufalusi, N., et al. (2015). Micro-RNA Binding Site Polymorphisms in the WFS1 Gene Are Risk Factors of Diabetes Mellitus. *PLoS One* 10 (10), e0139519. doi:10.1371/journal.pone.0139519

Elfaki, I., Mir, R., Mir, M. M., AbuDuhier, F. M., Babakr, A. T., and Barnawi, J. (2019). Potential Impact of MicroRNA Gene Polymorphisms in the Pathogenesis of Diabetes and Atherosclerotic Cardiovascular Disease. *J. Pers Med.* 9 (4), 51. doi:10.3390/jpm9040051

Erturk, E., Cecener, G., Polatkan, V., Gokgoz, S., Egeli, U., Tunca, B., et al. (2014). Evaluation of Genetic Variations in miRNA-Binding Sites of BRCA1 and BRCA2 Genes as Risk Factors for the Development of Early-Onset And/or Familial Breast Cancer. *Asian Pac. J. Cancer Prev.* 15 (19), 8319–8324. doi:10.7314/apjcp.2014.15.19.8319

Fan, B., Chopp, M., Zhang, Z. G., and Liu, X. S. (2020). Emerging Roles of microRNAs as Biomarkers and Therapeutic Targets for Diabetic Neuropathy. *Front. Neurol.* 11, 558758. doi:10.3389/fneur.2020.558758

Fawcett, K. A., Wheeler, E., Morris, A. P., Ricketts, S. L., Hallmans, G., Rolandsson, O., et al. (2010). Detailed Investigation of the Role of Common and Low-Frequency WFS1 Variants in Type 2 Diabetes Risk. *Diabetes* 59 (3), 741–746. doi:10.2337/db09-0920

Fawzy, M. S., Toraih, E. A., Ibrahiem, A., Abdeldayem, H., Mohamed, A. O., and Abdel-Daim, M. M. (2017). Evaluation of miRNA-196a2 and Apoptosis-Related Target Genes: ANXA1, DFFA and PDCD4 Expression in Gastrointestinal Cancer Patients: A Pilot Study. *PloS one* 12 (11), e0187310. doi:10.1371/journal.pone.0187310

Feng, Z., Scott, S. P., Bussen, W., Sharma, G. G., Guo, G., Pandita, T. K., et al. (2011). Rad52 Inactivation Is Synthetically Lethal with BRCA2 Deficiency. *Proc. Natl. Acad. Sci.* 108 (2), 686–691. doi:10.1073/pnas.1010959107

Ferracin, M., Querzoli, P., Calin, G. A., and Negrini, M. (2011). MicroRNAs: Toward the Clinic for Breast Cancer Patients. *Semin. Oncol.* 38 (6), 764–775. doi:10.1053/j.seminoncol.2011.08.005

Ferretti, E., De Smaele, E., Po, A., Di Marcotullio, L., Tosi, E., Espinola, M. S. B., et al. (2009). MicroRNA Profiling in Human Medulloblastoma. *Int. J. Cancer* 124 (3), 568–577. doi:10.1002/ijc.23948

Fonseca, S. G., Fukuma, M., Lipson, K. L., Nguyen, L. X., Allen, J. R., Oka, Y., et al. (2005). WFS1 Is a Novel Component of the Unfolded Protein Response and Maintains Homeostasis of the Endoplasmic Reticulum in Pancreatic β-Cells. *J. Biol. Chem.* 280 (47), 39609–39615. doi:10.1074/jbc.m507426200

Fowler, M. J. (2008). Microvascular and Macrovascular Complications of Diabetes. *Clin. Diabetes* 26, 77–82. doi:10.2337/diaclin.26.2.77

Galka-Marciniak, P., Urbanek-Trzeciak, M. O., Nawrocka, P. M., Dutkiewicz, A., Giefing, M., Lewandowska, M. A., et al. (2019). Somatic Mutations in miRNA Genes in Lung Cancer-Potential Functional Consequences of Non-coding Sequence Variants. *Cancers* 11 (6), 793. doi:10.3390/cancers11060793

Gao, F., Xiong, X., Pan, W., Yang, X., Zhou, C., Yuan, Q., et al. (2015). A Regulatory MDM4 Genetic Variant Locating in the Binding Sequence of Multiple MicroRNAs Contributes to Susceptibility of Small Cell Lung Cancer. *PLOS ONE* 10 (8), e0135647. doi:10.1371/journal.pone.0135647

Gao, W., Gu, Y., Li, Z., Cai, H., Peng, Q., Tu, M., et al. (2015). miR-615-5p Is Epigenetically Inactivated and Functions as a Tumor Suppressor in Pancreatic Ductal Adenocarcinoma. *Oncogene* 34 (13), 1629–1640. doi:10.1038/onc.2014.101

George, G. P., Gangwar, R., Mandal, R. K., Sankhwar, S. N., and Mittal, R. D. (2011). Genetic Variation in microRNA Genes and Prostate Cancer Risk in North Indian Population. *Mol. Biol. Rep.* 38 (3), 1609–1615. doi:10.1007/s11033-010-0270-4

Ghaedi, H., Bastami, M., Jahani, M. M., Alipoor, B., Tabasinezhad, M., Ghaderi, O., et al. (2016). A Bioinformatics Approach to the Identification of Variants Associated with Type 1 and Type 2 Diabetes Mellitus that Reside in Functionally Validated miRNAs Binding Sites. *Biochem. Genet.* 54 (3), 211–221. doi:10.1007/s10528-016-9713-5

Gholami, M., Asgarbeik, S., Razi, F., Esfahani, E. N., Zoughi, M., Vahidi, A., et al. (2020). Association of microRNA Gene Polymorphisms with Type 2 Diabetes Mellitus: A Systematic Review and Meta-Analysis. *J. Res. Med. Sci.* 25, 56. doi:10.4103/jrms.JRMS_751_19

Gilam, A., Conde, J., Weissglas-Volkov, D., Oliva, N., Friedman, E., Artzi, N., et al. (2016). Local microRNA Delivery Targets Palladin and Prevents Metastatic Breast Cancer. *Nat. Commun.* 7 (1), 12868–12914. doi:10.1038/ncomms12868

Goda, N., Murase, H., Kasezawa, N., Goda, T., and Yamakawa-Kobayashi, K. (2015). Polymorphism in microRNA-Binding Site in HNF1B Influences the Susceptibility of Type 2 Diabetes Mellitus: A Population Based Case-Control Study. *BMC Med. Genet.* 16, 75. doi:10.1186/s12881-015-0219-5

Goicoechea, S. M., Bednarski, B., García-Mata, R., Prentice-Dunn, H., Kim, H. J., and Otey, C. A. (2009). Palladin Contributes to Invasive Motility in Human Breast Cancer Cells. *Oncogene* 28 (4), 587–598. doi:10.1038/onc.2008.408

Gong, J., Tong, Y., Zhang, H.-M., Wang, K., Hu, T., Shan, G., et al. (2012). Genome-wide Identification of SNPs in microRNA Genes and the SNP Effects on microRNA Target Binding and Biogenesis. *Hum. Mutat.* 33 (1), 254–263. doi:10.1002/humu.21641

Grillari, J., and Grillari-Voglauer, R. (2010). Novel Modulators of Senescence, Aging, and Longevity: Small Non-coding RNAs Enter the Stage. *Exp. Gerontol.* 45 (4), 302–311. doi:10.1016/j.exger.2010.01.007

Guo, Z., Shu, Y., Zhou, H., and Zhang, W. (2016). Identification of Diagnostic and Prognostic Biomarkers for Cancer: Focusing on Genetic Variations in microRNA Regulatory Pathways (Review). *Mol. Med. Rep.* 13 (3), 1943–1952. doi:10.3892/mmr.2016.4782

Han, S., Gao, F., Yang, W., Ren, Y., Liang, X., Xiong, X., et al. (2015). Identification of an SCLC Susceptibility Rs7963551 Genetic Polymorphism in a Previously GWAS-Identified 12p13.33 RAD52 Lung Cancer Risk Locus in the Chinese Population. *Int. J. Clin. Exp. Med.* 8 (9), 16528–16535.

Harel, T., Pehlivan, D., Caskey, C. T., and Lupski, J. R. (2015). "Mendelian, Non-mendelian, Multigenic Inheritance, and Epigenetics," in *Rosenberg's Molecular and Genetic Basis of Neurological and Psychiatric Disease*. Editors R. N. Rosenberg and J. M. Pascual. 5th Edn (Boston. MA: Academic Press), 3–27. doi:10.1016/b978-0-12-410529-4.00001-2

Hasani, S.-S., Hashemi, M., Eskandari-Nasab, E., Naderi, M., Omrani, M., and Sheybani-Nasab, M. (2014). A Functional Polymorphism in the miR-146a Gene Is Associated with the Risk of Childhood Acute Lymphoblastic Leukemia: A Preliminary Report. *Tumor Biol.* 35 (1), 219–225. doi:10.1007/s13277-013-1027-1

Hashimoto, Y., Akiyama, Y., and Yuasa, Y. (2013). Multiple-to-multiple Relationships Between microRNAs and Target Genes in Gastric Cancer. *PLoS One* 8 (5), e62589. doi:10.1371/journal.pone.0062589

Hatefi, Z., Soltani, G., Khosravi, S., Kazemi, M., Salehi, A. R., and Salehi, R. (2018). Micro R-410 Binding Site Single Nucleotide Polymorphism Rs13702 in Lipoprotein Lipase Gene Is Effective to Increase Susceptibility to Type 2 Diabetes in Iranian Population. *Adv. Biomed. Res.* 7, 79. doi:10.4103/abr.abr_286_16

Haunerland, N. H., and Spener, F. (2004). Fatty Acid-Binding Proteins - Insights from Genetic Manipulations. *Prog. Lipid Res.* 43 (4), 328–349. doi:10.1016/j.plipres.2004.05.001

Hiratsuka, M., Inoue, T., Toda, T., Kimura, N., Shirayoshi, Y., Kamitani, H., et al. (2003). Proteomics-based Identification of Differentially Expressed Genes in Human Gliomas: Down-Regulation of SIRT2 Gene. *Biochem. biophysical Res. Commun.* 309 (3), 558–566. doi:10.1016/j.bbrc.2003.08.029

Hoffman, A. E., Zheng, T., Yi, C., Leaderer, D., Weidhaas, J., Slack, F., et al. (2009). microRNA miR-196a-2 and Breast Cancer: A Genetic and Epigenetic Association Study and Functional Analysis. *Cancer Res.* 69 (14), 5970–5977. doi:10.1158/0008-5472.can-09-0236

Hofmann, S., Philbrook, C., Gerbitz, K.-D., and Bauer, M. F. (2003). Wolfram Syndrome: Structural and Functional Analyses of Mutant and Wild-type Wolframin, the WFS1 Gene Product. *Hum. Mol. Genet.* 12 (16), 2003–2012. doi:10.1093/hmg/ddg214

Hori, M., Shimazaki, J., Inagawa, S., Itabashi, M., and Hori, M. (2002). Overexpression of MDM2 Oncoprotein Correlates with Possession of Estrogen Receptor Alpha and Lack of MDM2 mRNA Splice Variants in Human Breast Cancer. *Breast Cancer Res. Treat.* 71 (1), 77–84. doi:10.1023/a:1013350419426

Hoss, A. G., Kartha, V. K., Dong, X., Latourelle, J. C., Dumitriu, A., Hadzi, T. C., et al. (2014). MicroRNAs Located in the Hox Gene Clusters Are Implicated in huntington's Disease Pathogenesis. *Plos Genet.* 10 (2), e1004188. doi:10.1371/journal.pgen.1004188

Hribal, M. L., au, fnm., Presta, I., Procopio, T., Marini, M. A., Stančáková, A., et al. (2011). Glucose Tolerance, Insulin Sensitivity and Insulin Release in European Non-diabetic Carriers of a Polymorphism Upstream of CDKN2A and CDKN2B. *Diabetologia* 54 (4), 795–802. doi:10.1007/s00125-010-2038-8

Hu, Z., Chen, J., Tian, T., Zhou, X., Gu, H., Xu, L., et al. (2008). Genetic Variants of miRNA Sequences and Non-small Cell Lung Cancer Survival. *J. Clin. Invest.* 118 (7), 2600–2608. doi:10.1172/JCI34934

Hu, Z., Liang, J., Wang, Z., Tian, T., Zhou, X., Chen, J., et al. (2009). Common Genetic Variants in Pre-microRNAs Were Associated with Increased Risk of Breast Cancer in Chinese Women. *Hum. Mutat.* 30 (1), 79–84. doi:10.1002/humu.20837

Huan, T., Chen, G., Liu, C., Bhattacharya, A., Rong, J., Chen, B. H., et al. (2018). Age-associated microRNA Expression in Human Peripheral Blood Is Associated with All-Cause Mortality and Age-Related Traits. *Aging Cell* 17 (1), e12687. doi:10.1111/acel.12687

Hulf, T., Sibbritt, T., Wiklund, E. D., Bert, S., Strbenac, D., Statham, A. L., et al. (2011). Discovery Pipeline for Epigenetically Deregulated miRNAs in Cancer: Integration of Primary miRNA Transcription. *BMC genomics* 12 (1), 54–59. doi:10.1186/1471-2164-12-54

Iguchi, T., Nambara, S., Masuda, T., Komatsu, H., Ueda, M., Kidogami, S., et al. (2016). miR-146a Polymorphism (Rs2910164) Predicts Colorectal Cancer Patients' Susceptibility to Liver Metastasis. *PLOS ONE* 11 (11), e0165912. doi:10.1371/journal.pone.0165912

Inoue, T., Hiratsuka, M., Osaki, M., and Oshimura, M. (2007). The Molecular Biology of Mammalian SIRT Proteins: SIRT2 Functions on Cell Cycle Regulation. *Cell Cycle* 6 (9), 1011–1018. doi:10.4161/cc.6.9.4219

Iorio, M. V., and Croce, C. M. (2012). microRNA Involvement in Human Cancer. *Carcinogenesis* 33 (6), 1126–1133. doi:10.1093/carcin/bgs140

Jaakson, H., Karis, P., Ling, K., Ilves-Luht, A., Samarütel, J., Henno, M., et al. (2018). Adipose Tissue Insulin Receptor and Glucose Transporter 4 Expression, and Blood Glucose and Insulin Responses during Glucose Tolerance Tests in Transition Holstein Cows with Different Body Condition. *J. Dairy Sci.* 101 (1), 752–766. doi:10.3168/jds.2017-12877

Jeon, Y. J., Kim, O. J., Kim, S. Y., Oh, S. H., Oh, D., Kim, O. J., et al. (2013). Association of the miR-146a , miR-149 , miR-196a2 , and miR-499 Polymorphisms with Ischemic Stroke and Silent Brain Infarction Risk. *Arterioscler Thromb. Vasc. Biol.* 33 (2), 420–430. doi:10.1161/atvbaha.112.300251

Jiang, Y., Qin, Z., Hu, Z., Guan, X., Wang, Y., He, Y., et al. (2013). Genetic Variation in a Hsa-Let-7 Binding Site in RAD52 Is Associated with Breast Cancer Susceptibility. *Carcinogenesis* 34 (3), 689–693. doi:10.1093/carcin/bgs373

Jin, Y., Zeng, S. X., Sun, X.-X., Lee, H., Blattner, C., Xiao, Z., et al. (2008). MDMX Promotes Proteasomal Turnover of P21 at G 1 and Early S Phases Independently of, but in Cooperation with, MDM2. *Mol. Cel Biol* 28 (4), 1218–1229. doi:10.1128/mcb.01198-07

Johnson, M., Sharma, M., and Henderson, B. R. (2009). IQGAP1 Regulation and Roles in Cancer. *Cell Signal.* 21 (10), 1471–1478. doi:10.1016/j.cellsig.2009.02.023

Kim, H.-S., Vassilopoulos, A., Wang, R.-H., Lahusen, T., Xiao, Z., Xu, X., et al. (2011). SIRT2 Maintains Genome Integrity and Suppresses Tumorigenesis Through Regulating APC/C Activity. *Cancer cell* 20 (4), 487–499. doi:10.1016/j.ccr.2011.09.004

Kim, J. K., Fillmore, J. J., Chen, Y., Yu, C., Moore, I. K., Pypaert, M., et al. (2001). Tissue-specific Overexpression of Lipoprotein Lipase Causes Tissue-specific Insulin Resistance. *Proc. Natl. Acad. Sci.* 98 (13), 7522–7527. doi:10.1073/pnas.121164498

Klöting, N., Berthold, S., Kovacs, P., Schön, M. R., Fasshauer, M., Ruschke, K., et al. (2009). MicroRNA Expression in Human Omental and Subcutaneous Adipose Tissue. *PLoS One* 4 (3), e4699. doi:10.1371/journal.pone.0004699

Kovacs-Nagy, R., Elek, Z., Szekely, A., Nanasi, T., Sasvari-Szekely, M., and Ronai, Z. (2013). Association of Aggression with a Novel microRNA Binding Site Polymorphism in the Wolframin Gene. *Am. J. Med. Genet.* 162 (4), 404–412. doi:10.1002/ajmg.b.32157

Krishnamurthy, J., Ramsey, M. R., Ligon, K. L., Torrice, C., Koh, A., Bonner-Weir, S., et al. (2006). p16INK4a Induces an Age-dependent Decline in Islet Regenerative Potential. *Nature* 443 (7110), 453–457. doi:10.1038/nature05092

Labbaye, C., and Testa, U. (2012). The Emerging Role of MIR-146A in the Control of Hematopoiesis, Immune Function and Cancer. *J. Hematol. Oncol.* 5 (1), 13. doi:10.1186/1756-8722-5-13

Laragh, J., and Pickering, T. (1991). "Essential Hypertension," in *Brenner & Rector's the Kidney*. Editors B. M. Brenner and F. C. Rector (Philadelphia, PA: W.B. Saunders Company), 1913–1967.

Latini, A., Borgiani, P., De Benedittis, G., D'Amato, C., Greco, C., Lauro, D., et al. (2020). Mitochondrial DNA Copy Number in Peripheral Blood Is Reduced in Type 2 Diabetes Patients with Polyneuropathy and Associated with a MIR499A Gene Polymorphism. *DNA Cel Biol.* 39 (8), 1467–1472. doi:10.1089/dna.2019.5326

Lee, M. J., Yu, G. R., Yoo, H. J., Kim, J. H., Yoon, B. I., Choi, Y. K., et al. (2009). ANXA8 Down-Regulation by EGF-FOXO4 Signaling Is Involved in Cell Scattering and Tumor Metastasis of Cholangiocarcinoma. *Gastroenterology* 137 (3), 1138–1150.e1–e9. doi:10.1053/j.gastro.2009.04.015

Lehmann, T. P., Korski, K., Ibbs, M., Zawierucha, P., Grodecka-gazdecka, S., and Jagodziński, P. P. (2013). rs12976445 Variant in the Pri-miR-125a Correlates with a Lower Level of Hsa-miR-125a and ERBB2 Overexpression in Breast Cancer Patients. *Oncol. Lett.* 5 (2), 569–573. doi:10.3892/ol.2012.1040

Leong, P. L. (2013). *Single Nucleotide Polymorphism in the Seed Region of microRNA Alters the Expression of its Mature microRNA/Leong Pei Li*. Kuala Lumpur, Malaysia: University of Malaya.

Li, N., Chen, J., Zhao, J., and Wang, T. (2017). MicroRNA-3188 Targets ETS-Domain Protein 4 and Participates in RhoA/ROCK Pathway to Regulate the Development of Atherosclerosis. *Pharmazie* 72 (11), 687–693. doi:10.1691/ph.2017.7686

Li, X., Wang, J., Jia, Z., Cui, Q., Zhang, C., Wang, W., et al. (2013). MiR-499 Regulates Cell Proliferation and Apoptosis during Late-Stage Cardiac Differentiation via Sox6 and Cyclin D1. *PLOS ONE* 8 (9), e74504. doi:10.1371/journal.pone.0074504

Li, Z., Xie, Q. R., Chen, Z., Lu, S., and Xia, W. (2013). Regulation of SIRT2 Levels for Human Non-small Cell Lung Cancer Therapy. *Lung Cancer* 82 (1), 9–15. doi:10.1016/j.lungcan.2013.05.013

Lieberman, R., Xiong, D., James, M., Han, Y., Amos, C. I., Wang, L., et al. (2016). Functional Characterization of RAD52 as a Lung Cancer Susceptibility Gene in the 12p13.33 Locus. *Mol. Carcinog.* 55 (5), 953–963. doi:10.1002/mc.22334

Ling, H., Zhang, W., and Calin, G. A. (2011). Principles of microRNA Involvement in Human Cancers. *Chin. J. Cancer* 30 (11), 739–748. doi:10.5732/cjc.011.10243

Liu, F.-J., Wen, T., and Liu, L. (2012). MicroRNAs as a Novel Cellular Senescence Regulator. *Ageing Res. Rev.* 11 (1), 41–50. doi:10.1016/j.arr.2011.06.001

Liu, G., Schwartz, J. A., and Brooks, S. C. (2000). Estrogen Receptor Protects P53 from Deactivation by Human Double Minute-2. *Cancer Res.* 60 (7), 1810–1814.

Liu, X., Zhang, Z., Sun, L., Chai, N., Tang, S., Jin, J., et al. (2011). MicroRNA-499-5p Promotes Cellular Invasion and Tumor Metastasis in Colorectal Cancer by Targeting FOXO4 and PDCD4. *Carcinogenesis* 32 (12), 1798–1805. doi:10.1093/carcin/bgr213

Liu, X., Zhao, J., Liu, Q., Xiong, X., Zhang, Z., Jiao, Y., et al. (2016). MicroRNA-124 Promotes Hepatic Triglyceride Accumulation Through Targeting Tribbles Homolog 3. *Sci. Rep.* 6, 37170. doi:10.1038/srep37170

Liu, Y., He, A., Liu, B., Zhong, Y., Liao, X., Yang, J., et al. (2018). rs11614913 Polymorphism in miRNA-196a2 and Cancer Risk: an Updated Meta-Analysis. *Ott* 11, 1121–1139. doi:10.2147/ott.s154211

Liu, Y., Li, M.-j., Lee, E. Y.-H. P., and Maizels, N. (1999). Localization and Dynamic Relocalization of Mammalian Rad52 During the Cell Cycle and in Response to DNA Damage. *Curr. Biol.* 9 (17), 975–978. doi:10.1016/s0960-9822(99)80427-8

Liu, Y., Ma, Y., Zhang, B., Wang, S.-X., Wang, X.-M., and Yu, J.-M. (2014). Genetic Polymorphisms in Pre-microRNAs and Risk of Ischemic Stroke in a Chinese Population. *J. Mol. Neurosci.* 52 (4), 473–480. doi:10.1007/s12031-013-0152-z

Liu, Z., Li, G., Wei, S., Niu, J., El-Naggar, A. K., Sturgis, E. M., et al. (2010). Genetic Variants in Selected Pre-microRNA Genes and the Risk of Squamous Cell Carcinoma of the Head and Neck. *Cancer* 116 (20), 4753–4760. doi:10.1002/cncr.25323

Lok, B. H., Carley, A. C., Tchang, B., and Powell, S. N. (2013). RAD52 Inactivation Is Synthetically Lethal with Deficiencies in BRCA1 and PALB2 in Addition to BRCA2 through RAD51-Mediated Homologous Recombination. *Oncogene* 32 (30), 3552–3558. doi:10.1038/onc.2012.391

Lu, L.-F., Boldin, M. P., Chaudhry, A., Lin, L.-L., Taganov, K. D., Hanada, T., et al. (2010). Function of miR-146a in Controlling Treg Cell-Mediated Regulation of Th1 Responses. *Cell* 142 (6), 914–929. doi:10.1016/j.cell.2010.08.012

Luthra, R., Singh, R. R., Luthra, M. G., Li, Y. X., Hannah, C., Romans, A. M., et al. (2008). MicroRNA-196a Targets Annexin A1: A microRNA-Mediated Mechanism of Annexin A1 Downregulation in Cancers. *Oncogene* 27 (52), 6667–6678. doi:10.1038/onc.2008.256

Malhotra, P., Read, G. H., and Weidhaas, J. B. (2019). Breast Cancer and miR-SNPs: The Importance of miR Germ-Line Genetics. *Noncoding RNA* 5 (1), 27. doi:10.3390/ncrna5010027

Malhotra, P., Read, G., and Weidhaas, J. (2019). Breast Cancer and miR-SNPs: The Importance of miR Germ-Line Genetics. *ncRNA* 5 (1), 27. doi:10.3390/ncrna5010027

Marczak, E. D., Marzec, J., Zeldin, D. C., Kleeberger, S. R., Brown, N. J., Pretorius, M., et al. (2012). Polymorphisms in the Transcription Factor NRF2 and Forearm Vasodilator Responses in Humans. *Pharmacogenetics and genomics* 22 (8), 620–628. doi:10.1097/fpc.0b013e32835516e5

Mariella, E., Marotta, F., Grassi, E., Gilotto, S., and Provero, P. (2019). The Length of the Expressed 3' UTR Is an Intermediate Molecular Phenotype Linking Genetic Variants to Complex Diseases. *Front. Genet.* 10 (714), 714. doi:10.3389/fgene.2019.00714

Marine, J.-C., Francoz, S., Maetens, M., Wahl, G., Toledo, F., and Lozano, G. (2006). Keeping P53 in Check: Essential and Synergistic Functions of Mdm2 and Mdm4. *Cell Death Differ* 13 (6), 927–934. doi:10.1038/sj.cdd.4401912

Markey, M. P. (2011). Regulation of MDM4. *Front. Biosci.* 16, 1144–1156. doi:10.2741/3780

Martin, M. M., Buckenberger, J. A., Jiang, J., Malana, G. E., Nuovo, G. J., Chotani, M., et al. (2007). The Human Angiotensin II Type 1 Receptor +1166 A/C Polymorphism Attenuates microRNA-155 Binding. *J. Biol. Chem.* 282 (33), 24262–24269. doi:10.1074/jbc.m701050200

Marzec, J. M., Christie, J. D., Reddy, S. P., Jedlicka, A. E., Vuong, H., Lanken, P. N., et al. (2007). Functional Polymorphisms in the Transcription Factor NRF2 in Humans Increase the Risk of Acute Lung Injury. *FASEB j.* 21 (9), 2237–2246. doi:10.1096/fj.06-7759com

Matijasevic, Z., Steinman, H. A., Hoover, K., and Jones, S. N. (2008). MdmX Promotes Bipolar Mitosis to Suppress Transformation and Tumorigenesis in P53-Deficient Cells and Mice. *Mol. Cel Biol* 28 (4), 1265–1273. doi:10.1128/mcb.01108-07

Mead, J., Irvine, S., and Ramji, D. (2002). Lipoprotein Lipase: Structure, Function, Regulation, and Role in Disease. *J. Mol. Med.* 80 (12), 753–769. doi:10.1007/s00109-002-0384-9

Michan, S., and Sinclair, D. (2007). Sirtuins in Mammals: Insights into Their Biological Function. *Biochem. J.* 404 (1), 1–13. doi:10.1042/bj20070140

Morales, S., De Mayo, T., Gulppi, F., Gonzalez-Hormazabal, P., Carrasco, V., Reyes, J., et al. (2018). Genetic Variants in Pre-miR-146a, Pre-miR-499, Pre-miR-125a, Pre-miR-605, and Pri-miR-182 Are Associated with Breast Cancer Susceptibility in a South American Population. *Genes* 9 (9), 427. doi:10.3390/genes9090427

Moszyńska, A., Gebert, M., Collawn, J. F., and Bartoszewski, R. (2017). SNPs in microRNA Target Sites and Their Potential Role in Human Disease. *Open Biol.* 7 (4), 170019. doi:10.1098/rsob.170019

Moszynska, A., Gebert, M., Collawn, J. F., and Bartoszewski, R. (2017). SNPs in microRNA Target Sites and Their Potential Role in Human Disease. *Open Biol.* 7 (4), 170019. doi:10.1098/rsob.170019

Nabeshima, K., Shimao, Y., Inoue, T., and Koono, M. (2002). Immunohistochemical Analysis of IQGAP1 Expression in Human Colorectal Carcinomas: Its Overexpression in Carcinomas and Association with Invasion Fronts. *Cancer Lett.* 176 (1), 101–109. doi:10.1016/s0304-3835(01)00742-x

Nahand, J. S., Karimzadeh, M. R., Nezamnia, M., Fatemipour, M., Khatami, A., Jamshidi, S., et al. (2020). The Role of miR-146a in Viral Infection. *IUBMB Life* 72 (3), 343–360. doi:10.1002/iub.2222

Nakamura, H., Fujita, K., Nakagawa, H., Kishi, F., Takeuchi, A., Aute, I., et al. (2005). Expression Pattern of the Scaffold Protein IQGAP1 in Lung Cancer. *Oncol. Rep.* 13 (3), 427–431. doi:10.3892/or.13.3.427

Nicoloso, M. S., Sun, H., Spizzo, R., Kim, H., Wickramasinghe, P., Shimizu, M., et al. (2010). Single-nucleotide Polymorphisms Inside microRNA Target Sites Influence Tumor Susceptibility. *Cancer Res.* 70 (7), 2789–2798. doi:10.1158/0008-5472.can-09-3541

Noble, E. P. (2003). D2 Dopamine Receptor Gene in Psychiatric and Neurologic Disorders and its Phenotypes. *Am. J. Med. Genet.* 116B (1), 103–125. doi:10.1002/ajmg.b.10005

Nogueira, A., Fernandes, M., Catarino, R., and Medeiros, R. (2019). RAD52 Functions in Homologous Recombination and its Importance on Genomic Integrity Maintenance and Cancer Therapy. *Cancers* 11 (11), 1622. doi:10.3390/cancers11111622

Nojima, H., Adachi, M., Matsui, T., Okawa, K., Tsukita, S., and Tsukita, S. (2008). IQGAP3 Regulates Cell Proliferation Through the Ras/ERK Signalling cascade. *Nat. Cel Biol* 10 (8), 971–978. doi:10.1038/ncb1757

North, B. J., Marshall, B. L., Borra, M. T., Denu, J. M., and Verdin, E. (2003). The Human Sir2 Ortholog, SIRT2, Is an NAD+-dependent Tubulin Deacetylase. *Mol. Cel.* 11 (2), 437–444. doi:10.1016/s1097-2765(03)00038-8

Nussinov, R., Zhang, M., Tsai, C.-J., and Jang, H. (2018). Calmodulin and IQGAP1 Activation of PI3Kα and Akt in KRAS, HRAS and NRAS-Driven Cancers. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1864 (6, Part B), 2304–2314. doi:10.1016/j.bbadis.2017.10.032

O'Brien, J., Hayder, H., Zayed, Y., and Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. *Front. Endocrinol. (Lausanne)* 9 (402), 402. doi:10.3389/fendo.2018.00402

Ohnheiser, J., Ferlemann, E., Haas, A., Müller, J. P., Werwein, E., Fehler, O., et al. (2015). Programmed Cell Death 4 Protein (Pdcd4) and Homeodomain-Interacting Protein Kinase 2 (Hipk2) Antagonistically Control Translation of Hipk2 mRNA. *Biochim. Biophys. Acta (Bba) - Mol. Cel Res.* 1853 (7), 1564–1573. doi:10.1016/j.bbamcr.2015.03.008

Okubo, M., Tahara, T., Shibata, T., Yamashita, H., Nakamura, M., Yoshioka, D., et al. (2011). Association Study of Common Genetic Variants in Pre-microRNAs in Patients with Ulcerative Colitis. *J. Clin. Immunol.* 31 (1), 69–73. doi:10.1007/s10875-010-9461-y

Olefsky, J. M. (2001). Nuclear Receptor Minireview Series. *J. Biol. Chem.* 276 (40), 36863–36864. doi:10.1074/jbc.r100047200

Omrani, M., Hashemi, M., Eskandari-Nasab, E., Hasani, S.-S., Mashhadi, M. A., Arbabi, F., et al. (2014). hsa-mir-499 Rs3746444 Gene Polymorphism Is Associated with Susceptibility to Breast Cancer in an Iranian Population. *Biomarkers Med.* 8 (2), 259–267. doi:10.2217/bmm.13.118

Oparil, S., and Weber, M. A. (2000). *Hypertension: A Companion to Brenner & Rector's the Kidney*. Philadelphia, PA: W. B. Saunders Company.

Pal, D., Ghatak, S., and Sen, C. K. (2015). "Epigenetic Modification of MicroRNAs," in *MicroRNA in Regenerative Medicine*. Editor C. K. Sen (Amsterdam, Netherlands: Elsevier), 77–109. doi:10.1016/b978-0-12-405544-5.00003-4

Palmero, E. I., Campos, S. G. P. d., Campos, M., Souza, N. C. N. d., Guerreiro, I. D. C., Carvalho, A. L., et al. (2011). Mechanisms and Role of microRNA Deregulation in Cancer Onset and Progression. *Genet. Mol. Biol.* 34 (3), 363–370. doi:10.1590/s1415-47572011000300001

Park, D. H., Jeon, H. S., Lee, S. Y., Choi, Y. Y., Lee, H. W., Yoon, S., et al. (2015). MicroRNA-146a Inhibits Epithelial Mesenchymal Transition in Non-small

Cell Lung Cancer by Targeting Insulin Receptor Substrate 2. *Int. J. Oncol.* 47 (4), 1545–1553. doi:10.3892/ijo.2015.3111

Pelletier, C., Speed, W. C., Paranjape, T., Keane, K., Blitzblau, R., Hollestelle, A., et al. (2011). RareBRCA1haplotypes Including 3'UTR SNPs Associated with Breast Cancer Risk. *Cell Cycle* 10 (1), 90–99. doi:10.4161/cc.10.1.14359

Peng, Y., and Croce, C. M. (2016). The Role of MicroRNAs in Human Cancer. *Sig Transduct Target. Ther.* 1 (1), 15004. doi:10.1038/sigtrans.2015.4

Peterlongo, P., Caleca, L., Cattaneo, E., Ravagnani, F., Bianchi, T., Galastri, L., et al. (2011). The Rs12975333 Variant in the miR-125a and Breast Cancer Risk in Germany, Italy, Australia and Spain. *J. Med. Genet.* 48 (10), 703–704. doi:10.1136/jmedgenet-2011-100103

Peyrou, M., Bourgoin, L., Poher, A.-L., Altirriba, J., Maeder, C., Caillon, A., et al. (2015). Hepatic PTEN Deficiency Improves Muscle Insulin Sensitivity and Decreases Adiposity in Mice. *J. Hepatol.* 62 (2), 421–429. doi:10.1016/j.jhep.2014.09.012

Qi, P., Wang, L., Zhou, B., Yao, W. J., Xu, S., Zhou, Y., et al. (2015). Associations of miRNA Polymorphisms and Expression Levels with Breast Cancer Risk in the Chinese Population. *Genet. Mol. Res.* 14 (2), 6289–6296. doi:10.4238/2015.june.11.2

Rademakers, R., Eriksen, J. L., Baker, M., Robinson, T., Ahmed, Z., Lincoln, S. J., et al. (2008). Common Variation in the miR-659 Binding-Site of GRN Is a Major Risk Factor for TDP43-Positive Frontotemporal Dementia. *Hum. Mol. Genet.* 17 (23), 3631–3642. doi:10.1093/hmg/ddn257

Rahim, A., Afzal, M., and Naveed, A. K. (2019). Genetic Polymorphism of miRNA-196a and its Target Gene Annexin-A1 Expression Based on Ethnicity in Pakistani Female Breast Cancer Patients. *Pak J. Med. Sci.* 35 (6), 1598–1604. doi:10.12669/pjms.35.6.1322

Richardson, K., Nettleton, J. A., Rotllan, N., Tanaka, T., Smith, C. E., Lai, C.-Q., et al. (2013). Gain-of-function Lipoprotein Lipase Variant Rs13702 Modulates Lipid Traits Through Disruption of a microRNA-410 Seed Site. *Am. J. Hum. Genet.* 92 (1), 5–14. doi:10.1016/j.ajhg.2012.10.020

Ristori, E., and Nicoli, S. (2017). "Comparative Functions of miRNAs in Embryonic Neurogenesis and Neuronal Network Formation," in *Essentials of Noncoding RNA in Neuroscience*. Editor D. D. P. Tonelli (Amsterdam, Netherlands: Elsevier), 265–282. doi:10.1016/b978-0-12-804402-5.00015-7

Roy, M., Li, Z., and Sacks, D. B. (2005). IQGAP1 Is a Scaffold for Mitogen-Activated Protein Kinase Signaling. *Mol. Cel Biol* 25 (18), 7940–7952. doi:10.1128/mcb.25.18.7940-7952.2005

Rusca, N., and Monticelli, S. (2011). MiR-146a in Immunity and Disease. *Mol. Biol. Int.* 2011, 437301. doi:10.4061/2011/437301

Ruttan, C. C., and Glickman, B. W. (2002). Coding Variants in Human Double-Strand Break DNA Repair Genes. *Mutat. Res.* 509 (1-2), 175–200. doi:10.1016/s0027-5107(02)00218-x

Schaefer, A. S., Richter, G. M., Nothnagel, M., Laine, M. L., Rühling, A., Schäfer, C., et al. (2010). A 3′ UTR Transition within DEFB1 Is Associated with Chronic and Aggressive Periodontitis. *Genes Immun.* 11 (1), 45–54. doi:10.1038/gene.2009.75

Schmidt, P. S., Duvernell, D. D., and Eanes, W. F. (2000). Adaptive Evolution of a Candidate Gene for Aging in Drosophila. *Proc. Natl. Acad. Sci.* 97 (20), 10861–10865. doi:10.1073/pnas.190338897

Schmidt, V. A., Chiariello, C. S., Capilla, E., Miller, F., and Bahou, W. F. (2008). Development of Hepatocellular Carcinoma in Iqgap2 -Deficient Mice Is IQGAP1 Dependent. *Mol. Cel Biol* 28 (5), 1489–1502. doi:10.1128/mcb.01090-07

Scott, G. K., Goga, A., Bhaumik, D., Berger, C. E., Sullivan, C. S., and Benz, C. C. (2007). Coordinate Suppression of ERBB2 and ERBB3 by Enforced Expression of Micro-RNA miR-125a or miR-125b. *J. Biol. Chem.* 282 (2), 1479–1486. doi:10.1074/jbc.m609383200

Scutt, G., Overall, A., Bakrania, P., Krasteva, E., Parekh, N., Ali, K., et al. (2020). The Association of a Single-Nucleotide Polymorphism in the Nuclear Factor (Erythroid-Derived 2)-Like 2 Gene with Adverse Drug Reactions, Multimorbidity, and Frailty in Older People. *The Journals Gerontol. Ser. A* 75 (6), 1050–1057. doi:10.1093/gerona/glz131

Sebastiani, G., Po, A., Miele, E., Ventriglia, G., Ceccarelli, E., Bugliani, M., et al. (2015). MicroRNA-124a Is Hyperexpressed in Type 2 Diabetic Human Pancreatic Islets and Negatively Regulates Insulin Secretion. *Acta Diabetol.* 52 (3), 523–530. doi:10.1007/s00592-014-0675-y

Sethupathy, P., Borel, C., Gagnebin, M., Grant, G. R., Deutsch, S., Elton, T. S., et al. (2007). Human microRNA-155 on Chromosome 21 Differentially Interacts with its Polymorphic Target in the AGTR1 3′ Untranslated Region: A Mechanism for Functional Single-Nucleotide Polymorphisms Related to Phenotypes. *Am. J. Hum. Genet.* 81 (2), 405–413. doi:10.1086/519979

Shankaran, Z. S., Walter, C. E. J., Ramachandiran, K., Gurramkonda, V. B., and Johnson, T. (2020). Association of microRNA-146a Rs2910164 Polymorphism with Type II Diabetes Mellitus in a South Indian Population and a Meta-Analysis. *Gene Rep.* 18, 100567. doi:10.1016/j.genrep.2019.100567

Shimoyama, Y., Mitsuda, Y., Tsuruta, Y., Hamajima, N., and Niwa, T. (2014). Polymorphism of Nrf2, an Antioxidative Gene, Is Associated with Blood Pressure and Cardiovascular Mortality in Hemodialysis Patients. *Int. J. Med. Sci.* 11 (7), 726–731. doi:10.7150/ijms.8590

Shvarts, A., Steegenga, W. T., Riteco, N., van Laar, T., Dekker, P., Bazuine, M., et al. (1996). MDMX: A Novel P53-Binding Protein with Some Functional Properties of MDM2. *EMBO J.* 15 (19), 5349–5357. doi:10.1002/j.1460-2075.1996.tb00919.x

Si, W., Shen, J., Zheng, H., and Fan, W. (2019). The Role and Mechanisms of Action of microRNAs in Cancer Drug Resistance. *Clin. Epigenet* 11 (1), 25. doi:10.1186/s13148-018-0587-8

Slaby, O., Bienertova-Vasku, J., Svoboda, M., and Vyzula, R. (2012). Genetic Polymorphisms and microRNAs: New Direction in Molecular Epidemiology of Solid Cancer. *J. Cel Mol Med* 16 (1), 8–21. doi:10.1111/j.1582-4934.2011.01359.x

Solito, E., De Coupade, C., Canaider, S., Goulding, N. J., and Perretti, M. (2001). Transfection of Annexin 1 in Monocytic Cells Produces a High Degree of Spontaneous and Stimulated Apoptosis Associated with Caspase-3 Activation. *Br. J. Pharmacol.* 133 (2), 217–228. doi:10.1038/sj.bjp.0704054

Sommer, S., and Fuqua, S. A. W. (2001). Estrogen Receptor and Breast Cancer. *Semin. Cancer Biol.* 11 (5), 339–352. doi:10.1006/scbi.2001.0389

Sondermeijer, B. M., Bakker, A., Halliani, A., de Ronde, M. W. J., Marquart, A. A., Tijsen, A. J., et al. (2011). Platelets in Patients with Premature Coronary Artery Disease Exhibit Upregulation of miRNA340* and miRNA624*. *PloS one* 6 (10), e25946. doi:10.1371/journal.pone.0025946

Song, F.-J., and Chen, K.-X. (2011). Single-nucleotide Polymorphisms Among microRNA: Big Effects on Cancer. *Chin. J. Cancer* 30 (6), 381–391. doi:10.5732/cjc.011.10142

Sotiriou, S. K., Kamileri, I., Lugli, N., Evangelou, K., Da-Ré, C., Huber, F., et al. (2016). Mammalian RAD52 Functions in Break-Induced Replication Repair of Collapsed DNA Replication forks. *Mol. Cel.* 64 (6), 1127–1134. doi:10.1016/j.molcel.2016.10.038

Srivastava, K., Srivastava, A., and Mittal, B. (2010). Common Genetic Variants in Pre-microRNAs and Risk of Gallbladder Cancer in North Indian Population. *J. Hum. Genet.* 55 (8), 495–499. doi:10.1038/jhg.2010.54

Stanelle, J., and Pützer, B. M. (2006). E2F1-induced Apoptosis: Turning Killers into Therapeutics. *Trends Mol. Med.* 12 (4), 177–185. doi:10.1016/j.molmed.2006.02.002

Stanford, J. L., Szklo, M., and Brinton, L. A. (1986). Estrogen Receptors and Breast Cancer. *Epidemiol. Rev.* 8, 42–59. doi:10.1093/oxfordjournals.epirev.a036295

Stegeman, S., Moya, L., Selth, L. A., Spurdle, A. B., Clements, J. A., and Batra, J. (2015). A Genetic Variant of MDM4 Influences Regulation by Multiple microRNAs in Prostate Cancer. *Endocr. Relat. Cancer* 22 (2), 265–276. doi:10.1530/erc-15-0013

Stenholm, L., Stoehlmacher-Williams, J., Al-Batran, S. E., Heussen, N., Akin, S., Pauligk, C., et al. (2013). Prognostic Role of microRNA Polymorphisms in Advanced Gastric Cancer: A Translational Study of the Arbeitsgemeinschaft Internistische Onkologie (AIO). *Ann. Oncol.* 24 (10), 2581–2588. doi:10.1093/annonc/mdt330

Strachan, G. D., Jordan-Sciutto, K. L., Rallapalli, R., Tuan, R. S., and Hall, D. J. (2003). The E2F-1 Transcription Factor Is Negatively Regulated by its Interaction with the MDMX Protein. *J. Cel. Biochem.* 88 (3), 557–568. doi:10.1002/jcb.10318

Sun, T., Aceto, N., Meerbrey, K. L., Kessler, J. D., Zhou, C., Migliaccio, I., et al. (2011). Activation of Multiple Proto-Oncogenic Tyrosine Kinases in Breast Cancer via Loss of the PTPN12 Phosphatase. *Cell* 144 (5), 703–718. doi:10.1016/j.cell.2011.02.003

Sun, Y.-M., Lin, K.-Y., and Chen, Y.-Q. (2013). Diverse Functions of miR-125 Family in Different Cell Contexts. *J. Hematol. Oncol.* 6, 6. doi:10.1186/1756-8722-6-6

Sussan, T. E., Jun, J., Thimmulappa, R., Bedja, D., Antero, M., Gabrielson, K. L., et al. (2008). Disruption of Nrf2, A Key Inducer of Antioxidant Defenses, Attenuates ApoE-Mediated Atherosclerosis in Mice. *PloS one* 3 (11), e3791. doi:10.1371/journal.pone.0003791

Taganov, K. D., Boldin, M. P., and Baltimore, D. (2007). MicroRNAs and Immunity: Tiny Players in a Big Field. *Immunity* 26 (2), 133–137. doi:10.1016/j.immuni.2007.02.005

Taganov, K. D., Boldin, M. P., Chang, K.-J., and Baltimore, D. (2006). NF-B-dependent Induction of microRNA miR-146, An Inhibitor Targeted to Signaling Proteins of Innate Immune Responses. *Proc. Natl. Acad. Sci.* 103 (33), 12481–12486. doi:10.1073/pnas.0605298103

Tan, Z., Randall, G., Fan, J., Camoretti-Mercado, B., Brockman-Schneider, R., Pan, L., et al. (2007). Allele-specific Targeting of microRNAs to HLA-G and Risk of Asthma. *Am. J. Hum. Genet.* 81 (4), 829–834. doi:10.1086/521200

Thakur, N., Singhal, P., Mehrotra, R., and Bharadwaj, M. (2019). Impacts of Single Nucleotide Polymorphisms in Three microRNAs (miR-146a, miR-196a2 and miR-499) on the Susceptibility to Cervical Cancer Among Indian Women. *Biosci. Rep.* 39 (4), BSR20180723. doi:10.1042/BSR20180723

Thumser, A. E., Moore, J. B., and Plant, N. J. (2014). Fatty Acid Binding Proteins. *Curr. Opin. Clin. Nutr. Metab. Care* 17 (2), 124–129. doi:10.1097/mco.0000000000000031

Tian, T., Shu, Y., Chen, J., Hu, Z., Xu, L., Jin, G., et al. (2009). A Functional Genetic Variant in microRNA-196a2 Is Associated with Increased Susceptibility of Lung Cancer in Chinese. *Cancer Epidemiol. Biomarkers Prev.* 18 (4), 1183–1187. doi:10.1158/1055-9965.epi-08-0814

Tsai, I.-T., Wu, C.-C., Hung, W.-C., Lee, T.-L., Hsuan, C.-F., Wei, C.-T., et al. (2020). FABP1 and FABP2 as Markers of Diabetic Nephropathy. *Int. J. Med. Sci.* 17 (15), 2338–2345. doi:10.7150/ijms.49078

Urbánek, P., and Klotz, L.-O. (2017). Posttranscriptional Regulation ofFOXOexpression: microRNAs and Beyond. *Br. J. Pharmacol.* 174 (12), 1514–1532. doi:10.1111/bph.13471

Usiello, A., Baik, J.-H., Rougé-Pont, F., Picetti, R., Dierich, A., LeMeur, M., et al. (2000). Distinct Functions of the Two Isoforms of Dopamine D2 Receptors. *Nature* 408 (6809), 199–203. doi:10.1038/35041572

Valinezhad Orang, A., Safaralizadeh, R., and Kazemzadeh-Bavili, M. (2014). Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-specific Upregulation. *Int. J. Genomics* 2014, 970607. doi:10.1155/2014/970607

Vaquero, A., Scher, M. B., Lee, D. H., Sutton, A., Chen, H.-L., Alt, F. W., et al. (2006). SirT2 Is a Histone Deacetylase with Preference for Histone H4 Lys 16 During Mitosis. *Genes Develop.* 20 (10), 1256–1261. doi:10.1101/gad.1412706

Vincent, A. M., Hayes, J. M., McLean, L. L., Vivekanandan-Giri, A., Pennathur, S., and Feldman, E. L. (2009). Dyslipidemia-induced Neuropathy in Mice: The Role of oxLDL/LOX-1. *Diabetes* 58 (10), 2376–2385. doi:10.2337/db09-0047

Vitale, A. V., Tan, H., and Jin, P. (2011). MicroRNAs, SNPs and Cancer. *J. Nucleic Acids Invest.* 2 (1), e6. doi:10.4081/jnai.2011.2236

von Nandelstadh, P., Gucciardo, E., Lohi, J., Li, R., Sugiyama, N., Carpen, O., et al. (2014). Actin-associated Protein Palladin Promotes Tumor Cell Invasion by Linking Extracellular Matrix Degradation to Cell Cytoskeleton. *MBoC* 25 (17), 2556–2570. doi:10.1091/mbc.e13-11-0667

von Otter, M., Bergström, P., Quattrone, A., De Marco, E. V., Annesi, G., Söderkvist, P., et al. (2014). Genetic Associations of Nrf2-Encoding NFE2L2 Variants with Parkinson's Disease - A Multicenter Study. *BMC Med. Genet.* 15, 131. doi:10.1186/s12881-014-0131-4

Walton, R. G., Zhu, B., Unal, R., Spencer, M., Sunkara, M., Morris, A. J., et al. (2015). Increasing Adipocyte Lipoprotein Lipase Improves Glucose Metabolism in High Fat Diet-Induced Obesity. *J. Biol. Chem.* 290 (18), 11547–11556. doi:10.1074/jbc.m114.628487

Wang, G., van der Walt, J. M., Mayhew, G., Li, Y.-J., Züchner, S., Scott, W. K., et al. (2008). Variation in the miRNA-433 Binding Site of FGF20 Confers Risk for Parkinson Disease by Overexpression of α-Synuclein. *Am. J. Hum. Genet.* 82 (2), 283–289. doi:10.1016/j.ajhg.2007.09.021

Wang, J.-X., Jiao, J.-Q., Li, Q., Long, B., Wang, K., Liu, J.-P., et al. (2011). miR-499 Regulates Mitochondrial Dynamics by Targeting Calcineurin and Dynamin-Related Protein-1. *Nat. Med.* 17 (1), 71–78. doi:10.1038/nm.2282

Wang, L., Zhang, N., Pan, H.-p., Wang, Z., and Cao, Z.-y. (2015). MiR-499-5p Contributes to Hepatic Insulin Resistance by Suppressing PTEN. *Cell Physiol Biochem* 36 (6), 2357–2365. doi:10.1159/000430198

Wang, P.-Y., Gao, Z.-H., Jiang, Z.-H., Li, X.-X., Jiang, B.-F., and Xie, S.-Y. (2013). The Associations of Single Nucleotide Polymorphisms in miR-146a, miR-196a and miR-499 with Breast Cancer Susceptibility. *PLoS One* 8 (9), e70656. doi:10.1371/journal.pone.0070656

Wang, X., Li, W., Ma, L., Gao, J., Liu, J., Ping, F., et al. (2015). Association Study of the miRNA-Binding Site Polymorphisms of CDKN2A/B Genes with Gestational Diabetes Mellitus Susceptibility. *Acta Diabetol.* 52 (5), 951–958. doi:10.1007/s00592-015-0768-2

Wang, X., Li, W., Ma, L., Ping, F., Liu, J., Wu, X., et al. (2017). Investigation of miRNA-Binding Site Variants and Risk of Gestational Diabetes Mellitus in Chinese Pregnant Women. *Acta Diabetol.* 54 (3), 309–316. doi:10.1007/s00592-017-0969-y

Wang, X., Qin, X., Yan, M., Shi, J., Xu, Q., Li, Z., et al. (2019). Loss of Exosomal miR-3188 in Cancer-Associated Fibroblasts Contributes to HNC Progression. *J. Exp. Clin. Cancer Res.* 38 (1), 151. doi:10.1186/s13046-019-1144-9

Wang, Y.-Q., Guo, R.-D., Guo, R.-M., Sheng, W., and Yin, L.-R. (2013). MicroRNA-182 Promotes Cell Growth, Invasion, and Chemoresistance by Targeting Programmed Cell Death 4 (PDCD4) in Human Ovarian Carcinomas. *J. Cel. Biochem.* 114 (7), 1464–1473. doi:10.1002/jcb.24488

Wei, N., Liu, S. S., Chan, K. K. L., and Ngan, H. Y. S. (2012). Tumour Suppressive Function and Modulation of Programmed Cell Death 4 (PDCD4) in Ovarian Cancer. *PLoS One* 7 (1), e30311. doi:10.1371/journal.pone.0030311

Weissbach, L., Settleman, J., Kalady, M. F., Snijders, A. J., Murthy, A. E., Yan, Y. X., et al. (1994). Identification of a Human rasGAP-Related Protein Containing Calmodulin-Binding Motifs. *J. Biol. Chem.* 269 (32), 20517–20521. doi:10.1016/s0021-9258(17)32023-9

Wilson, K. D., Hu, S., Venkatasubrahmanyam, S., Fu, J.-D., Sun, N., Abilez, O. J., et al. (2010). Dynamic MicroRNA Expression Programs during Cardiac Differentiation of Human Embryonic Stem Cells. *Circ. Cardiovasc. Genet.* 3 (5), 426–435. doi:10.1161/circgenetics.109.934281

Wojcicka, A., de la Chapelle, A., and Jazdzewski, K. (2014). MicroRNA-related Sequence Variations in Human Cancers. *Hum. Genet.* 133 (4), 463–469. doi:10.1007/s00439-013-1397-x

Wu, B., Liu, G., He, F., Liu, R., Wang, Z., Wang, Y., et al. (2019). miR-3188 (rs7247237-C>T) Single-Nucleotide Polymorphism Is Associated with the Incidence of Vascular Complications in Chinese Patients with Type 2 Diabetes. *J. Cardiovasc. Pharmacol.* 74 (1), 62–70. doi:10.1097/fjc.0000000000000681

Wunderlich, M., Ghosh, M., Weghorst, K., and Berberich, S. J. (2004). MdmX Represses E2F1 Transactivation. *Cell Cycle* 3 (4), 472–478. doi:10.4161/cc.3.4.746

Wynendaele, J., Böhnke, A., Leucci, E., Nielsen, S. J., Lambertz, I., Hammer, S., et al. (2010). An Illegitimate microRNA Target Site Within the 3′ UTR of MDM4 Affects Ovarian Cancer Progression and Chemosensitivity. *Cancer Res.* 70 (23), 9641–9649. doi:10.1158/0008-5472.can-10-0527

Xiang, Y., Fan, S., Cao, J., Huang, S., and Zhang, L.-p. (2012). Association of the microRNA-499 Variants with Susceptibility to Hepatocellular Carcinoma in a Chinese Population. *Mol. Biol. Rep.* 39 (6), 7019–7023. doi:10.1007/s11033-012-1532-0

Xu, C., Zhu, J., Fu, W., Liang, Z., Song, S., Zhao, Y., et al. (2016). MDM4 Rs4245739 A > C Polymorphism Correlates with Reduced Overall Cancer Risk in a Meta-Analysis of 69477 Subjects. *Oncotarget* 7 (44), 71718–71726. doi:10.18632/oncotarget.12326

Xu, Y., Li, L., Xiang, X., Wang, H., Cai, W., Xie, J., et al. (2013). Three Common Functional Polymorphisms in microRNA Encoding Genes in the Susceptibility to Hepatocellular Carcinoma: A Systematic Review and Meta-Analysis. *Gene* 527 (2), 584–593. doi:10.1016/j.gene.2013.05.085

Yan, W., Gao, X., and Zhang, S. (2017). Association of miR-196a2 Rs11614913 and miR-499 Rs3746444 Polymorphisms with Cancer Risk: A Meta-Analysis. *Oncotarget* 8 (69), 114344–114359. doi:10.18632/oncotarget.22547

Yang, B., Chen, J., Li, Y., Zhang, J., Li, D., Huang, Z., et al. (2012). Association of Polymorphisms in Pre-miRNA with Inflammatory Biomarkers in Rheumatoid Arthritis in the Chinese Han Population. *Hum. Immunol.* 73 (1), 101–106. doi:10.1016/j.humimm.2011.10.005

Yang, F., Chen, F., Xu, J., and Guan, X. (2016). Identification and Frequency of the Rs12516 and Rs8176318 BRCA1 Gene Polymorphisms Among Different Populations. *Oncol. Lett.* 11 (4), 2481–2486. doi:10.3892/ol.2016.4252

Yang, H.-S., Matthews, C. P., Clair, T., Wang, Q., Baker, A. R., Li, C.-C. H., et al. (2006). Tumorigenesis Suppressor Pdcd4 Down-Regulates Mitogen-Activated Protein Kinase Kinase Kinase Kinase 1 Expression to Suppress Colon Carcinoma Cell Invasion. *Mol. Cel Biol* 26 (4), 1297–1306. doi:10.1128/mcb.26.4.1297-1306.2006

Yekta, S., Shih, I.-h., and Bartel, D. P. (2004). MicroRNA-directed Cleavage of HOXB8 mRNA. *Science* 304 (5670), 594–596. doi:10.1126/science.1097434

Zhang, H., Zhang, Y., Yan, W., Wang, W., Zhao, X., Ma, X., et al. (2017). Association Between Three Functional microRNA Polymorphisms (miR-499 Rs3746444, miR-196a Rs11614913 and miR-146a Rs2910164) and Breast Cancer Risk: A Meta-Analysis. *Oncotarget* 8 (1), 393–407. doi:10.18632/oncotarget.13426

Zhang, S., Li, J., Jiang, Y., Xu, Y., and Qin, C. (2009). Programmed Cell Death 4 (PDCD4) Suppresses Metastastic Potential of Human Hepatocellular Carcinoma Cells. *J. Exp. Clin. Cancer Res.* 28 (1), 71–11. doi:10.1186/1756-9966-28-71

Zhang, Y., Bai, R., Liu, C., Ma, C., Chen, X., Yang, J., et al. (2019). MicroRNA Single-Nucleotide Polymorphisms and Diabetes Mellitus: A Comprehensive Review. *Clin. Genet.* 95 (4), 451–461. doi:10.1111/cge.13491

Zhao, D. (2020). Single Nucleotide Alterations in MicroRNAs and Human Cancer-A Not Fully Explored Field. *Non-coding RNA Res.* 5 (1), 27–31. doi:10.1016/j.ncrna.2020.02.003

Zhao, X., Ye, Q., Xu, K., Cheng, J., Gao, Y., Li, Q., et al. (2013). Single-nucleotide Polymorphisms Inside microRNA Target Sites Influence the Susceptibility to Type 2 Diabetes. *J. Hum. Genet.* 58 (3), 135–141. doi:10.1038/jhg.2012.146

Zheng, H., Song, F., Zhang, L., Yang, D., Ji, P., Wang, Y., et al. (2011). Genetic Variants at the miR-124 Binding Site on the Cytoskeleton-Organizing IQGAP1 Gene Confer Differential Predisposition to Breast Cancer. *Int. J. Oncol.* 38 (4), 1153–1161. doi:10.3892/ijo.2011.940

Zhou, B., Rao, L., Peng, Y., Wang, Y., Chen, Y., Song, Y., et al. (2010). Common Genetic Polymorphisms in Pre-microRNAs Were Associated with Increased Risk of Dilated Cardiomyopathy. *Clin. Chim. Acta* 411 (17-18), 1287–1290. doi:10.1016/j.cca.2010.05.010

Zhou, J., Lv, R., Song, X., Li, D., Hu, X., Ying, B., et al. (2012). Association Between Two Genetic Variants in miRNA and Primary Liver Cancer Risk in the Chinese Population. *DNA Cel. Biol.* 31 (4), 524–530. doi:10.1089/dna.2011.1340

Zhou, L., Zhang, X., Li, Z., Zhou, C., Li, M., Tang, X., et al. (2013). Association of a Genetic Variation in a miR-191 Binding Site in MDM4 with Risk of Esophageal Squamous Cell Carcinoma. *PloS one* 8 (5), e64331. doi:10.1371/journal.pone.0064331

Zhuang, G. Q., and Wang, Y. X. (2017). A Tiny RNA Molecule with a Big Impact on Type 2 Diabetes Mellitus Susceptibility. *Biomed. Environ. Sci.* 30 (11), 855–861. doi:10.3967/bes2017.116

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 Genetics

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fgene.2021.755629

Prevalence and Spectrum of Predisposition Genes With Germline Mutations Among Chinese Patients With Bowel Cancer

*Zhengyong Xie†, Yongli Ke†, Junyong Chen, Zehang Li, Changzheng Wang, Yuhong Chen, Hongliang Ding and Liyang Cheng**

General Surgery Department, General Hospital of Southern Theatre Command, People's Liberation Army of China (PLA), Guangzhou, China

OPEN ACCESS

Edited by:
Rajkumar S. Kalra, Okinawa Institute of Science and Technology Graduate University, Japan

Reviewed by:
Aijaz Ahmad Naik, National Institute of Mental Health (NIH), United States
Manish Kumar Dwivedi, Indian Institute of Science Education and Research, Bhopal, India
Devivasha Bordoloi, Wistar Institute, United States

***Correspondence:**
Liyang Cheng
chliyang2021_1@163.com

†These authors share first authorship

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *09 August 2021*
Accepted: *04 October 2021*
Published: *27 January 2022*

Citation:
Xie Z, Ke Y, Chen J, Li Z, Wang C, Chen Y, Ding H and Cheng L (2022) Prevalence and Spectrum of Predisposition Genes With Germline Mutations Among Chinese Patients With Bowel Cancer. Front. Genet. 12:755629. doi: 10.3389/fgene.2021.755629

Background: Bowel cancer is the third-most common cancer and the second leading cause of cancer-related death worldwide. Bowel cancer has a substantial hereditary component; however, additional hereditary risk factors involved in bowel cancer pathogenesis have not been systematically defined.

Materials and Methods: A total of 573 patients with bowel cancer were enrolled in the present study, of whom 93.72% had colorectal cancer (CRC). Germline mutations were integrated with somatic mutation information via utilizing target next-generation sequencing.

Results: Pathogenic/Likely Pathogenic (P/LP) germline alterations were identified in 47 (8.2%) patients with bowel cancer and the ratio of the number of these patients with family history was significantly higher in the P/LP group than that noted in the non-pathogenic (Non-P) group. Certain rare germline alterations were noted, such as those noted in the following genes: *FANCD2, CDH1*, and *FLCN*. A total of 32 patients (68.1%) had germline alterations in the DNA-damage repair (DDR) genes and homologous recombination (HR) accounted for the highest proportion of this subgroup. By comparing 573 patients with bowel cancer with reference controls (China_MAPs database), significant associations ($p < 0.01$) were observed between the incidence of bowel cancer and the presence of mutations in *APC, ATM, MLH1, FANCD2, MSH3, MSH6, PMS1*, and *RAD51D*. Somatic gene differential analysis revealed a marked difference in 18 genes and a significant difference was also noted in tumor mutation burden (TMB) between germline mutation carriers and non-germline mutation subjects ($p < 0.001$). In addition, TMB in DDR mutation groups indicated a dramatic difference compared with the non-DDR mutation group ($p < 0.01$). However, no statistically significant differences in TMB were noted among detailed DDR pathways for patients with bowel cancer, irrespective of the presence of germline mutations. Moreover, a significantly higher level ($p < 0.0001$) of mutation count was observed in the DDR group from The Cancer Genome Atlas (TCGA) database and the DDR and non-DDR alteration groups displayed various immune profiles.

Conclusion: Chinese patients with bowel cancer exhibited a distinct spectrum of germline variants, with distinct molecular characteristics such as TMB and DDR. Furthermore, the information on somatic mutations obtained from TCGA database indicated that a deeper

understanding of the interactions among DDR and immune cells would be useful to further investigate the role of DDR in bowel cancer.

Keywords: bowel cancer, germline, somatic, P/LP (pathogenic/likely-pathogenic), TMB, DDR

INTRODUCTION

Bowel cancer ranks third with regard to cancer morbidity and mortality worldwide (Thanikachalam and Khan, 2019). According to the Chinese Cancer Registration Report of 2018, 387,600 bowel cancer new cases and 187,100 bowel cancer-related deaths occurred in China during 2015, ranking it the fourth (9.87%) and fifth (8.01%) highest incidence and mortality rates, respectively, among all cancers (Arnold et al., 2020; Yang et al., 2020). In addition, the rates of bowel cancer steadily increased from 2000 to 2018 (Bhui et al., 2009; Bray et al., 2018). Although lower rates compared with the world average (incidence rate of 17.81/100,000 persons and mortality rate of 8.12/100,000 persons) (Bray et al., 2018), the number of new bowel cancer cases and bowel cancer-related deaths in China is the highest in the world due to its relatively large population. Genetic factors resulting in the early development of cancers account for a substantial number of bowel cancer (Medina Pabón and Babiker, 2021). Therefore, it is necessary to explore the prevalence of hereditary bowel cancer and the contribution of the pathogenic germline variants in the development of this disease in the Chinese population.

In general, hereditary cancer syndromes have been implicated in 3–5% of overall cases with bowel cancer (Medina Pabón and Babiker, 2021). Individuals who harbor germline mutations in specific genes are at high risk for developing bowel cancer. Clinically, individuals with hereditary bowel cancer syndromes may be alert to this situation and more likely to undertake frequent early screening (Sokic-Milutinovic and eng, 2019). It is known that germline mismatch repair (MMR) gene mutations, together with *APC* gene mutations, contribute significantly to inherited bowel cancer (Sa et al., 2020). Lynch Syndrome, the most common hereditary cancer syndrome associated with predisposition to bowel cancer, is associated with germline mutations in DNA mismatch repair (MMR) genes such as *MLH1, MSH2, MSH6, PMS2,* and *EPCAM* (Engel et al., 2020). Familial adenomatous polyposis (FAP) is associated with germline mutations in the *APC* tumor suppressor gene and has been implicated in 1% of cases with bowel cancer. Germline mutations in additional high and moderate penetrance cancer genes such as *BRCA1, CDH1,* and *MUTYH*, have also been associated with increased risk for the developing colorectal neoplasia (Ma et al., 2018). Recent studies have demonstrated that germline variants in various cancer predisposition genes have been identified in 1 out of 10 adults and children diagnosed with advanced cancer types, as well as those with colorectal (Mueller et al., 2021), pancreatic (Gentiluomo et al., 2020), and metastatic prostate (Giri et al.eng, 2019) cancers. The impact of an individual germline variant for clinical decision-making depends on the specific characteristics of the variant (Bertelsen et al., 2019), which classify whether the variant is pathogenic/likely pathogenic (P/LP), or whether it is known and/or it is likely to affect the function of its gene.

Accurate interpretation of genetic test results is of vital importance, notably for patients who are identified with one or more P/LP germline variant associated with a hereditary cancer syndrome. Due to the development of next-generation sequencing (NGS) technology, it has been found that various ratio of patients with bowel cancer harbor germline mutations (Bien et al., 2019). However, despite germline variants in genes related to cancer susceptibility being more common than initially expected, identification of germline mutations of Chinese patients with bowel cancer and the correlation between germline mutations and somatic mutations has not been studied in detail. The present study sought to determine the characteristics of P/LP germline variants in Chinese patients with bowel cancer. The results revealed that a wider panel of predisposition genes are recommended for Chinese patients with bowel cancer, which will be helpful to aid the establishment of prevention and surveillance strategies that can be used to reduce the incidence of this disease.

MATERIALS AND METHODS

Samples Source and Ethic Data

Patients with bowel cancer gave written informed consent prior to their participation in General Hospital of Southern Theatre Command, PLA. Formalin-fixed, paraffin-embedded (FFPE) tumor tissues and matched blood samples in EDTA tubes (for germline tests) from 573 diagnosed bowel cancer patients (Information on clinicopathological status of patients is provided in **Supplementary Table S2**) were collected. All tumor FFPE sections were evaluated by pathologist to contain at least 20% tumor cells. Family history here is defined as the confirmed bowel cancer patient who has at least one family member (first and second-degree relatives) who had a history of tumor diagnosis. The immediate family member includes father, mother, brother(s), sister(s), son(s), daughter(s); second degree relatives include grandparent(s), uncle(s), aunt(s).

Deoxy Ribonucleic Acid Isolation and Targeted Next-Generation Sequencing

The FFPE tissues and peripheral white blood cells were collected to extract DNA using QIAamp DNA FFPE Tissue Kit and DNeasy Blood and Tissue Kit (Qiagen, Inc.),

respectively. And the purified gDNA was quantified using the Qubit 3.0 Fluorometer (Life Technologies, Inc.).

For the matched germline and tumor samples, 100 ng of DNA was shared with a Covaris E210 system (Covaris, Inc.) to obtain an average of 200 bp fragments. Accel-NGS 2S DNA Library Kit (Swift Biosciences, Inc.) and xGen Lockdown Probes kit (IDT, Inc.) were used for NGS library preparation of the tumor gDNA matched germline gDNA. The custom xGen Lockdown probe was synthesized by IDT, Inc. to target the exons and selected intronic regions of 499 genes (Gene list is provided in **Supplementary Table S1**).

Interpretation of Pathogenicity of Germline Mutations and Calculation of Somatic Tumor Mutation Burden

Pathogenicity of germline mutations was defined and predicted based on the five-grade classification system according to the American College of Medical Genetics and Genomics (ACMG) Guidelines for the Interpretation of Sequence (Li et al., 2017). It was modified here that pathogenic/likely-pathogenic germline variants were depicted as P/LP and the variant of undetermined significance (VUS), benign, likely benign, and somatic mutations were defined as the non-pathogenic group (Non-P) in our results. Therefore, all mutations were categorized into P/LP or Non-P groups in this study.

Tumor mutation burden of each sample was calculated according to a published and the method of Chalmers et al (2017).

Data Processing

Germline mutation data and incidence rates were obtained from the ChinaMAP database (http://www.GenomAD.org). The Cancer Genome Atlas (TCGA) database (https://tcga-data.nci.nih.gov/tcga/) provides several expression profiles and mutation data of CRC, as well as corresponding clinical data. Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway terms were considered statistically significant when FDR < 0.01. CIBERSORT was used for evaluating diverse immune cell types in the cancer microenvironment. The violin software package was used to visualize differentially infiltrated immune cells between the two groups through the Wilcoxon test.

Statistical Analysis

Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) statistical package and Graphpad (Prism 8). Student's t-test was performed when two groups were compared, and ANOVA and post hoc tests were performed when three or more groups were compared. Gene prevalence between different groups was analyzed by Chi-Square test or Fisher exact test under/with a dominant model. A two-sided p value of less than 0.05 was considered to be statistically significant.

TABLE 1 | Description of cohort.

Characteristic	Subgroups	Total evaluable cohort, no. (%)	
No. of participants		573	
Age, year	<50	124	21.64%
	≥50	449	78.36%
Sex	female	227	39.62%
	male	346	60.38%
Family history	Yes*	121	21.11%
	First-degree	108	18.85%
	Second-degree	13	2.27%
	No	275	48%
	NA	177	30.89%
Stage	I	7	1.22%
	II	202	35.25%
	III	63	10.99%
	IV	301	52.53%
Tumor Location	Colon cancer	380	66.32%
	Rectal cancer	157	27.40%
	Duodenal Cancer	27	4.71%
	Small bowel cancer	6	1.05%
	Cecal cancer	3	0.52%

Yes: the confirmed colorectal cancer patient has at least one family member (first- and second-degree relatives) who had a history of tumor diagnosis.*

RESULTS

Demographic Characteristics and Landscape of Mutational Profiles in Chinese Patients With Bowel Cancer

Briefly, paired tumor/germline analysis was conducted using a customized next generation sequencing (NGS) panel of 499 selected genes (**Supplementary Table S1**). Somatic variants were determined by comparing the data between tumor and blood samples and all participants were included as a result of successful germline sequencing, which resulted in an evaluable population of 573 patients with bowel cancer. The demographic, clinical, and pathological characteristics of this patient cohort are shown in **Table 1**. In this cohort (n = 573), 21.64% (n = 124) participants who were diagnosed with bowel cancers were < 50 years old and 78.36% (n = 449) > 50 years old. Approximately 39.62% (n = 227) patients were female, 66.32% (n = 380) exhibited colon cancer, 27.4% (n = 157) owned rectal cancer, and 52.53% (n = 301) patients were reported to have the stage IV tumors. A total of 108 (18.85%) participants had one or more first-degree relatives with a history of tumor diagnosis, and 13 (2.27%) had one or more second-degree relatives with a history of tumor diagnosis.

Characteristics of Pathogenic Germline Mutations in the Chinese Cohort and Their Impact on Bowel Cancer Risk

Overall, 47 patients were found to carry P/LP germline mutations (**Table 2**). The age and sex of the patients were not associated with the presence or absence of a P/PL germline mutation (p = 0.19

TABLE 2 | The summary of clinicopathological and history information for bowel cancer patients with or without distinct germline mutation pathogenicity.

Characteristic	Subgroups	P/LP (N = 47)		Non-pathogenic (N = 526)		*p* Value
Age, year	<50	14	29.79%	110	20.91%	0.194
	≥50	33	70.21%	416	79.09%	—
Sex	female	23	48.93%	204	38.78%	0.213
	male	24	51.07%	322	61.22%	—
Family history	Yes*	14	29.79%	107	20.34%	0.037
	First-degree	12	25.53%	96	18.25%	—
	Second-degree	2	4.26%	11	2.09%	—
	No	15	31.91%	260	49.43%	—
	NA	18	38.30%	159	30.23%	—

Yes: the confirmed colorectal cancer patient has at least one family member (first and second degree relatives) who had a history of tumor diagnosis.*

and $p = 0.21$, respectively, **Table 2**). Interestingly, the ratio of patients with bowel cancer with at least one family member (first- and second-degree relatives) with family history of cancer(s), such as colon, breast, endometrium, ovary, and/or pancreas was significantly higher in the P/LP group than that of the Non-P group ($p = 0.037$, **Table 2**).

A total of 30 genes with P/PL germline variants were detected among 47 patients (**Figure 1A**). Besides, this study identified 5 out of 47 patients who carried *MUTYH* P/LP mutations (5/47), followed by *APC* (4/47) *MLH1* (3/47), *TP53*(3/47), and *ATM* (3/47) (**Supplementary Table S2**, **Figure 2A**). Moreover, it was found that 8 patients carried LS related mutations (3 *MLH1*, 2 *MSH2*, 2 *MSH6*, and 1 *EPCAM*) and 39 carried non-LS mutations (**Figure 1A**). These findings were consistent with those reported in the previous studies (Latham et al., 2019). Notably, certain novel P/LP mutations were present, including those in *FANCD2, RAD51D, BLM, CDH1, FLCN, MEN1, SDHB,* and *SLX4* were newly discovered in our cohort and have been rarely reported in the previous publications. The detailed distribution data and information on the germline mutations are presented in **Supplementary Table S2**. The functions of these genes with newly discovered P/LP mutations were mainly involved in DNA damage repair pathways (DDR-related genes were shown in **Supplementary Table S3**) and exerted a very broad impact. They included the following components: i) Those contributing to homologous recombination (HR), such as *BRCA1, RAD51D, MRE11A*, and *RAD51B*; ii) Those involved in Fanconi anemia (FA), such as *FANCG, FANCA, SLX4, BLM, FANCD2, FANCL,* and *BRIP1*; iii) Those involved in base excision repair (BER), such as *MUTYH*; iv) those involved in nucleotide excision repair (NER), such as *ERCC2*; v) those involved in mismatch repair (MMR), such as *MLH1, MSH2, MSH3, MSH6, PMS1,* and *EPCAM*; vi) those contributing to DNA sensor (DS), such as *ATM* and *CHEK2* (**Figure 1B**).

Mutations in the DNA damage repair genes increase the risk of subsequent mutations and therefore confer high cancer susceptibility. Previous studies verified the association between alterations in 34 DDR genes that exhibited higher TMB levels in urothelial cancer and demonstrated that DDR inactivation was associated with higher levels of TMB (Wang et al., 2018). Therefore, the TMB of CRC patients harboring germline mutations was compared among different DDR subgroups. Although there were no statistically significant differences with regard to the incidence of TMB in these pathways, the average value of MMR was the highest with the exception of the group of others, while HR exhibited the lowest (**Figure 1C**). This phenomenon is in line with the previous studies showing that over 90% of MMR-deficient tumors exhibit high TMB levels (Stadler et al., 2016). In addition, the data indicated that the majority of germline mutations were located in or affecting protein functional domains and that they may have a significant impact on protein function (**Supplementary Figure S1**).

To investigate the risk of bowel cancer in individuals carrying P/LP germline mutations, the mutation prevalence of all germline mutations was searched in the total population and in different populations derived from the China_MAPs database (**Table 3**). Eight genes were significantly associated with CRC compared with control subjects derived from the China_MAPs database. These included the following: i) *APC*, with mutations in 0.70% of cases and in 0.01% of control subjects (OR, 74.32); ii) *ATM*, with mutations in 0.52% of cases and in 0.03% of control subjects (OR, 18.56); iii) *MLH1*, with mutations in 0.52% of cases and in 0.01% of control subjects (OR, 111.50); iv) *FANCD2*, with mutations in 0.35% of cases and in 0.03% of control subjects (OR, 12.35); v) *MSH3*, with mutations in 0.35% of cases and in 0.04% of control subjects (OR, 9.27); vi) *MSH6*, with mutations in 0.35% of cases and in 0.01% of control subjects (OR, 74.07); vii) *PMS1*, with mutations in 0.35% of cases and in 0.30% of control subjects (OR, 10.59); viii) *RAD51D*, with mutations in 0.35% of cases and in 0.01% of control subjects (OR, 74.07) (**Table 3**). These findings suggested that these P/LP germline mutations were risk factors for the development of bowel cancer.

Furthermore, the bowel cancer germline mutation frequency found in the present study was compared with that reported in other studies including the investigations performed in Japan (Fujita et al., 2020) and in America (Yurgelun et al., 2017). For the incidence of *APC* and *MUTYH* in the Japanese cohort were significantly lower than those noted in the present ($p = 0.02$, $p < 0.001$ respectively). The P/PL prevalence did not differ significantly of genes including *MLH1, MSH2, MSH6, BRCA1, TP53, CHEK2, ATM,* and *BRIP1* among the three cohorts investigated. (**Table 4**).

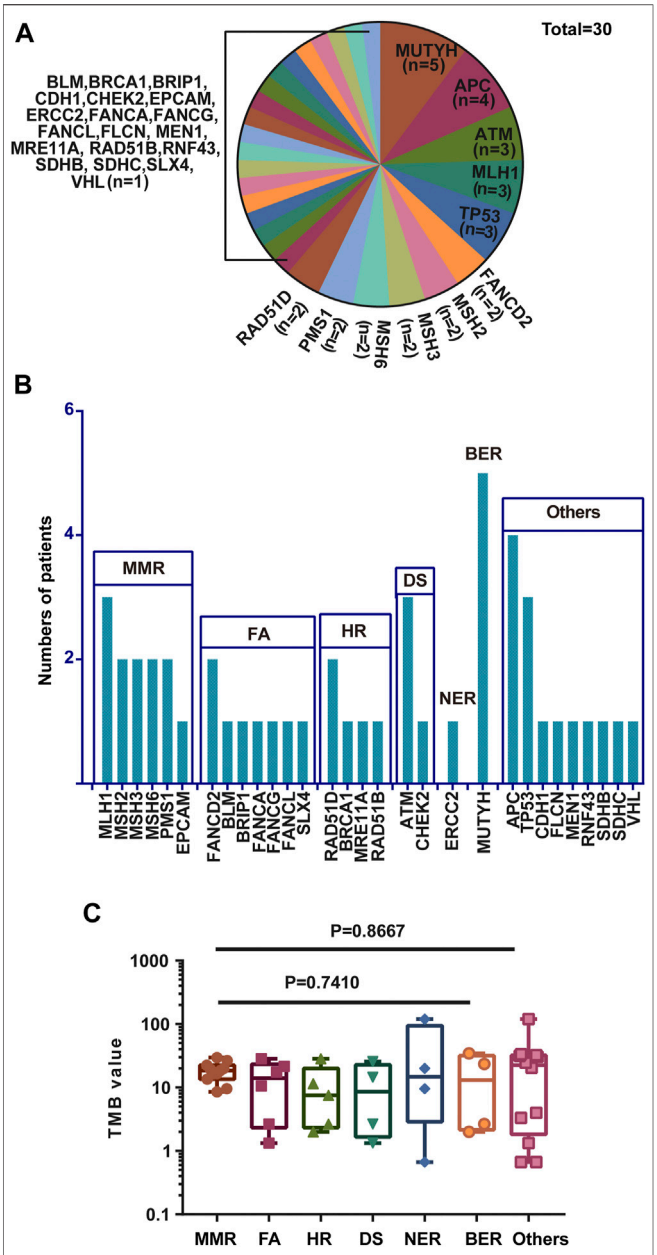


FIGURE 1 | Gene names, functions, and number of variations of all P/LP germline mutations. **(A)** Gene names and the ratio of mutations of P/LP germline variations. **(B)** Specific gene mutation number with each DDR pathway in the P/LP group. **(C)** Comparison of the TMB with somatic mutations among different pathways in the P/LP group. pathogenic/likely-pathogenic: P/LP; DDR: DNA damage response; TMB, tumor mutation burden; HR, homologous recombination; FA, fanconi anemia; MMR, mismatch repair; NER, nucleotide excision repair; BER, base excision repair; DS, DNA sensor.

Molecular Analysis of Somatic Mutations of Patients With Bowel Cancer Carrying Germline P/LP Mutations

The relationship between germline mutation carriers and patients with somatic mutations has been studied in other cancer types, such as lung cancer (Liu et al., 2020). However, the connection of germline variations and somatic mutations in bowel cancer has not been explored in detail. The somatic mutation spectrum was classified by pathogenicity/likely pathogenicity of germline mutations for all patients with bowel cancer (P/LP and Non-P groups) (**Figures 2A,B**). *TP53, APC, KRAS, NOTCH1, ARID1A, FBXW7, PIK3CA, KMT2C, NOTCH3,* and *TCF7L2*, were found to be the top 10 mutated genes in the P/LP group. With regard to the Non-P group, the top 10 mutated genes were *TP53, APC, PIK3CA, SMAD4, NOTCH3, FAT1, ARID1A, KMT2D, RECQL4,* and *FBXW7*. According to the different mutation classification categories, the missense mutation was the one that obtained the highest proportion in the bowel cancer samples in the presence or absence of germline variants.

To determine the presence of germline variants in the patients examined, the comparison of somatic alterations was conducted between P/LP and Non-P groups, and the results revealed dramatic differences ($p < 0.001$) in several gene mutations between these two groups (**Figure 3A**). A total of 17 gene mutations such as *TCF7L2, KMT2D, PRKDC, NOTCH1, KMT2C, ERBB3,* and *TSC2*, and others were more common in patients with bowel cancer with P/LP germline mutations; strikingly, *SMAD4* exhibited the opposite prevalence and its mutation frequency in the Non-P group was significantly higher than that in the P/LP group (**Figure 3A**). Subsequently, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were further conducted to explore the biological roles of the identified differential genes. GO as well as KEGG enrichment analysis revealed that these genes were mainly enriched in the biological process (BP) terms, such as PI3K signaling, DNA recombination, HR, and MMR (**Supplementary Figure S2**). Furthermore, the identification of specific driver genes in the present study has been previously reported (Eklöf et al., 2013; Dow et al., 2015). The mutation rate (frequency) of *APC* was the highest among all genes both in the P/LP and non-P groups followed by the mutation rate of *TP53*, *KRAS*, *PIK3CA*, *SMAD4* and *BRAF* (**Figure 3B**).

Association Between Tumor Mutation Burden and the Incidence of DNA-Damage Repair Mutations in the Chinese Cohort

TMB is considered a vital biomarker in a variety of cancer types (Samstein et al., 2019; Wu et al., 2019), which may reflect the degree of genomic instability at the nucleotide level. Therefore, the TMB was initially compared in the P/LP and Non-P groups and the data indicated a significantly higher level of TMB in the P/LP group (14.56 vs. 6.39 mutations/Mb, $p = 0.0056$, **Figure 4**). Furthermore, the DDR system plays an important role in maintaining genome stability, based on this notion, a focused analysis was performed, specifically on the DDR-altered genes in all bowel cancer cases. In the P/LP group, DDR alterations were present in 68.1% ($n = 32$) of cases and were involved in the DDR pathways including HR (31.36%), DS (20.82%), FA (19.28%), MMR (12.85%), BER (9%) and NER (6.68%), respectively (**Figure 5A**). In the Non-P group, DDR gene mutations occurred in 30% ($n = 157$) of cases, including HR (26.97%),

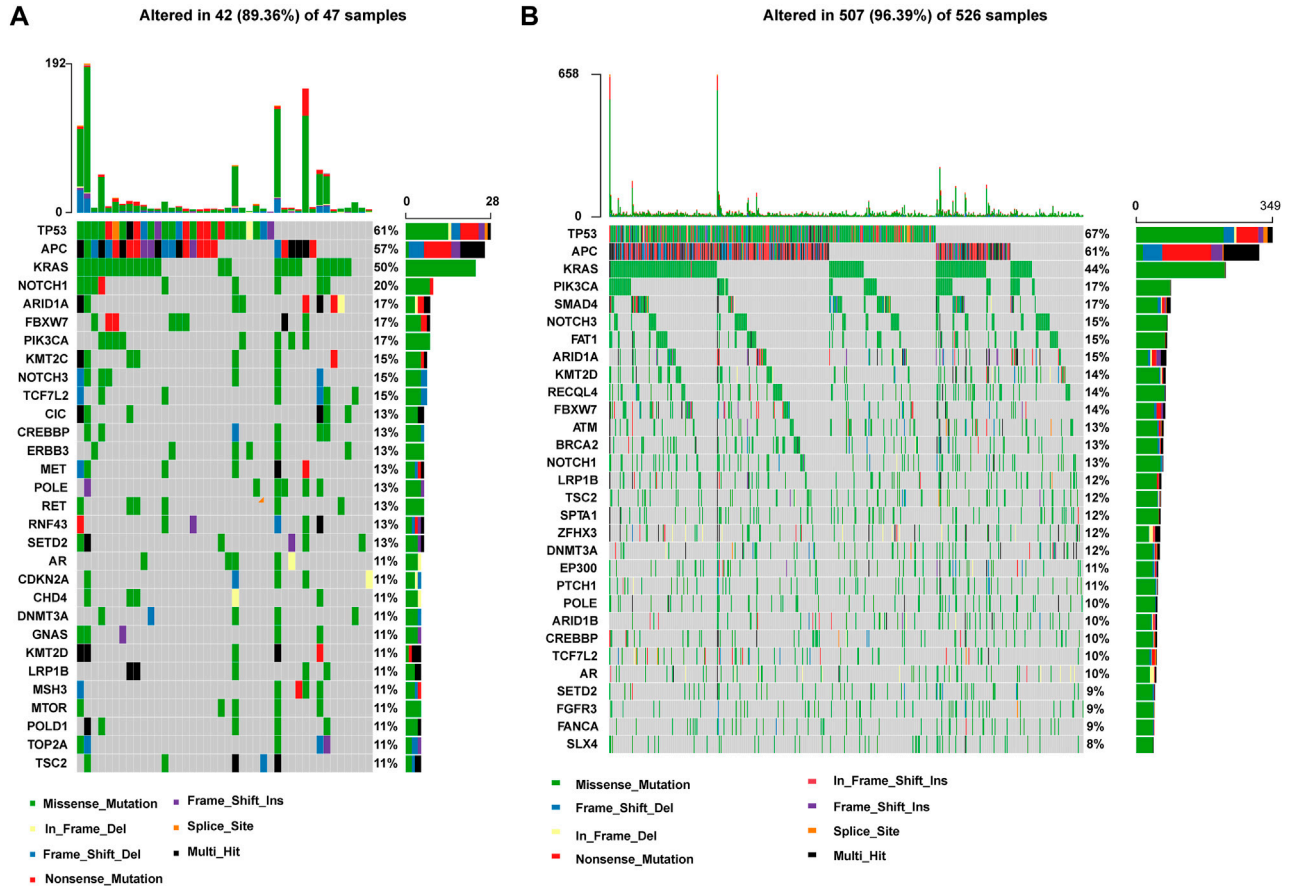


FIGURE 2 | Heatmap of full SNV and INDEL somatic mutation for patients with P/LP **(A)** or Non-P **(B)** germline mutations. **(A)** Somatic mutation spectrum for 47 patients with P/LP germline mutations. **(B)** Somatic mutation spectrum for 526 patients with Non-P germline mutations. Details of mutations are labeled beneath each panel, and somatically mutated genes are listed in the left of the figures. The right bars represent the overall number of mutations for each gene. *Y*-axis above each panel represents the number of somatic mutations detected for each patient. Colors represent mutation types as indicated by the figure legend.

MMR (20.22%), FA (20.22%), BER (13.48%), DS (11.24%) and NER (7.67%) (**Figure 5B**). In the P/LP group, *FANCA, ATM,* and *MUTYH* were the most commonly altered DDR genes, followed by *BRCA2, MSH3,* and *POLE*; whereas in the Non-P group, *BRCA2* exhibited the highest DDR gene alterations, followed by *ATM, POLE,* and *MSH6* (**Figure 5C**). Furthermore, the highest average TMB value was observed in DDR-altered Non-P cases with bowel cancer compared with that of DDR-altered P/LP cases pf bowel cancer or Non-DDR altered cases ($p < 0.001$, **Figure 5D**). The data demonstrated that tumors with DDR mutations in the Non-P group exhibited increased genomic instability than that of the P/LP group. The lowest average TMB appeared in the non-DDR group, which is in line with the evidence reported in the previous studies.

Comparison of Somatic Mutations Between the Chinese Cohort and the Independent The Cancer Genome Atlas Cohort

A total of 223 bowel cancer samples with somatic mutation profiles were downloaded from the TCGA database. The relevant clinical information was listed in **Supplementary Table S4**. DDR gene mutation analysis was also conducted in the TCGA cohort, and the FA pathway accounted for the highest proportion (25.22%), while NER accounted for the lowest proportion (7.42%) (**Figure 6A**). Moreover, a similar finding was obtained regarding specific gene mutations in these pathways between TCGA cohort and the Chinese cohort, such as *ATM* (the highest mutation frequency) and *EPCAM* (relative lower mutation frequency) (**Figure 6B**).

Subsequently, the TMB in the Chinese cohort was compared with the mutation count in the TCGA cohort (TMB was not accessible and therefore the mutation count was adopted here). The selection of mutation genes with P/LP was performed *via* the cBioportal website. In concordance with the previous results, a significantly higher level of TMB was observed in bowel cancer harboring DDR somatic mutations compared with cases with non-DDR bowel cancer in TCGA cohort ($p < 0.0001$, **Figure 7**). A similar finding was obtained from the Chinese cohort (**Figure 5D**). Especially, although there was no statistically significant difference among the DDR pathways examined in the TCGA and the Chinese cohort, the lowest TMB or mutation

TABLE 3 | Comparison of mutation carriers by panel gene between colorectal cancer cases and China_MAPs control cases.

Genes	Cases			China_MAPs Cases			Cancer Risk	
	Cases With Mutations, No	Individuals Tested, No	Carrier Frequency, %	Controls With Mutations, No	Individuals Tested, No	Carrier Frequency, %	Odds Ratio(95%CI)	*p* Value
Genes Significantly Associated with Colorectal Cancer								
APC	4	573	0.70	2	21,176	0.01	74.32 (10.63–833.35)	˂0.001
ATM	3	573	0.52	6	21,176	0.03	18.56 (3.00–87.19)	0.001
MLH1	3	573	0.52	1	21,176	0.01	111.50 (8.92–5,589.64)	˂0.001
	2	573	0.35	6	21,176	0.03	12.35 (1.22–69.42)	0.02
FANCD2								
MSH3	2	573	0.35	8	21,176	0.04	9.27 (0.96–46.56)	0.03
MSH6	2	573	0.354	1	21,176	0.01	74.07 (3.85–4,220.60)	0.002
PMS1	2	573	0.35	7	21,176	0.03	10.59 (1.07–55.73)	0.02
	2	573	0.35	1	21,176	0.01	74.07 (3.85–4,220.60)	0.002
RAD51D								
Genes Not Significantly Associated with Colorectal Cancer								
BLM	**1**	573	0.17	11	21,176	0.05	4.62(0.10–34.60)	3.36
BRCA1	**1**	573	0.17	8	21,176	0.04	5.29 (0.12–41.26)	4.62
BRIP1	1	573	0.17	7	21,176	0.03	12.33 (0.23–153.20)	5.29
CHEK2	1	573	0.17	3	21,176	0.01	18.50 (0.31–354.76)	12.33
EPCAM	1	573	0.17	2	21,176	0.01	6.17 (0.13–50.98)	18.50
ERCC2	1	573	0.17	6	21,176	0.03	5.29 (0.12–41.26)	6.17
FANCA	1	573	0.17	7	21,176	0.03	9.25 (0.19–93.72)	5.29
FANCL	1	573	0.17	4	21,176	0.02	36.98 (0.47–2,826.49)	9.25
SLX4	1	573	0.17	9	21,176	0.01	4.11 (0.09–29.76)	36.98
VHL	1	573	0.17	3	21,176	0.04	12.33 (0.23–153.20)	4.11
MEN1	1	573	0.17	1	21,176	0.01	4.62(0.10–34.60)	12.33

TABLE 4 | Comparison of germline mutation carriers with specific genes among different countries.

Genes	China		Japan			America		
	Individuals Tested, No	Carrier Frequency, %	Individuals Tested, No	Carrier Frequency, %	China vs. Japan *p* Value	Individuals Tested, No	Carrier Frequency, %	China vs. America *p* Value
APC	573	0.70	12,503	0.16	0.02	1,058	0.47	0.73
MUTYH	573	0.87	12,503	0.10	˂0.001	1,058	1.80	0.20
MLH1	573	0.52	12,503	0.28	0.23	1,058	1.18	0.20
MSH2	573	0.35	12,503	0.29	0.68	1,058	0.66	0.51
MSH6	573	0.35	12,503	0.31	0.70	1,058	0.57	0.72
BRCA1	573	0.17	12,503	0.17	1.00	1,058	0.28	1.00
TP53	573	0.52	12,503	0.15	0.07	1,058	0.09	0.13
CHEK2	573	0.17	12,503	0.12	0.51	1,058	0.19	1.00
ATM	573	0.52	12,503	0.37	0.47	1,058	0.95	0.56
BRIP1	573	0.17	12,503	0.14	0.57	1,058	0.28	1.00

count was observed in NER pathway. In addition, the HR pathway displayed an increased mutation count in the TCGA cohort, while the Chinese cohort harbored the highest TMB of somatic mutations in BER pathway (**Supplementary Figure S3**).

Association Between DNA-Damage Repair Mutation and Immune Cell Infiltration Pattern in The Cancer Genome Atlas Cohort

Previous study has demonstrated that mutations can generate novel peptide sequences, which may affect the immune response (Chalmers et al., 2017). The TMB noted in the DDR mutation group was higher in the Chinese cohort and the TCGA cohort (**Figure 5D**, **Figure 7**). Thus, higher immune cell abundance was expected in the DDR somatic mutation. By applying the CIBERSORT algorithm, the differential variation of immune cell infiltration was estimated in the DDR and non-DDR groups of bowel cancer. The Wilcoxon rank-sum test indicated that the proportion of B cell naive (p = 0.017), T cell follicular helper (p = 0.0069), Macrophage M1 (p = 0.0038), and Neutrophils (p = 0.0163) were significantly elevated in the DDR group. By contrast, the infiltration levels of T cell

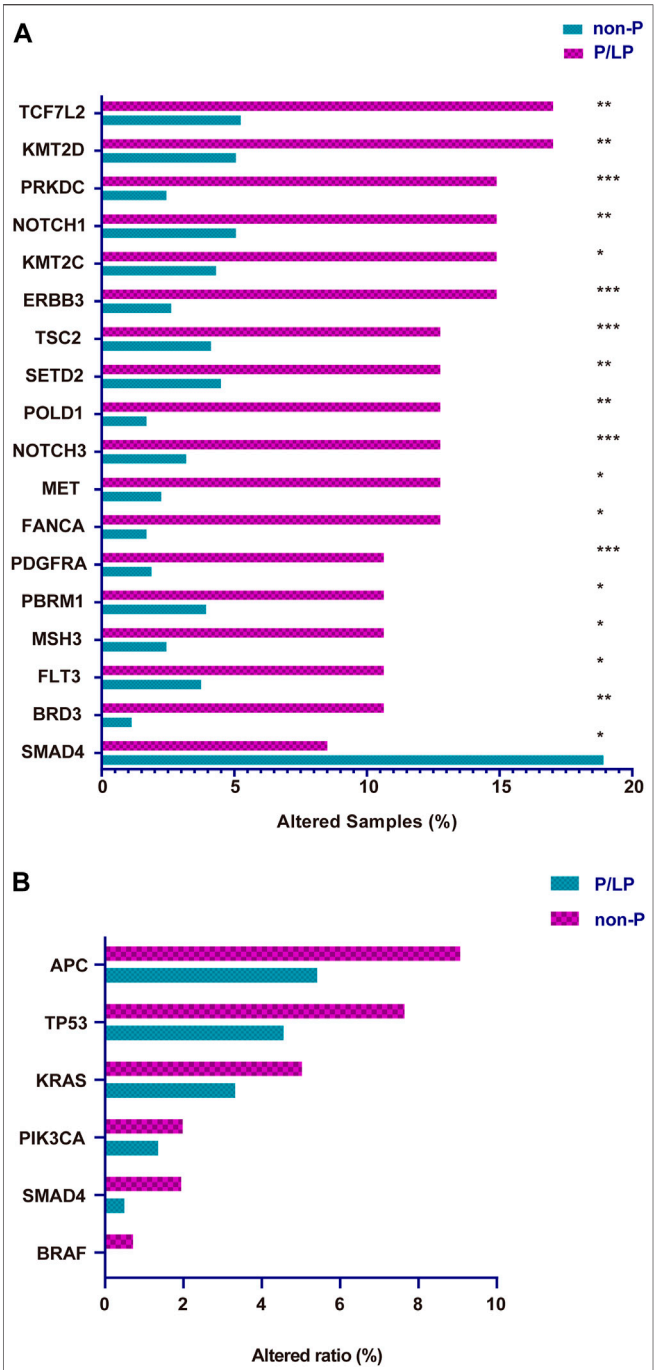


FIGURE 3 | The somatic gene variation rate between P/LP and Non-P groups for all patients in this study. **(A)** Differential expressing Mutation genes in P/LP and Non-P groups. **(B)** Comparison of the variation rate (mutational frequency) for driver genes with somatic mutations between P/LP and Non-P groups. P/LP: pathogenic/likely-pathogenic; Non-P: non-pathogenic.

regulatory Tregs (p = 0.0291) and Myeloid dendritic cell resting (p = 0.0163) were lower in the DDR group (**Figure 8A**). However, no association was observed between the different DDR pathway alterations and immune cell abundance (**Figure 8B**).

Association Between DNA-Damage Repair Mutations and Survival Outcomes

In addition, the present study further investigated whether the number of somatic mutations in the DDR genes of P/LP group was associated with improved survival following programmed cell death protein 1 (PD-1)/programmed death-ligand 1 (PD-L1) or cytotoxic T-lymphocyte-associated protein 4 immunotherapy in CRC patients. The clinical data of immunogenomic studies were downloaded from the cBioportal website. The overall survival (OS) was defined as the time from initial surgery to the date of death or last contact (censored). As expected, alteration in the mutation status of P/LP DDR conferred superior OS compared with patients without altered DDR gene(s) (HR, 0.3358; 95% CI, 0.1767 to 0.638; p = 0.0009) in the immunotherapy cohort (**Figure 9A**). However, patients with DDR mutations did not obtain a significantly prolonged OS when the MMR pathway was excluded (HR = 0.2998, 95%CI 0.1051 to 0.8556; p = 0.0549, **Figure 9B**).

DISCUSSION

Germline variants transmit genetic information that determines the heritability of complex disorders (McClellan et al., 2010). The presence of individually-rare but collectively common germline variants can explain a fraction of the missing genetic predisposition to bowel cancer. However, the major percentage of bowel cancer heritability is still not fully characterized, especially for the Chinese population. In the present study, an NGS-based analysis of germline mutations was performed for 573 Chinese patients with various stages of bowel cancer. The analysis provided a representative germline mutation landscape. The present study is the first to elucidate a more comprehensive germline mutation profile of Chinese patients with bowel cancer with the aim of identifying the novel candidate genes for hereditary bowel cancer. In addition, genetic testing and identification of germline mutations may have implications for the relatives of patients with bowel cancer because of the associated risks of CRC and other cancer types. In the present

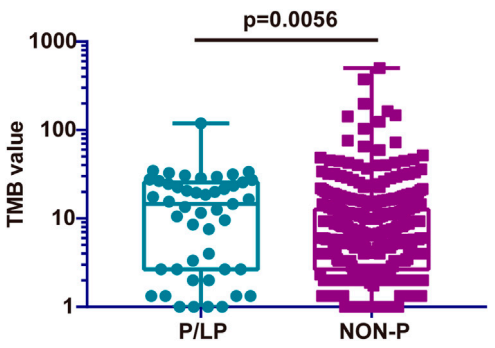


FIGURE 4 | Comparison of the TMB from somatic mutations of the P/LP and the Non-P groups. P/LP: pathogenic/likely-pathogenic; Non-P: non-pathogenic; TMB, tumor mutation burden.

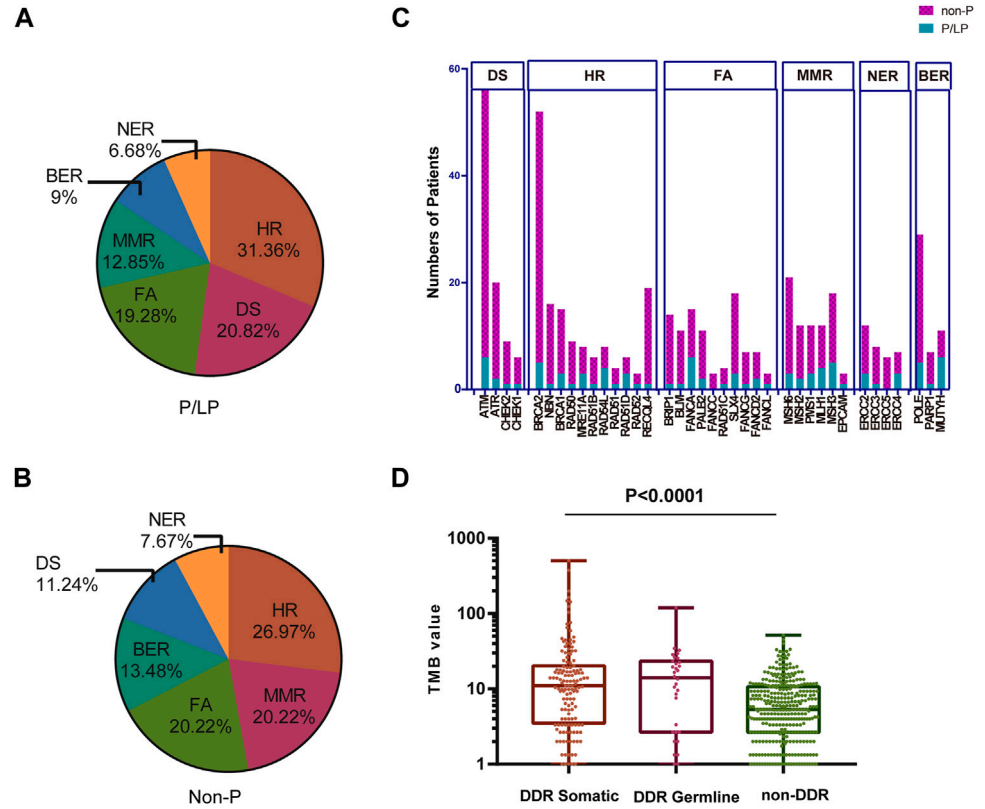


FIGURE 5 | Percentage of patients with bowel cancer harboring DDR gene mutations and the TMB in our cohort. The proportion of different DDR pathways in P/LP **(A)** and Non-P groups **(B)**. **(C)** Comparison of the number of patients with gene mutations in DDR pathways for P/LP and Non-P groups. **(D)** Comparison of the TMB from somatic mutations of the DDR somatic, DDR germline, and non-DDR groups. P/LP: pathogenic/likely-pathogenic; DDR: DNA damage response; TMB, tumor mutation burden.

Chinese cohort, the ratio of patients with bowel cancer with at least one family member (first- and second-degree relatives) with tumor history was significantly higher in the P/LP group than that in the Non-P group, indicating that pathogenic cancer-predisposing variants were associated with the incidence of bowel cancer and resulted in familial clustering. On the other hand, the patients with bowel cancer and their family members with a history of other cancers were included in the family history examined in the current study indicating that the presence of pathogenic germline mutations increased the incidence of other cancer types.

Germline pathogenic variants in genes encoding for DNA mismatch repair proteins cause Lynch syndrome, which is considered the most prevalent form of hereditary bowel cancer (Arora et al., 2015). Classic hereditary bowel cancer syndromes, including Lynch syndrome are mainly due to germline mutations in *APC, MUTYH*, and the MMR genes (such as *MLH1, MSH2, MSH6,* and *PMS2*). The data reported in the current study demonstrated that the top three genes with the highest number of germline mutations were *MUTYH* (5/47), *APC* (4/47), and *MLH1* (3/47), which was consistent with the findings of previous publications (Esteban-Jurado et al., 2015; Reilly et al., 2019). Meanwhile, the recent study also revealed the most common mutated genes of *TP53, APC, KRAS, SMAD4, PIK3CA* etc., besides, the mutation frequencies of *TP53* and *APC* in the left CRC were significantly higher than that of right CRC (Huang et al., 2021). While the germline alterations in certain susceptibility genes were also detected in the bowel cancer samples including *FANCD2, CDH1, FLCN, MEN1, SDHB,* and *SLX4,* which have been rarely reported in the Chinese population. As is known, *FANCD2* is the frequently mutated gene in colorectal cancer (Offman et al., 2005). *CDH1* mutations are more predisposed to familial colorectal cancer (Richards et al., 1999). Besides, it was reported that the frameshift mutations in the *FLCN* exon 11 which would suppress the activation of *FLCN* could lead to the increased incidence of colorectal cancer (Nahorski et al., 2010). While *MEN1* was reported that it could be a novel driver causing the dysregulation of Wnt signaling pathway in colorectal cancer (Fennell et al., 2020). SDHB is the catalytic core of succinate dehydrogenase (SDH), of which dysfunction would exert an influence on the TGF-beta signaling pathway contributing to the colorectal cancer formation (Wang et al., 2016). Moreover, the mutation of the tumor suppressor gene SLX4 was recently shown to be associated with the early-onset of CRC in the population of Kazakhstan (Zhunussova et al., 2019).Among these germline variants, mutations in *SLX4, FANCD2,* and *FLCN* are associated with FA pathway alteration (provided by RefSeq, NCBI). In other

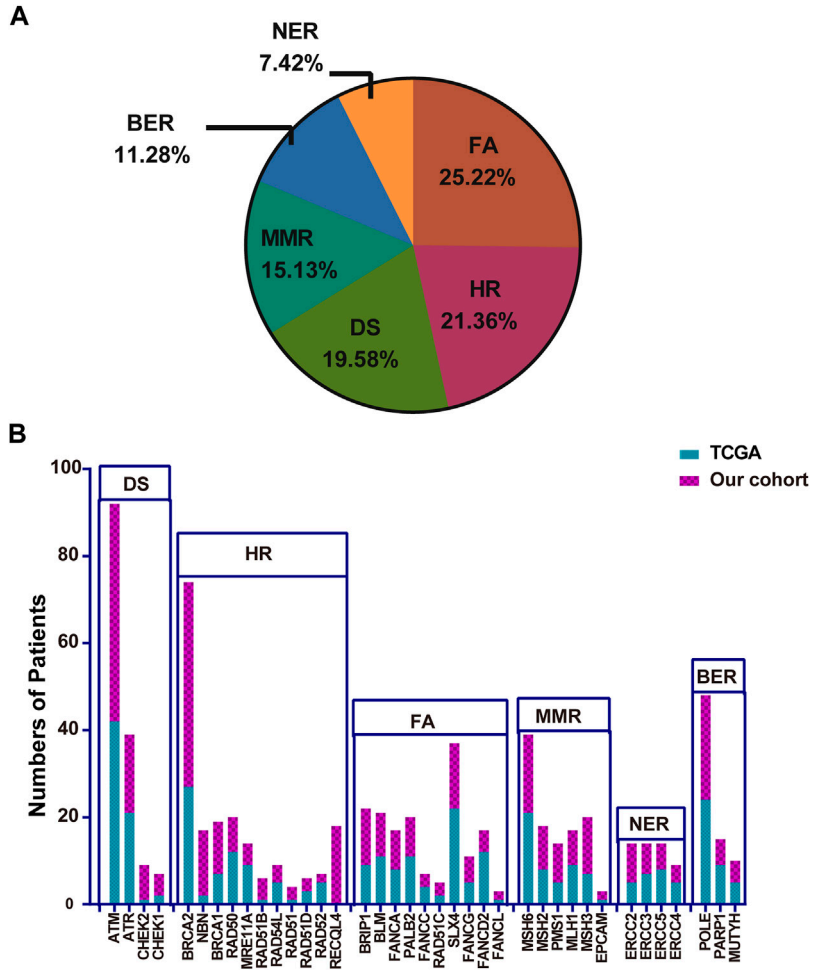


FIGURE 6 | The proportion of the detailed DDR pathways **(A)** and the number of patients with specific DDR pathway genes in TCGA (B). TCGA: The Cancer Genome Atlas; DDR: DNA damage response.

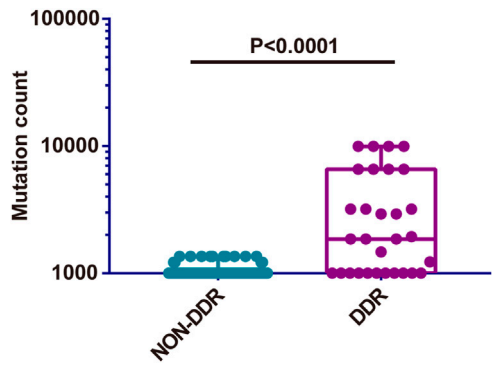


FIGURE 7 | Comparison of somatic mutation count between non-DDR and DDR groups from TCGA cohort. DDR: DNA damage response; TCGA: The Cancer Genome Atlas.

geographical regions, germline mutations of CRC patients mainly occur in the HR, DS, NER, BER, and MMR pathways (Berginc et al., 2009; Yurgelun et al., 2017; Ma et al., 2018; Schneider et al., 2018; Fujita et al., 2020). Although the present study detected germline mutations in these DDR pathways, a higher number of genes were identified which were involved in the FA pathway compared with the previous studies, such as *FANCD2, FANCA, FANCG, FANCL,* and *SLX4*. The identification of a wider causative mutation in bowel cancer has implications that can apply to genetic counseling practices that are of vital importance for the family under investigation (Wells and Wise, 2017). Once established in a particular family with carriers and non-carriers, prevention strategies can be directed more precisely to those subjects carrying the causative mutation and who are therefore at risk of developing bowel cancer and other related malignancies.

The frequency of mutations queried in the China_MAPs database represents the frequency of a certain mutation site in the general population. The OR of the cases investigated in the present study suggested that the germline mutations were risk factors for bowel cancer. In this case-control study, mutations in 8 genes (*APC, ATM, MLH1, FANCD2, MSH3, MSH6, PMS1,* and *RAD51D*) were found to be associated with bowel cancer and were present in 3.5% of patients with bowel cancer. Mutations in

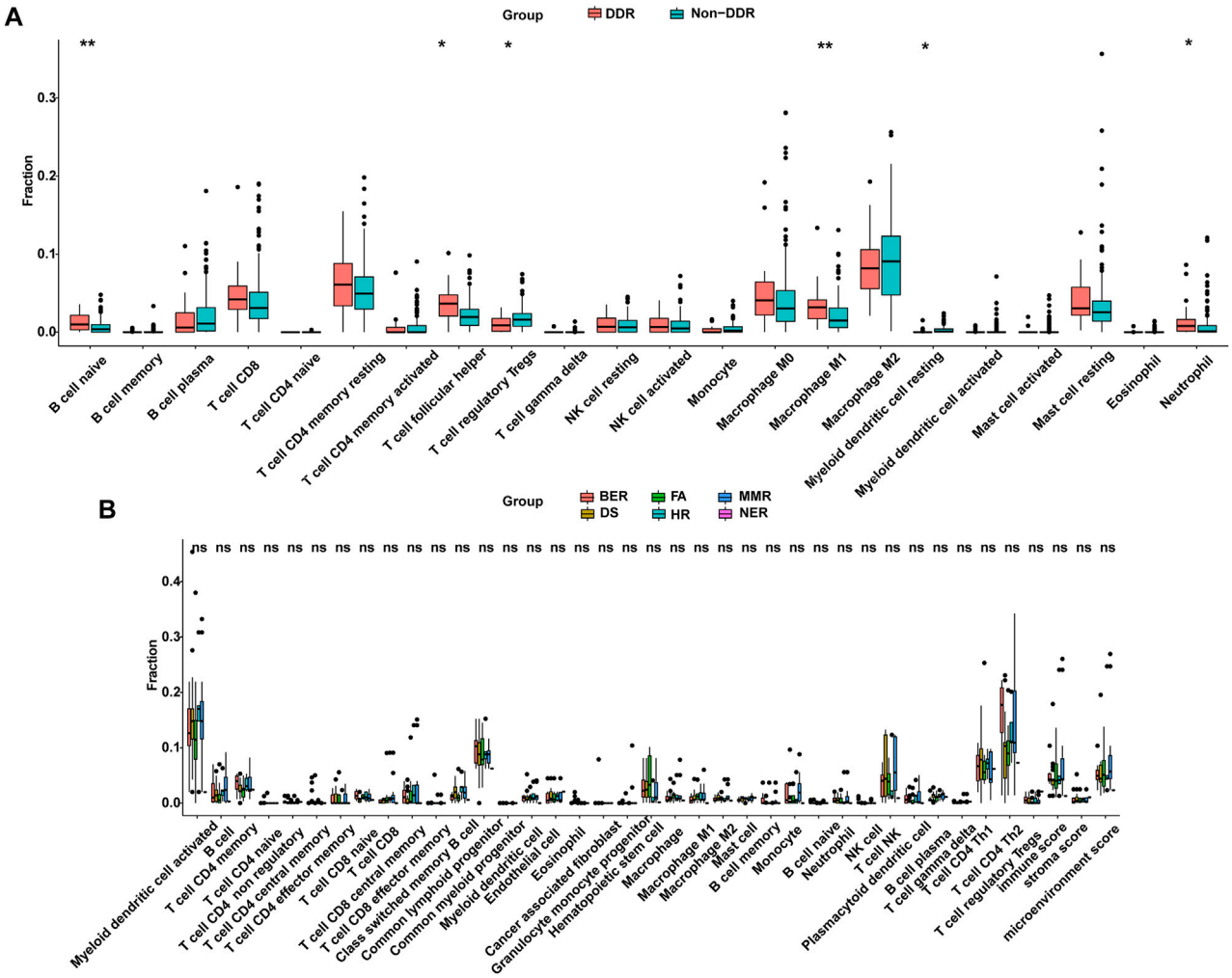


FIGURE 8 | DDR was correlated with the infiltration of immune cells from the TCGA cohort. **(A)** Differentially infiltrated immune cells between DDR and non-DDR groups. **(B)** Comparison of infiltrated immune cells in different altered DDR pathways. DDR: DNA damage response; TCGA: The Cancer Genome Atlas.

APC, ATM, and *MLH1* were associated with the highest risks of bowel cancer. The frequency of latter two germline variants was also the highest among patients with bowel cancer (0.7, 0.52, and 0.52%, respectively). In addition to commonly mutated genes such as *APC, ATM,* and LS-related genes that have been previously reported (Bernstein and Concannon, 2017; Snowsill et al., 2017; Aghabozorgi et al., 2019), the current analysis further revealed significantly higher rates of *FANCD2* and *RAD51D* mutations in bowel cancer germline mutation carriers than those of the general population, suggesting that *FANCD2* and *RAD51D* can be considered as bowel cancer-susceptibility genes. FANCD2 is monoubiquinated in response to DNA damage, resulting in its localization to nuclear foci with other proteins (BRCA1 and BRCA2) involved in homology-directed DNA repair. Furthermore, *RAD51D* is involved in the homologous recombination and repair of DNA (provided by RefSeq, NCBI). However, these genes are rarely reported in bowel cancer. Therefore, the current results provide preliminary evidence of potential susceptibility genes that can be used for hereditary bowel cancer.

The integration of germline and somatic genomic data can provide insight into the mechanisms that drive tumor progression (Ramroop et al., 2019). Therefore, an in-depth integrated analysis was performed on germline and somatic NGS data derived from the patients with bowel cancer. The present study identified several distinct somatic mutation rates between the carriers of germline mutations and the non-carriers. For example, differential expression gene analysis indicated that the mutation incidence of specific genes, such as *TCF7L2, KMT2D, PRKDC,* and *NOTCH1* was significantly higher in the P/LP group than that of the Non-P group, suggesting a higher risk compared with subjects without germline mutations. It is interesting to note that the *SMAD4* mutation rate in the P/LP group was dramatically lower than that of the Non-P group. Mutations or deletions in this gene have been shown to result in the development of juvenile polyposis syndrome, whereas weak expression of *SMAD4* is known to associate with poor survival in

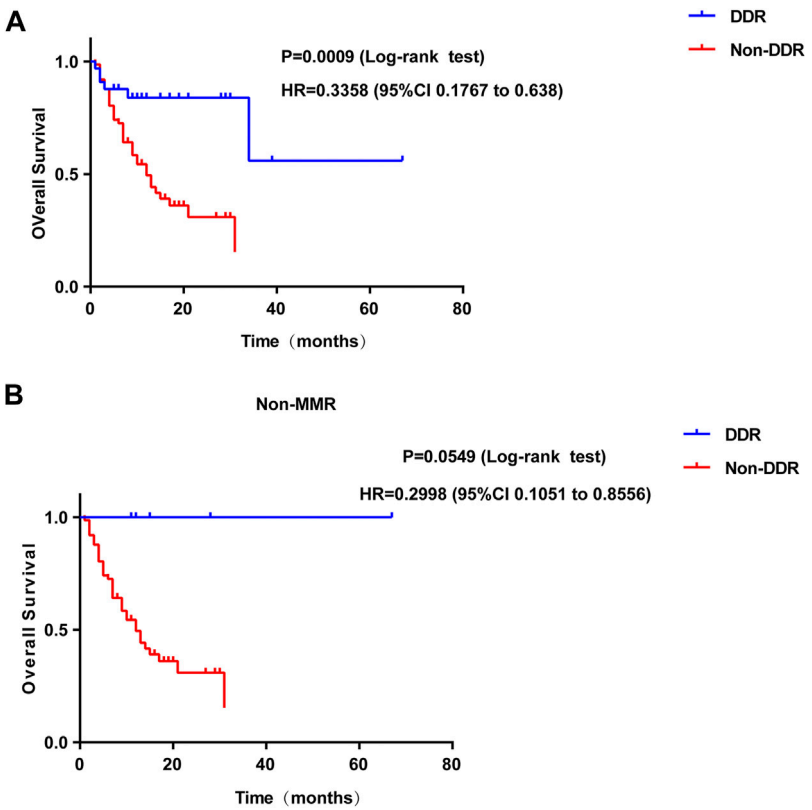


FIGURE 9 | Survival curves of OS comparing DDR mutation **(A)** or DDR without MMR mutation **(B)** (blue) with Non-DDR (red) in colorectal patients from immunotherapy cohort. OS: overall survival; DDR: DNA damage response.

patients with bowel cancer (Yan et al., 2016). The results in this study confirmed that *SMAD4* mutation mainly occurred in somatic mutation patients rather than germline carriers. One possible explanation for this outcome could be that the environmental factors mainly affected the Non-P germline mutation carriers.

It is noteworthy that somatic mutations in patients with P/LP germline mutations showed distinct characteristics in the present study. TMB, which is considered to be a predictive biomarker of immune checkpoint inhibitor (ICIs) (Wang and Li, 2019), may reflect the degree of genomic instability at the nucleotide level. The TMB in the P/LP group was significantly lower than that in the non-germline mutant group. It was speculated that somatic mutations in patients with P/LP germline mutations may be more focused on certain key genes and key pathways, whereas somatic mutations in patients without P/LP germline mutations may be more sporadic. Therefore, patients with P/LP germline mutations may be more likely to have abnormalities in key genes and pathways, leading to an increased risk of bowel cancer. Moreover, TMB in the non-altered DDR group was significantly lower than that noted in the altered DDR groups. In addition, higher TMB was observed in the DDR of the non-germline mutation group compared with that of the DDR germline mutation group. A possible interpretation is that the further alteration in DDR may induce a hypermutated phenotype with a higher TMB, which could be established as a biomarker of ICI treatment (Rizvi et al., 2015).

In the current comprehensive analysis, we also explored the association between DDR mutation and the number of immune cells among patients with bowel cancer from the TCGA cohort. The results found that the number of B cell naive, T cell follicular helper, Macrophage M1, and Neutrophil were elevated in the DDR group. Previous studies have demonstrated that the infiltration of B cells plays an important role in tumor immunotherapy (Cabrita et al., 2020; Helmink et al., 2020; Petitprez et al., 2020). Wei, et al. (2016) reported that B cell naive could suppress the antitumor adaptive immune response in the patients with ovarian cancer. An additional study proposed that B cell naive could be considered as a readily available and effective source of antigen-presenting cells in clinical research on tumor immunotherapy (Ren et al., 2014). Based on these findings, it was speculated that DDR alteration in bowel cancer might modulate the response to ICIs to a certain extent (Ren et al., 2014). In addition, the present study indicated that patients with DDR gene alterations were more likely to experience improved OS than patients with unaltered DDR genes. Although MMR mutations were excluded from the DDR pathway, patients with DDR

alterations did not show a significantly prolonged OS, and the p value(p = 0.0549) is nearly close to the significance cutoff. The limited samples (n = 6) may contribute to this phenomenon. Therefore, it is suggested that the predicted value of this association is investigated further in larger data sets from randomized studies that have led to the FDA approval of several anti-PD-1/PD-L1 agents. DDR alterations may represent a useful predictive bowel cancer biomarker for of the patient response to anti-PD-1/PD-L1 provided these findings are validated in a larger cohort (Samstein and Riaz, 2018; Zhang et al., 2020).

The ratio of DDR pathway alteration and other molecular results varied between the TCGA cohort and the Chinese cohort. The baseline characteristics, such as sex and tumor stage, may also contribute to the distinction.

One limitation of the present study is that the germline mutations were unavailable from TCGA cohort. This cannot comprehensive assessment of the differences between P/LP germline mutation carriers and Non-P germline mutation carriers in distinct bowel cancer populations. In addition, although the mutations were identified in the genomic sequences, their exact effects on the altered protein function were not assessed. Despite these limitations, the results can still provide a reasonable basis for exploring the applications of the DDR germline mutations in the prognosis of hereditary bowel cancer.

In conclusion, the present study identified unique genomic and molecular characteristics such as TMB and DDR between P/LP germline alteration carriers and Non-P bowel cancer patients. A preliminary basis was provided for the assessment of a wider range of susceptibility genes in Chinese CRC patients. Moreover, the TCGA database indicated that a deeper understanding of the interactions between DDR and immune cell infiltration would be useful to further investigate the role of DDR in bowel cancer.

DATA AVAILABILITY STATEMENT

The raw data of next generation sequencing in this study are deposited in: https://bigd.big.ac.cn/gsa-human/browse/HRA001620 (the Chinese National Genomics Data Center) repository, accession number: HRA001620. The control data from the China Maps database were downloaded from the following: http://www.GenomAD.org.

AUTHOR CONTRIBUTIONS

ZX, YK contributed equally to this manuscript. ZX, YK, and LC designed this work; ZX: Data collection and drafting the article, YK: Data analysis and article revision, JC: Data collection and analysis, ZL: Data collection, CW, YC, and HD: Data collection, LC: write the manuscript draft and revise the article. All authors have read and approved the article.

FUNDING

This study was supported by the Special Funding of PLA.

ACKNOWLEDGMENTS

We are grateful of all participants in General Hospital of Southern Theatre Command, PLA.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.755629/full#supplementary-material

REFERENCES

Aghabozorgi, A. S., Bahreyni, A., Soleimani, A., Bahrami, A., Khazaei, M., Ferns, G. A., et al. (2019). "Role of Adenomatous Polyposis Coli (APC) Gene Mutations in the Pathogenesis of Colorectal Cancer; Current Status and Perspectives,". *Biochimie*, 157, 64–71. doi:10.1016/j.biochi.2018.11.003

Arnold, M., Abnet, C. C., Neale, R. E., Vignat, J., Giovannucci, E. L., McGlynn, K. A., et al. (2020). Global Burden of 5 Major Types of Gastrointestinal Cancer, *Gastroenterology* 159 (1), 335–349. doi:10.1053/j.gastro.2020.02.068

Arora, S., Yan, H., Cho, I., Fan, H.-Y., Luo, B., Gai, X., et al. (2015). "Genetic Variants that Predispose to DNA Double-Strand Breaks in Lymphocytes from a Subset of Patients with Familial Colorectal Carcinomas,". *Gastroenterol*, 149, 7, 1872–1883. e9. doi:10.1053/j.gastro.2015.08.052

Berginc, G., Bračko, M., Ravnik-Glavač, M., and Glavač, D. (2009). "Screening for Germline Mutations of MLH1, MSH2, MSH6 and PMS2 Genes in Slovenian Colorectal Cancer Patients: Implications for a Population Specific Detection Strategy of Lynch Syndrome,". *Fam. Cancer*, 8, 4, 421–429. doi:10.1007/s10689-009-9258-4

Bernstein, J. L., and Concannon, P. (2017). ATM, Radiation, and the Risk of Second Primary Breast Cancer. *Int. J. Radiat. Biol.* 93 (10), 1121–1127. doi:10.1080/09553002.2017.1344363

Bertelsen, B., Tuxen, I. V., Yde, C. W., Gabrielaite, M., Torp, M. H., Kinalis, S., et al. (2019). High Frequency of Pathogenic Germline Variants within Homologous Recombination Repair in Patients with Advanced Cancer. *Npj Genom. Med.* 4, 13. doi:10.1038/s41525-019-0087-6

Bhui, K., Prasad, S., George, J., and Shukla, Y. (2009). Bromelain Inhibits COX-2 Expression by Blocking the Activation of MAPK Regulated NF-Kappa B against Skin Tumor-Initiation Triggering Mitochondrial Death Pathway. *Cancer Letterscancer Lett.* 282 (2), 167–176. doi:10.1016/j.canlet.2009.03.00310.1016/j.canlet.2020.01.017

Bien, S. A., Su, Y. R., Conti, D. V., Harrison, T. A., Qu, C., Guo, X., et al. (2019). Genetic Variant Predictors of Gene Expression Provide New Insight into Risk of Colorectal Cancer. *Hum. Genet.* 138 (4), 307–326. doi:10.1007/s00439-019-01989-8

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L, Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.*, 68, 6, 394–424. doi:10.3322/caac.21492

Cabrita, R., Lauss, M., Sanna, A., Donia, M., Skaarup Larsen, M., Mitra, S., et al. (2020). "Tertiary Lymphoid Structures Improve Immunotherapy and Survival in Melanoma,". *Nature*, 577, 7791, 561–565. doi:10.1038/s41586-019-1914-8

Chalmers, Z. R., Connelly, C. F., Fabrizio, D., Gay, L., Ali, S. M., Ennis, R., et al. (2017). Analysis of 100,000 Human Cancer Genomes Reveals the Landscape of Tumor Mutational burden. *Genome Med.* 9 (1), 34. doi:10.1186/s13073-017-0424-2

Dow, L. E., O'Rourke, K. P., Simon, J., Tschaharganeh, D. F., van Es, J. H., Clevers, H., et al. (2015). Apc Restoration Promotes Cellular Differentiation and Reestablishes Crypt Homeostasis in Colorectal Cancer. *Cell* 161 (7), 1539–1552. doi:10.1016/j.cell.2015.05.033

Eklöf, V., Wikberg, M. L., Edin, S., Dahlin, A. M., Jonsson, B.-A., Öberg, A., et al. (2013). The Prognostic Role of KRAS, BRAF, PIK3CA and PTEN in Colorectal Cancer. *Br. J. Cancer* 108 (10), 2153–2163. doi:10.1038/bjc.2013.212

Engel, C., Ahadova, A., Seppälä, T. T., Aretz, S., Bigirwamungu-Bargeman, M., Bläker, H., et al. (2020). "Associations of Pathogenic Variants in MLH1, MSH2, and MSH6 with Risk of Colorectal Adenomas and Tumors and with Somatic Mutations in Patients with Lynch Syndrome,"*Gastroenterology*, 158, 5, 1326–1333. doi:10.1053/j.gastro.2019.12.032

Esteban-Jurado, C., au, fnm., Vila-Casadesús, M., Garre, P., Lozano, J. J., Pristoupilova, A., et al. (2015). Whole-exome Sequencing Identifies Rare Pathogenic Variants in New Predisposition Genes for Familial Colorectal Cancer. *Genet. Med.* 17 (2), 131–142. doi:10.1038/gim.2014.89

Fennell, L. J., Kane, A., Liu, C., McKeone, D., Fernando, W., Su, C., et al. (2020). APC Mutation Marks an Aggressive Subtype of BRAF Mutant Colorectal Cancers. *Cancers* 12 (5), 1171. doi:10.3390/cancers12051171

Fujita, M., Liu, X., Iwasaki, Y., Terao, C., Mizukami, K., Kawakami, E., et al. (2020). Population-based Screening for Hereditary Colorectal Cancer Variants in Japan. *Clin. Gastroenterol. Hepatol.* doi:10.1016/j.cgh.2020.12.007

Gentiluomo, M., Canzian, F., Nicolini, A., Gemignani, F., Landi, S., and Campa, D. (2020). "Germline Genetic Variability in Pancreatic Cancer Risk and Prognosis,". *Semin. Cancer Biol.* doi:10.1016/j.semcancer.2020.08.003

Giri, V. N., Knudsen, K., Kelly, W., Cheng, H., Cooney, K., Cookson, M., et al. "Implementation of Germline Testing for Prostate Cancer: Philadelphia Prostate Cancer Consensus Conference 2019,". *J. Clin. Oncol.*, 38, 24, 2798–2811. doi:10.1200/jco.20.00046

Helmink, B. A., Reddy, S. M., Gao, J., Zhang, S., Basar, R., Thakur, R., et al. (2020). B Cells and Tertiary Lymphoid Structures Promote Immunotherapy Response. *Nature* 577 (7791), 549–555. doi:10.1038/s41586-019-1922-8

Huang, W., Li, H., Shi, X., Lin, M., Liao, C., Zhang, S., et al. (2021). Characterization of Genomic Alterations in Chinese Colorectal Cancer Patients. *Jpn. J. Clin. Oncol.* 51 (1), 120–129. doi:10.1093/jjco/hyaa182

Latham, A., Srinivasan, P., Kemel, Y., Shia, J., Bandlamudi, C., Mandelker, D., et al. (2019). Microsatellite Instability Is Associated with the Presence of Lynch Syndrome Pan-Cancer. *J. Clin. Oncol.* 37 (4), 286–295. doi:10.1200/jco.18.00283

Li, M. M., Datto, M., Duncavage, E. J., Kulkarni, S., Lindeman, N. I., Roy, S., et al. (2017). Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J. Mol. Diagn.* 19 (1), 4–23. doi:10.1016/j.jmoldx.2016.10.002

Liu, M., Liu, X., Suo, P., Gong, Y., Qu, B., Peng, X., et al. (2020). The Contribution of Hereditary Cancer-Related Germline Mutations to Lung Cancer Susceptibility. *Transl Lung Cancer Res.* 9 (3), 646–658. doi:10.21037/tlcr-19-403

Ma, H., Brosens, L. A. A., Offerhaus, G. J. A., Giardiello, F. M., de Leng, W. W. J., and Montgomery, E. A. (2018). Pathology and Genetics of Hereditary Colorectal Cancer. *Pathology* 50 (1), 49–59. doi:10.1016/j.pathol.2017.09.004

McClellan, J., and King, M.-C. (2010). "Genetic Heterogeneity in Human Disease,". *Cell*, 141, 2, 210–217. doi:10.1016/j.cell.2010.03.032

Medina Pabón, M. A., and Babiker, H. M. (2021). *StatPearls. Treasure Island (FL).* StatPearls Publishing Copyright © 2021StatPearls Publishing LLC., Treasure Island, FL, USA, .

Mueller, M., Schneider, M. A., Deplazes, B., Cabalzar-Wondberg, D., Rickenbacher, A., and Turina, M. (2021). "Colorectal Cancer of the Young Displays Distinct Features of Aggressive Tumor Biology: A Single-center Cohort Study,". *Wjgs*, 13, 2, 164–175. doi:10.4240/wjgs.v13.i2.164

Nahorski, M. S., Lim, D. H. K., Martin, L., Gille, J. J. P., McKay, K., Rehal, P. K., et al. (2010). Investigation of the Birt-Hogg-Dube Tumour Suppressor Gene (FLCN) in Familial and Sporadic Colorectal Cancer. *J. Med. Genet.* 47 (6), 385–390. doi:10.1136/jmg.2009.073304

Offman, J., Gascoigne, K., Bristow, F., Macpherson, P., Bignami, M., Casorelli, I., et al. (2005). Repeated Sequences in CASPASE-5 and FANCD2 but Not NF1 Are Targets for Mutation in Microsatellite-Unstable Acute Leukemia/myelodysplastic Syndrome. *Mol. Cancer Res.* 3 (5), 251–260. doi:10.1158/1541-7786.Mcr-04-0182

Petitprez, F., de Reyniès, A., Keung, E. Z., Chen, T. W.-W., Sun, C.-M., Calderaro, J., et al. (2020). B Cells Are Associated with Survival and Immunotherapy Response in Sarcoma. *Nature* 577 (7791), 556–560. doi:10.1038/s41586-019-1906-8

Ramroop, J. R., Gerber, M. M., and Toland, A. E. (2019). Germline Variants Impact Somatic Events during Tumorigenesis. *Trends Genet.* 35 (7), 515–526. doi:10.1016/j.tig.2019.04.005

Reilly, N. M., Novara, L., Di Nicolantonio, F., and Bardelli, A. (2019). Exploiting DNA Repair Defects in Colorectal Cancer. *Mol. Oncol.* 13 (4), 681–700. doi:10.1002/1878-0261.12467

Ren, H., Zhao, S., Li, W., Dong, H., Zhou, M., Cao, M., et al. (2014). Therapeutic Antitumor Efficacy of B Cells Loaded with Tumor-Derived Autophagasomes Vaccine (DRibbles). *J. Immunother.* 37 (8), 383–393. doi:10.1097/cji.0000000000000051

Richards, F. M., McKee, S. A., Rajpar, M. H., Cole, T. R. P., Evans, D. G. R., Jankowski, J. A., et al. (1999). Germline E-Cadherin Gene (CDH1) Mutations Predispose to Familial Gastric Cancer and Colorectal Cancer. *Hum. Mol. Genet.* 8 (4), 607–610. doi:10.1093/hmg/8.4.607

Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., et al. (2015). Cancer Immunology. Mutational Landscape Determines Sensitivity to PD-1 Blockade in Non-small Cell Lung Cancer. *Science* 348 (6230), 124–128. doi:10.1126/science.aaa1348

Sa, R., Song, H., Wei, M., Su, H., Hong, Y., Zhou, L., et al. (2020). The Impact of APC Polymorphisms on the Transition from Polyps to Colorectal Cancer (CRC). *Gene* 740, 144486. doi:10.1016/j.gene.2020.144486

Samstein, R. M., Lee, C.-H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor Mutational Load Predicts Survival after Immunotherapy across Multiple Cancer Types. *Nat. Genet.* 51 (2), 202–206. doi:10.1038/s41588-018-0312-8

Samstein, R. M., and Riaz, N. (2018). The DNA Damage Response in Immunotherapy and Radiation. *Adv. Radiat. Oncol.* 3 (4), 527–533. doi:10.1016/j.adro.2018.08.017

Schneider, N. B., Pastor, T., Paula, A. E. d., Achatz, M. I., Santos, Â. R. d., Vianna, F. S. L., et al. (2018). Germline MLH1, MSH2 and MSH6 Variants in Brazilian Patients with Colorectal Cancer and Clinical Features Suggestive of Lynch Syndrome. *Cancer Med.* 7 (5), 2078–2088. doi:10.1002/cam4.1316

Snowsill, T., Coelho, H., Huxley, N., Jones-Hughes, T., Briscoe, S., Frayling, I. M., et al. (2017). "Molecular Testing for Lynch Syndrome in People with Colorectal Cancer: Systematic Reviews and Economic Evaluation,". *Health Technol. Assess.*, 21, 51, 1–238. doi:10.3310/hta21510

Sokic-Milutinovic, A. (2019). "Appropriate Management of Attenuated Familial Adenomatous Polyposis: Report of a Case and Review of the Literature,". *Dig. Dis.*, 37, 5, 400–405. doi:10.1159/000497207

Stadler, Z. K., Battaglin, F., Middha, S., Hechtman, J. F., Tran, C., Cercek, A., et al. (2016). Reliable Detection of Mismatch Repair Deficiency in Colorectal Cancers Using Mutational Load in Next-Generation Sequencing Panels. *Jco* 34 (18), 2141–2147. doi:10.1200/jco.2015.65.1067

Thanikachalam, K., and Khan, G. (2019). Colorectal Cancer and Nutrition. *Nutrients* 11 (1), 164. doi:10.3390/nu11010164

Wang, H., Chen, Y., and Wu, G. (2016). SDHB Deficiency Promotes TGFβ-Mediated Invasion and Metastasis of Colorectal Cancer through Transcriptional Repression Complex SNAIL1-Smad3/4. *Translational Oncol.* 9 (6), 512–520. doi:10.1016/j.tranon.2016.09.009

Wang, X., and Li, M. (2019). Correlate Tumor Mutation burden with Immune Signatures in Human Cancers. *BMC Immunol.* 20 (1), 4. doi:10.1186/s12865-018-0285-5

Wang, Z., Zhao, J., Wang, G., Zhang, F., Zhang, Z., Zhang, F., et al. (2018). Comutations in DNA Damage Response Pathways Serve as Potential Biomarkers for Immune Checkpoint Blockade. *Cancer Res.* 78 (22), 6486–6496. doi:10.1158/0008-5472.Can-18-1814

Wei, X., Jin, Y., Tian, Y., Zhang, H., Wu, J., Lu, W., et al. (2016). Regulatory B Cells Contribute to the Impaired Antitumor Immunity in Ovarian Cancer Patients. *Tumor Biol.* 37 (5), 6581–6588. doi:10.1007/s13277-015-4538-0

Wells, K., and Wise, P. E. (2017). Hereditary Colorectal Cancer Syndromes. *Surg. Clin. North America* 97 (3), 605–625. doi:10.1016/j.suc.2017.01.009

Wu, H.-X., Wang, Z.-X., Zhao, Q., Wang, F., and Xu, R.-H. (2019). "Designing Gene Panels for Tumor Mutational burden Estimation: the Need to Shift from 'correlation' to 'accuracy',"*J. Immunotherapy Cancer*. 7. 1. doi:10.1186/s40425-019-0681-2

Yan, P., Klingbiel, D., Saridaki, Z., Ceppa, P., Curto, M., McKee, T. A., et al. (2016). Reduced Expression of SMAD4 Is Associated with Poor Survival in Colon Cancer. *Clin. Cancer Res.* 22 (12), 3037–3047. doi:10.1158/1078-0432.Ccr-15-0939

Yang, Y., Han, Z., Han, Z., Li, X., Huang, A., Shi, J., et al. (2020). Epidemiology and Risk Factors of Colorectal Cancer in China. *Chin. J. Cancer Res.* 32 (6), 729–741. doi:10.21147/j.issn.1000-9604.2020.06.06

Yurgelun, M. B., Kulke, M. H., Fuchs, C. S., Allen, B. A., Uno, H., Hornick, J. L., et al. (2017). Cancer Susceptibility Gene Mutations in Individuals with Colorectal Cancer. *Jco* 35 (10), 1086–1095. doi:10.1200/jco.2016.71.0012

Zhang, J., Shih, D. J. H., and Lin, S.-Y. (2020). Role of DNA Repair Defects in Predicting Immunotherapy Response. *Biomark Res.* 8, 23. doi:10.1186/s40364-020-00202-7

Zhunussova, G., Afonin, G., Abdikerim, S., Jumanov, A., Perfilyeva, A., Kaidarova, D., et al. (2019). Mutation Spectrum of Cancer-Associated Genes in Patients with Early Onset of Colorectal Cancer. *Front. Oncol.* 9, 673. doi:10.3389/fonc.2019.00673

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 Genetics

ORIGINAL RESEARCH
published: 09 March 2022
doi: 10.3389/fgene.2022.823728

Identification of a Novel Glycosyltransferase Prognostic Signature in Hepatocellular Carcinoma Based on LASSO Algorithm

Zhiyang Zhou[1], *Tao Wang*[2], *Yao Du*[1], *Junping Deng*[1], *Ge Gao*[1]* *and Jiangnan Zhang*[1]*

[1]Department of General Surgery, The First Affiliated Hospital of Nanchang University, Nanchang, China, [2]Department of Day Ward, The First Affiliated Hospital of Nanchang University, Nanchang, China

OPEN ACCESS

Edited by:
Jaspreet Kaur Dhanjal,
Indraprastha Institute of Information Technology Delhi, India

Reviewed by:
Mriganko Das,
Ministry of Science and Technology, United States
Alok Kumar,
Kyoto University, Japan

***Correspondence:**
Ge Gao
gaogebj@hotmail.com
Jiangnan Zhang
zjnss@sina.com

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 28 November 2021
Accepted: 23 February 2022
Published: 09 March 2022

Citation:
Zhou Z, Wang T, Du Y, Deng J, Gao G and Zhang J (2022) Identification of a Novel Glycosyltransferase Prognostic Signature in Hepatocellular Carcinoma Based on LASSO Algorithm. Front. Genet. 13:823728. doi: 10.3389/fgene.2022.823728

Although many prognostic models have been developed to help determine personalized prognoses and treatments, the predictive efficiency of these prognostic models in hepatocellular carcinoma (HCC), which is a highly heterogeneous malignancy, is less than ideal. Recently, aberrant glycosylation has been demonstrated to universally participate in tumour initiation and progression, suggesting that dysregulation of glycosyltransferases can serve as novel cancer biomarkers. In this study, a total of 568 RNA-sequencing datasets of HCC from the TCGA database and ICGC database were analysed and integrated via bioinformatic methods. LASSO regression analysis was applied to construct a prognostic signature. Kaplan–Meier survival, ROC curve, nomogram, and univariate and multivariate Cox regression analyses were performed to assess the predictive efficiency of the prognostic signature. GSEA and the "CIBERSORT" R package were utilized to further discover the potential biological mechanism of the prognostic signature. Meanwhile, the differential expression of the prognostic signature was verified by western blot, qRT–PCR and immunohistochemical staining derived from the HPA. Ultimately, we constructed a prognostic signature in HCC based on a combination of six glycosyltransferases, whose prognostic value was evaluated and validated successfully in the testing cohort and the validation cohort. The prognostic signature was identified as an independent unfavourable prognostic factor for OS, and a nomogram including the risk score was established and showed the good performance in predicting OS. Further analysis of the underlying mechanism revealed that the prognostic signature may be potentially associated with metabolic disorders and tumour-infiltrating immune cells.

Keywords: glycosyltransferase, hepatocellular carcinoma, overall survival, prognostic signature, lasso regression analysis

INTRODUCTION

Hepatocellular carcinoma (HCC) is a highly aggressive solid malignancy and the fourth leading cause of cancer-related death, which imposes a tremendous health and socioeconomic burden globally (Singal, et al., 2020). Studies have shown that hepatitis virus infection, alcohol-related liver disease (ALD), non-alcoholic fatty liver disease (NAFLD) and non-alcoholic liver steatohepatitis (NASH) are the main aetiological risk factors for the development of HCC. Chronic infections with hepatitis virus are still the strongest risk factors for HCC in developing countries, nevertheless, NAFLD is gradually becoming the leading cause of HCC in Western countries (Huang, et al., 2021a). Despite all efforts made in the past to improve the prognosis of HCC, the prognosis remains poor, with an overall 5-years survival rate of approximately 18%, which is only slightly better than that of pancreatic cancer (Jemal, et al., 2017).

Of note, the prediction of clinical outcomes provides vital and necessary medical information. The traditional TNM staging system, which mainly relies on clinicopathological parameters, cannot provide a precise prediction of prognosis in clinical practice. In particular, HCC is a malignant tumour with the characteristic of high heterogeneity, which adds to the complexity of accurately predicting prognosis. One possible strategy to improve predictive outcome is to better understand the fundamental biological processes of cancer cells, and to identify prognostic signatures to stratify patients for individualized precision therapies based on prognosis and metastatic potential.

Glycosylation, the most universal protein post-translational modification, is an enzymatic process that catalyses the transfer of carbohydrate chains to proteins by glycosyltransferases (GTs) and glycosidases (Pinho and Reis, 2015). So far, 14 distinct human protein glycosylation pathways have been outlined, which are directed by at least 173 different GTs (Schjoldager, et al., 2020). They are divided into four main types: N-glycosylation, O- glycosylation, C-mannosylation and glypiation. Modified proteins are involved in nearly all biological processes, especially intercellular signal transduction and the immune response (Johannssen and Lepenies, 2017; Indellicato and Trinchera, 2021).

Alterations in cellular glycosylation have been recognized as hallmarks of malignant tumours, which contribute to sustaining proliferative signalling and metabolism, promoting invasion and metastasis, and immune evasion (Munkley and Elliott, 2016; Thomas, et al., 2021; Dobie and Skropeta, 2021; Rodrigues, et al., 2021). The under- or overexpression of GTs is the main contributor to cancer initiation and progression. Fucosylation is one of the most common modifications in the glycosylation pattern of HCC (Zhang, et al., 2017). The core fucosylation of α-fetoprotein (AFP-L3), a typical modified product, has already been confirmed as a biomarker in detecting early HCC (Wu, et al., 2014a; Noda, et al., 1998). As key enzymes of fucosylation, FUT1 (Kuo, et al., 2017), FUT8 (Cheng, et al., 2016) and POFUT1 (Ma, et al., 2016) are highly expressed and positively associated with advanced stage and poor prognosis in HCC. Other members of the FUT family, such as FUT2 (Wu, et al., 2014b), FUT4 (Cheng, et al., 2013), FUT6 (Guo, et al., 2012), and FUT7 (Wang, et al., 2005), are also known to support the development of HCC. Similarly, aberrant O-GlcNAcylation due to dysregulation of O-linked N-acetylglucosamine (GlcNAc) transferase (OGT) expression has been shown in HCC (Makwana, et al., 2019). OGT can promote migration by regulating FOXA2 stability and transcriptional activity (Huang, et al., 2021b), and the stem-like cell potential through O-GlcNAcylation of eIF4E (Cao, et al., 2019). In addition, abnormal expression of other kinds of GTs has been described in HCC, including C1GALT1, GALNT1, GALNT2, GALNT4, MGAT4A, MGAT5, B3GALT5, B4GALT4, ST3GAL1, ST3GAL2, ST3GAL6 and ST6GAL1. Cumulative findings indicate that abnormal expression of GTs seems to be a general feature of cancer cells and contributes to tumorigenesis and additional malignant characteristics.

Given the diversity of GTs and the high heterogeneity in individuals, a comprehensive understanding of the crucial role of aberrant glycosylation in HCC progression can further provide assistance in predicting prognosis. Therefore, the development of a novel evaluation index of glycosylation may be very useful for prognosis research. In this study, we developed a 6-gene prognostic signature that focused on the prognostic value of GT in HCC and validated its predictive capability through a variety of computational approaches.

MATERIALS AND METHODS

Data Screening and Gene Integration

Both the complete clinicopathologic information and matched RNA-sequencing FPKM data (HTSeq-FPKM) of HCC samples were extracted from The Cancer Genome Atlas (TCGA) data portal (up to September 23, 2021,https://portal.gdc.cancer.gov/). We also downloaded clinical and mRNA expression data of a Japanese cohort from the International Cancer Genome Consortium (ICGC) data portal (up to November 27,2019,https://dcc.icgc.org/).

Two glycosylation-related gene sets were obtained from the GlycoGene DataBase (GGDB: https://acgg.asia/ggdb2/index) and Hugo Gene Nomenclature Committee (HGNC: https://www.genenames.org/data/genegroup/#!/group/424). Gene set intersections were regarded as GT sets. Differentially expressed genes (DEGs) of the GT set between HCC and normal samples were identified using the "limma" package in R, and the screening criteria were FDR <0.05. Meanwhile, univariable Cox regression was employed to evaluate the association of each DEG with survival and results with a p value <0.05 were selected as prognosis-related genes. Finally, prognosis-related differentially expressed GTs were obtained.

Construction and Validation of the Prognostic Signature

All incorporated HCC samples from the TCGA database were randomly assigned to training and testing cohorts at a 1:1 ratio. A prognostic signature was constructed by applying the Least

Absolute Shrinkage and Selection Operator (LASSO) regression method, and the product of gene expression i and the corresponding coefficient β_i of each gene were added to establish the risk score: risk score = $\sum_{i=1}^{n}\beta_i * i$.

Utilizing the risk score formula, samples in the training cohort were categorized into high-risk and low-risk groups via the threshold of the median score. The Kaplan-Meier method was performed to compare survival differences between the two groups, and the prognostic value of the prognostic signature was shown by the receiver operating characteristic (ROC) curve. Simultaneously, we validated its prognostic performance with the TCGA testing cohort and ICGC external validation cohort.

Clinicopathological Features and Development of a Nomogram

Univariate and multivariate Cox regression analyses were performed to display the prognostic performance of this signature with other clinicopathological features.

A nomogram was developed to calculate individual's probability of overall survival (OS) by using the risk scores and clinical indicators. The final sum of the scores was expected to be the corresponding 1-, 2-, and 3-years survival probability.

Gene Set Enrichment Analysis and Correlation of Tumour-Infiltrating Immune Cells

GSEA was performed based on the gene matrix ("c2.cp.kegg.v7.4.symbols" and "c5.go.v7.4.symbols") between the high-risk and low-risk groups.

The CIBERSORT algorithm was used to calculate the relative abundance of 22 tumour-infiltrating immune cells in each sample of the TCGA dataset and ICGC dataset.

Cell Culture and the Experimental Validation *in vitro*

The LO2 human hepatocyte cell line and HepG2 human hepatoma cell line were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco, United States) supplemented with 10% foetal bovine serum (FBS, Gibco, United States) and incubated in a humidified atmosphere at 37 ˚C with 5% CO2.

Total RNA was extracted by using TransZol Up (TransGen Biotech, China) following the manufacturer's protocols. cDNA was synthesized using the PrimeScript RT reagent kit with gDNA Eraser (Takara, Japan), and mixed with primers (**Supplementary Table S1**) and TB Green Premix Ex Tap II (Takara, Japan), and run in the CFX96 Real-Time PCR Detection System (Bio–Rad, United States). The relative expression of the prognostic signature mRNA was calculated by the $2^{-\Delta\Delta Ct}$ method with GAPDH as the reference.

Total protein was prepared with RIPA buffer (Solarbio, China) with protease and phosphatase inhibitor cocktails (Solarbio, China). The protein levels were quantified by the BCA protein assay kit (Solarbio, China). Next, proteins were loaded onto 10% SDS-PAGE gel, separated electrophoretically, and transferred to PVDF membranes (Millipore, United States). After blocking with 5% non-fat milk for 1 h, the membranes were incubated at 4˚C overnight with primary antibodies against POMGNT1 (Immunoway, YT6311), B4GALT3 (Immunoway, YT5009), DPM1 (Proteintech, 12403-2-AP), B4GALT2 (Proteintech, 20330-1-AP), B4GALNT1 (Proteintech, 13396-1-AP), B3GAT3 (ABclonal, A20618), and GAPDH (Immunoway, YT5052). The next day, we incubated the PVDF membranes with HRP-conjugated secondary antibodies (mouse or rabbit) at room temperature for 1 h. The immunoblot signals were visualized using the hypersensitive ECL chemiluminescence detection kit (Proteintech, PK10003).

The protein expression levels of the prognostic signature were verified between normal tissues and cancer tissues from The Human Protein Atlas (HPA: https://www.proteinatlas.org/).

Statistical Analysis

All statistical analyses were performed with R software (version 4.0.4) and Strawberry Perl (version 5.32.0.1). LASSO regression analysis was applied to construct the prognostic signature. Nomogram construction and validation were performed using Iasonos' guide. The survival predictive accuracy of the risk assessment model was evaluated using time-dependent ROC curve analysis. Differences with $p < 0.05$ were considered statistically significant.

RESULTS

Dataset Characteristics and Candidate Gene Identification

The flow chart of this study design is depicted in **Figure 1**. In the TCGA dataset, 337 primary HCC samples and 39 normal samples were screened as training cohort and testing cohort. In the ICGC dataset, a total of 231 tumour samples with HCC were available for external validation. The detailed characteristics are shown in **Table 1**.

A total of 154 GT-related genes were generated by merging two gene sets, (**Supplementary Table S2**). Through initial analysis, 34 prognosis-related differentially expressed GTs overlapped as candidate genes for further analysis (**Supplementary Table S3** and **Figures 2A,B**).

Construction and Validation of Prognostic Signatures

Thirty-four candidate genes were used in LASSO regression analysis to confirm the core prognostic genes and to fit a risk prognosis model in the training cohort (n = 169). Finally, a prognostic risk score model comprising six genes (POMGNT1, DPM1, B4GALT3, B4GALT2, B4GALNT1, and B3GAT3) was constructed (**Table 2** and **Figures 2C,D**). The following formula was utilized: risk score = (0.002*expression level of POMGNT1) + (0.231*expression level of DPM1) + (0.222*expression level of B4GALT3) + (0.122* expression level of B4GALT2) + (0.212*

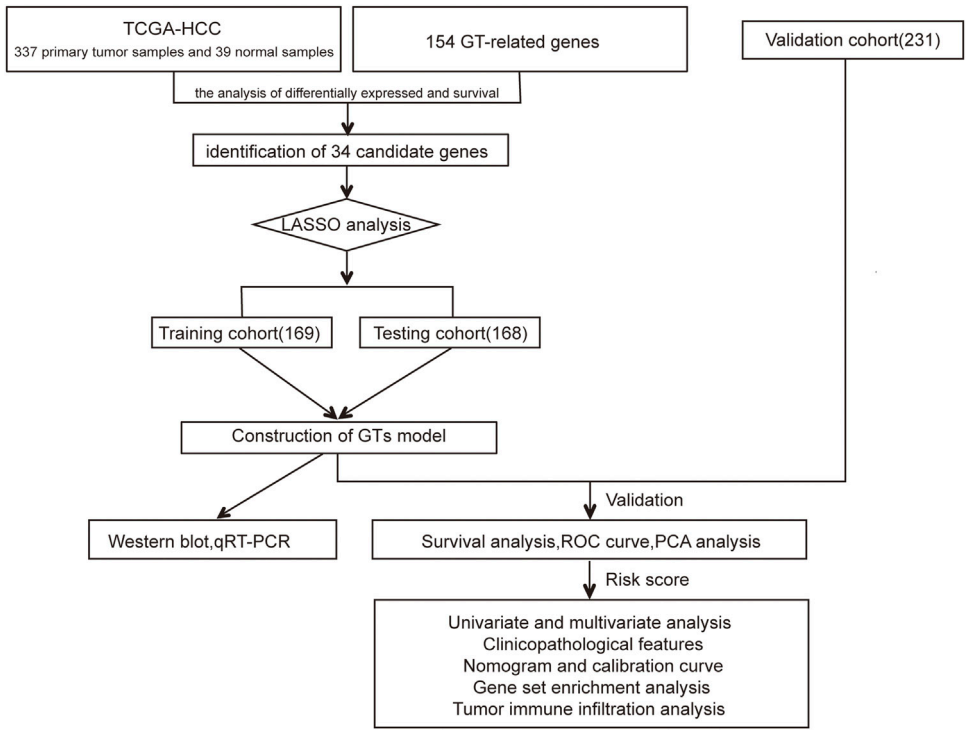


FIGURE 1 | Flowchart of our study.

TABLE 1 | Clinical characteristics of samples involved in this study.

Characteristics	TCGA dataset		ICGC dataset
	Training cohort	Testing cohort	Validation cohort
No. of samples	169	168	231
Age at diagnosis, years			
≤65	112	106	89
>65	57	62	142
Gender			
Female	52	55	61
Male	117	113	170
Grade			
G1-2	115	96	NA
G3-4	54	72	NA
TNM-stage			
StageI-II	127	123	141
StageIII-IV	42	45	90
T classification			
T1-2	128	125	NA
T3-4	41	43	NA

expression level of B4GALNT1) + (0.304*expression level of B3GAT3).

This formula was used to evaluate outcomes in each sample and the optimal cut-off value for samples in the high-risk group and low-risk group was set at the median risk score in the training cohort. Kaplan-Meier analysis revealed that a significantly inferior OS was reflected in the high-risk group than in the low-risk group in the training cohort, testing cohort and validation cohort (**Figures 3A–C**). Then, ROC curves were plotted to verify how well the risk score predicted the risk of death at years 1, 2, and 3 (**Figures 3D–F**). In **Figures 4A–C**, the risk score curves, risk gene expression heatmap and patient survival status are shown based on the risk score values. Furthermore, principal component analysis (PCA) was implemented to visualize the sample information by risk group (**Figure 4D**). The results proved that the prognostic

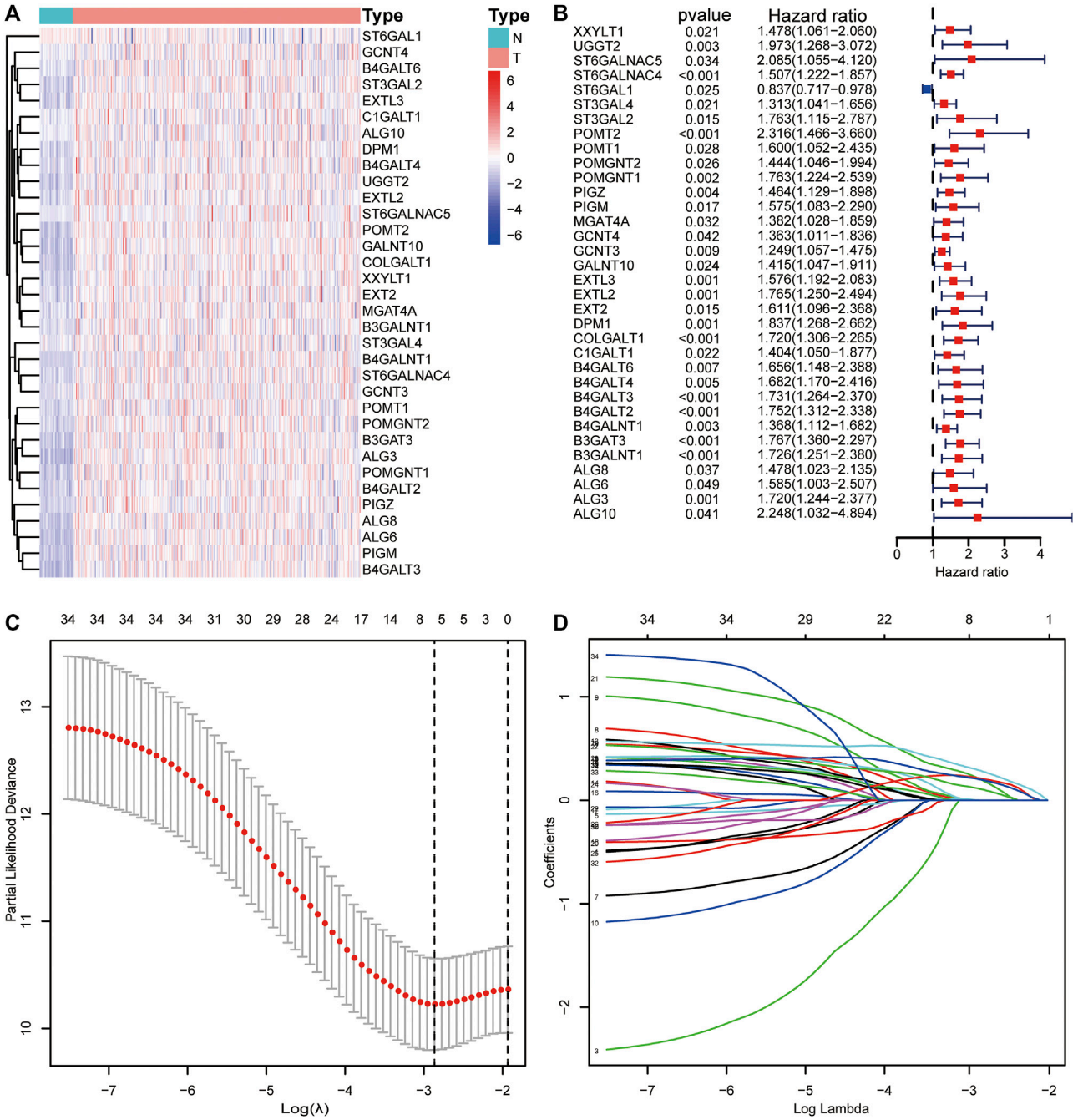


FIGURE 2 | Visualization of candidate genes. **(A)** Heatmap of the expression levels of candidate genes. **(B)** Forest plot of candidate genes. **(C)** Partial likelihood deviance of different combinations of variables calculated via the LASSO Cox regression model. **(D)** LASSO coefficient profiles of candidate genes.

signature based on these 6 candidate genes had good predictive performance for HCC patients.

Associations Between Risk Score and Clinicopathological Features

In the TCGA dataset, the risk score was significantly associated with OS in univariate Cox regression analysis (HR = 3.915, 95% CI = 2.516–6.092, $p < 0.001$, **Figure 5A**). Likewise, multivariate analysis showed that the risk score was an independent prognostic indicator in HCC (HR = 3.443, 95% CI = 2.163–5.481, $p < 0.001$, **Figure 5B**). The results from the ICGC dataset were consistent with the above (**Figures 5C,D**).

For further analyses, we created prognostic subgroups of patients based on multiple classification approaches in both datasets. The results showed that OS between the two groups

TABLE 2 | Detail information of the prognostic gene signatures.

Gene symbol	Gene name	Lasso coefficient
POMGNT1	Protein O-linked mannose beta 1,2- N-acetylglucosaminyltransferase 1	0.00221148
DPM1	Dolichyl-phosphate mannosyltransferase polypeptide 1	0.23073783
B4GALT3	Beta-1,4-galactosyltransferase 3	0.22198892
B4GALT2	Beta-1,4-galactosyltransferase 2	0.12189376
B4GALNT1	Beta-1,4-N-acetyl-galactosaminyltransferase 1	0.21161694
B3GAT3	Beta-1,3-glucuronyltransferase 3	0.30439163

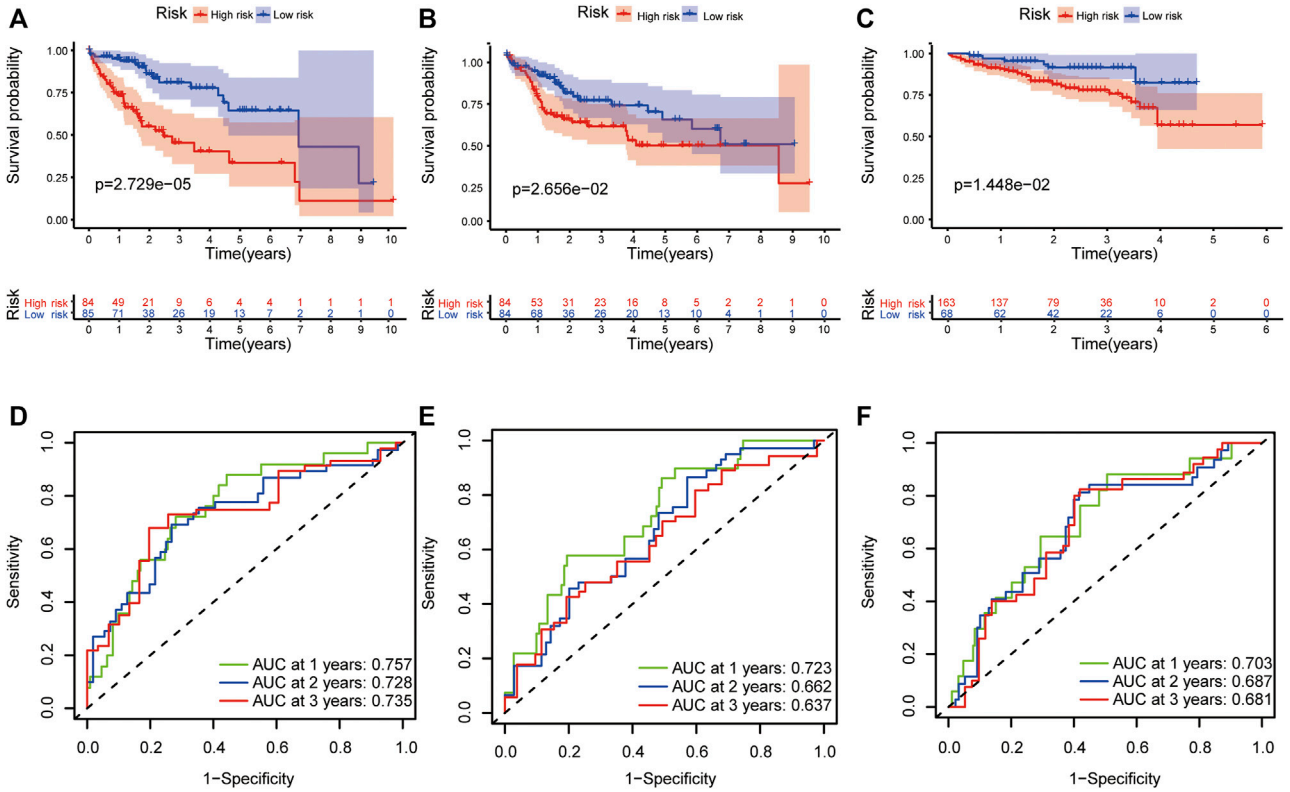


FIGURE 3 | Validation of the prognostic signature. The Kaplan–Meier survival plots of high-risk and low-risk groups in the training cohort **(A)**, testing cohort **(B)** and validation cohort **(C)**. The ROC curves of the prognostic signature in 1-, 2-, and 3-years survival in the training cohort **(D)**, testing cohort **(E)** and validation cohort **(F)**.

was significantly different in patients aged >65 years (**Figure 6A**, $p = 0.018$), ≤65 years (**Figure 6B**, $p < 0.001$), males (**Figure 6C**, $p < 0.001$), stage I-II (**Figure 6E**, $p = 0.010$) and stage III-IV (**Figure 6F**, $p = 0.010$). The difference in females did not reach significance (**Figure 6D**, $p = 0.084$). Furthermore, we used additional information from the TCGA to verify the above result (**Figures 6G–J**).

The Construction of a Nomogram and Calibration Curve

The nomogram was constructed by integrating the risk score with other clinicopathological features (**Figures 7A,B**). Furthermore, the calibration curve displayed linear concordance in the predicted and actual survival rates at 1, 2, and 3 years (**Figures 7C,D**). The findings suggested that the nomogram had high accuracy in predicting OS.

Gene Set Enrichment Analysis

Gene set enrichment analysis (GSEA) was done using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). GO term analysis was used to evaluate the functional assessment of the different risk score groups, and the results demonstrated that the high-risk group was reportedly associated with protein folding, protein targeting to mitochondrion, endoplasmic reticulum protein containing complex, vacuolar membrane, catalytic activity acting on a tRNA and chaperone binding (**Figures 8A–C**). Additionally, monocarboxylic acid catabolic process, amino acid betaine metabolic process, microbody lumen, high density lipoprotein

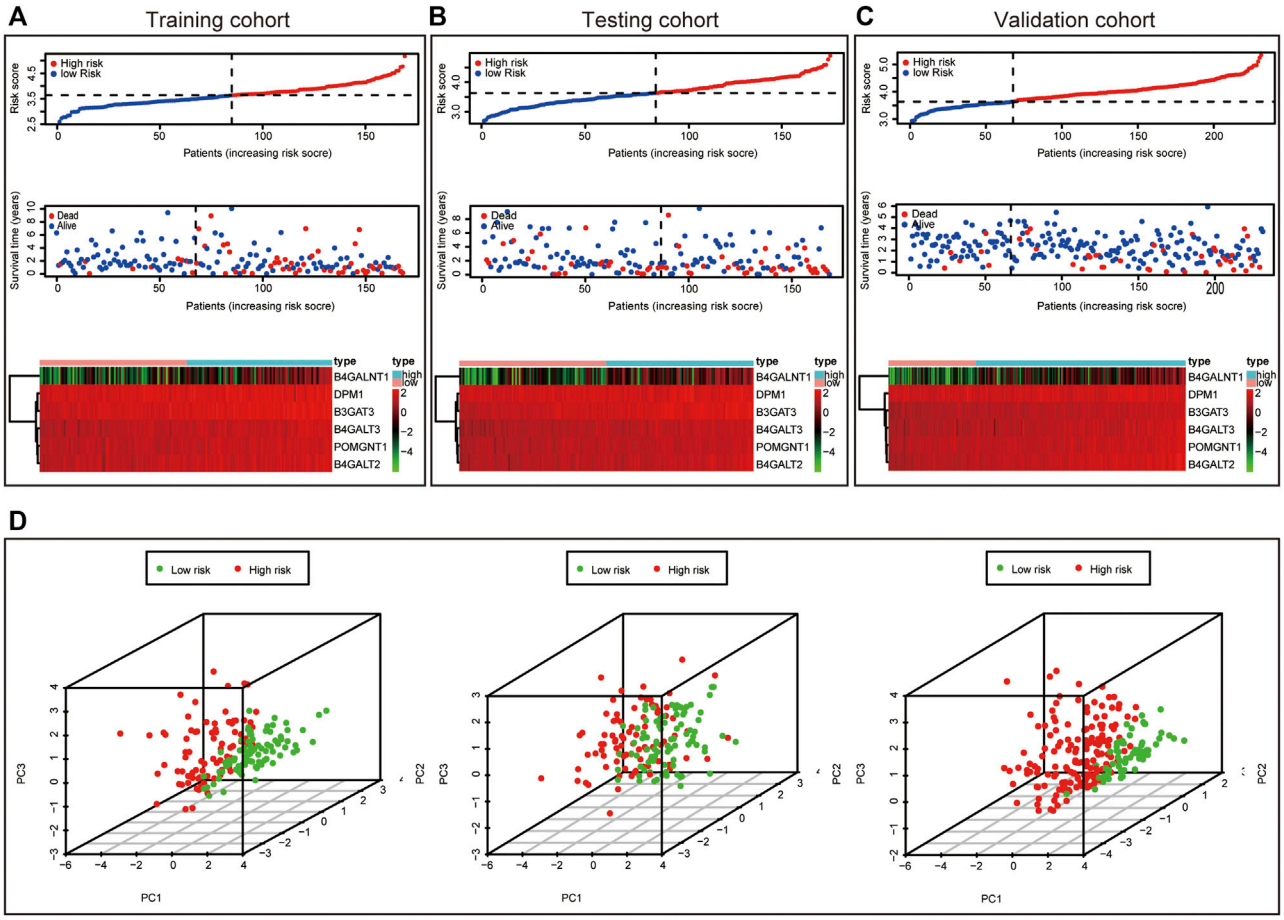


FIGURE 4 | Characteristics of prognostic signature. Distribution of risk score, Survival status of HCC samples and Heat map of the expression of prognostic signature in the training cohort **(A)**, testing cohort **(B)** and validation cohort **(C)**. Principal component analysis (PCA) plot in the training cohort, testing cohort and validation cohort **(D)**.

particle, aromatase activity and steroid hydroxylase activity were significantly downregulated in the low-risk group (**Figures 8D–F**). KEGG analysis showed that pyrimidine metabolism, purine metabolism, and N-glycan biosynthesis pathways were enriched in the high-risk group (**Figure 8G**); in contrast, some pathways in the low-risk group were enriched, including drug metabolism cytochrome P450, retinol metabolism and tryptophan metabolism (**Figure 8H**). We hypothesized that the prognostic signature may be potentially associated with metabolic disorders.

Correlation of Risk Score With Tumour-Infiltrating Immune Cells

The relative abundance of 22 infiltrating immune cells was calculated by the CIBERSORT algorithm between the groups in both datasets. In the ICGC dataset, the infiltration levels of follicular helper T cells, regulatory T cells (Tregs) and M0 macrophages were higher in the high-risk group; however, naive B cells and gamma delta T cells were significantly enriched in the low-risk group; meanwhile, the correlation between immune cell infiltration and risk score was analysed (**Figures 9A,B**). Then, we observed higher levels of immune-infiltrating of M0 macrophages, regulatory T cells (Tregs), memory B cells, activated CD4 memory T cells, follicular helper T cells and resting dendritic cells in the TCGA's high-risk group. In contrast, increased levels of naive B cells, resting CD4 memory T cells, resting NK cells, monocytes, M2 macrophages and resting Mast cells were found in the low-risk group. Similarly, we analysed the correlation between the risk score and TICS in the TCGA dataset (**Figures 9D,E**). By taking the intersection of results, the two most relevant TICS were identified as M0 macrophages and naive B cells (**Figures 9C,F–G**).

Validating the Expression of Six Genes

To validate the differential expression at the mRNA level, we used qRT–PCR to compare the expression of the six genes in the HCC cell line (HepG2) and normal liver cells (LO2) (**Figure 10**). The mRNA levels of DPM1, B4GALT3, B4GALT2, B4GALNT1, and B3GAT3 were significantly higher in HepG2 cells than in LO2 cells. Subsequently, to validate the differential expression at the

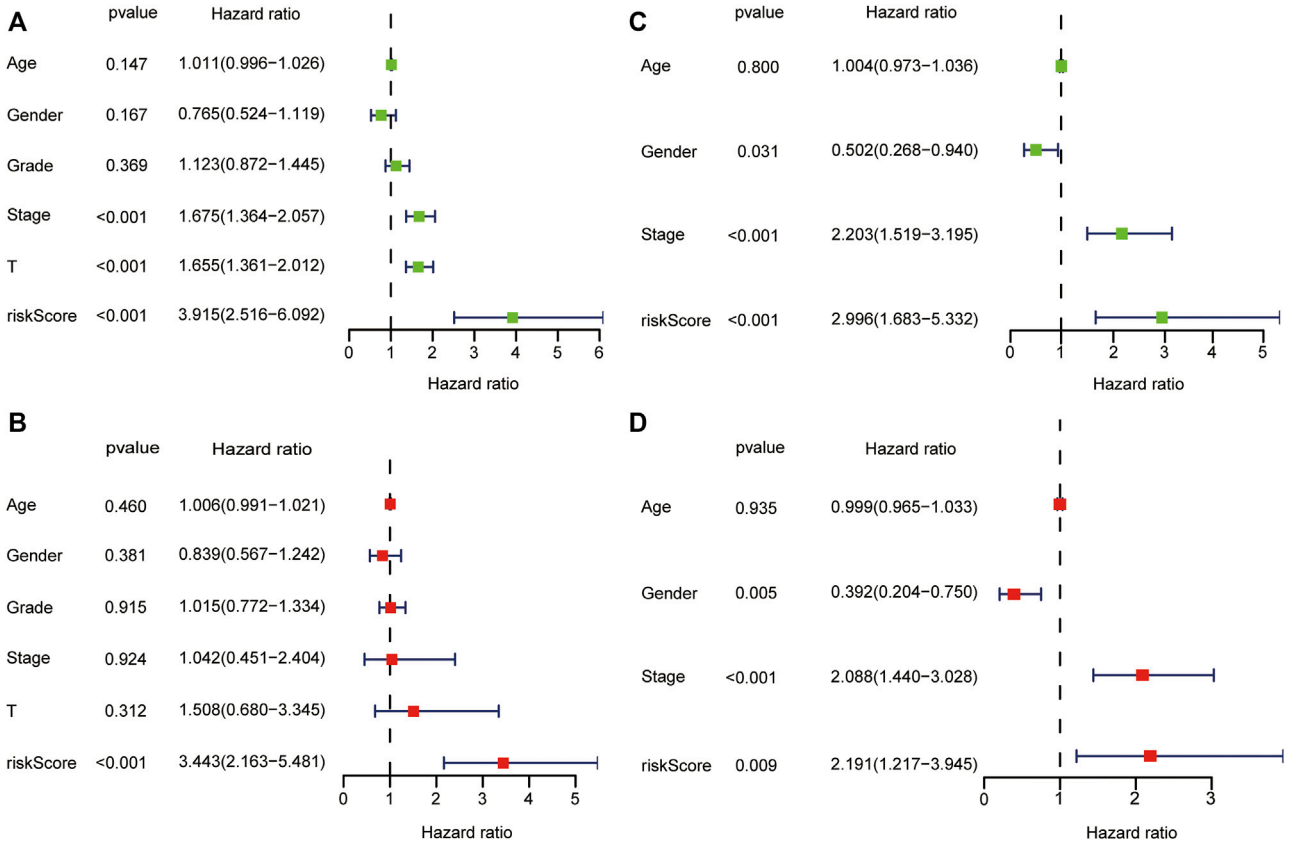


FIGURE 5 | Forest plot of prognostic signature and clinical risk factors. The univariate Cox regression analysis in the TCGA dataset **(A)** and ICGC dataset **(C)**. The multivariate Cox regression analysis in the TCGA dataset **(B)** and ICGC dataset **(D)**.

protein level, total cellular protein was analysed for the signature gene's expression by western blot (**Figure 11**). Likewise, these six genes were compared in normal versus cancer tissues derived from the HPA, and the results are shown in **Figure 12**. Similar expression of POMGNT1 and B3GAT3 was observed in normal versus cancer tissues by immunohistochemical staining. However, the degree of staining in DPM1, B4GALT3, B4GALT2, and B4GALNT1 was stronger in cancer tissues than normal tissues.

DISCUSSIONS

Protein glycosylation, as the most common post-translational modification, plays an indispensable regulatory role in diverse biological functions (Pinho and Reis, 2015). Almost all proteins exert their functions through one or more of the 14 distinct glycosylation pathways (Schjoldager, et al., 2020). Given its critical role in tumour biology, aberrant glycosylation is regarded as a new hallmark of cancer (Munkley and Elliott, 2016), and offers a novel direction to predict cancer prognosis and treat cancer. In essence, aberrant glycosylation is due to dysregulation of GTs, and many of them are implicated in tumorigenesis as tumour suppressors or oncogenes. For instance, as a major metabolic integration point, OGT is upregulated in many tumours, including HCC, and it has been shown to be involved in the regulation of stem-like cell potential through modification of eIF4E (Cao, et al., 2019). Likewise, several studies have reported associations between dysregulated GTs and patient outcomes. Wu et al. detected the expression level of the sialyltransferase ST3GAL1 in 273 patients with HCC and found that upregulation of ST3GAL1 was an independent predictor of OS and disease-free survival (DFS) (Wu, et al., 2016). Liu et al. demonstrated that the polypeptide N-acetylgalactosaminyltransferase GALNT4 promoted the development of cancer as a tumour suppressor gene, and the level of expression could act as an independent favourable prognostic factor for recurrence-free survival (RFS) and OS (Liu Y. et al., 2017). In addition, integrating multiple genes could better predict the clinical outcome in the study of Kuo et al. (Kuo, et al., 2017). For this purpose, it is necessary to explore the prognostic signature of GTs for the accurate prediction of prognosis or response to therapy, which provides a reliable basis and reference for cancer management.

In this study, 34 prognosis-related differentially expressed GTs were first obtained. Then, LASSO regression was applied to construct a prognostic signature, as used in previous studies(Ueno, et al., 2021). The final screening result identified

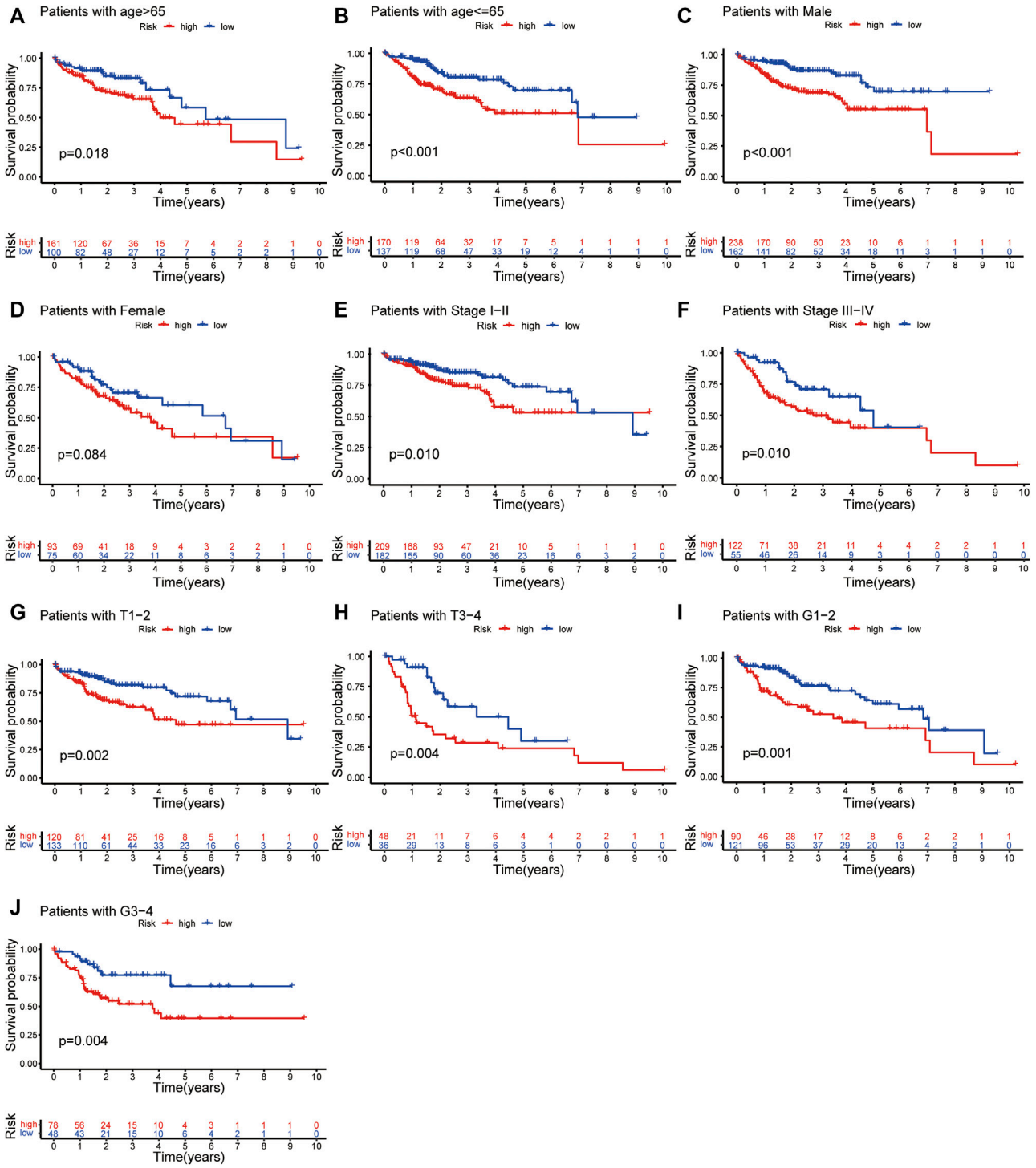


FIGURE 6 | Independent prognostic analysis of risk scores and clinicopathological features. The Kaplan–Meier survival plots of patients with age >65 and ≤65 **(A,B)**; Males and females **(C,D)**; Stage I-II and Stage III-IV **(E,F)** in both TCGA and ICGC dataset. The Kaplan–Meier survival plots of patients with tumour stage T1-2 and T3-4 **(G,H)**; tumour grading G1-2 and G3-4 **(I,J)**.

6 genes (POMGNT1, DPM1, B4GALT3, B4GALT2, B4GALNT1, and B3GAT3); consistent with the screening results, we confirmed their differential expression in cells and tissues. The prognostic signature had strong robustness and stable prediction performance in the training and validation cohorts by a series of significance tests, and the risk score was identified as an

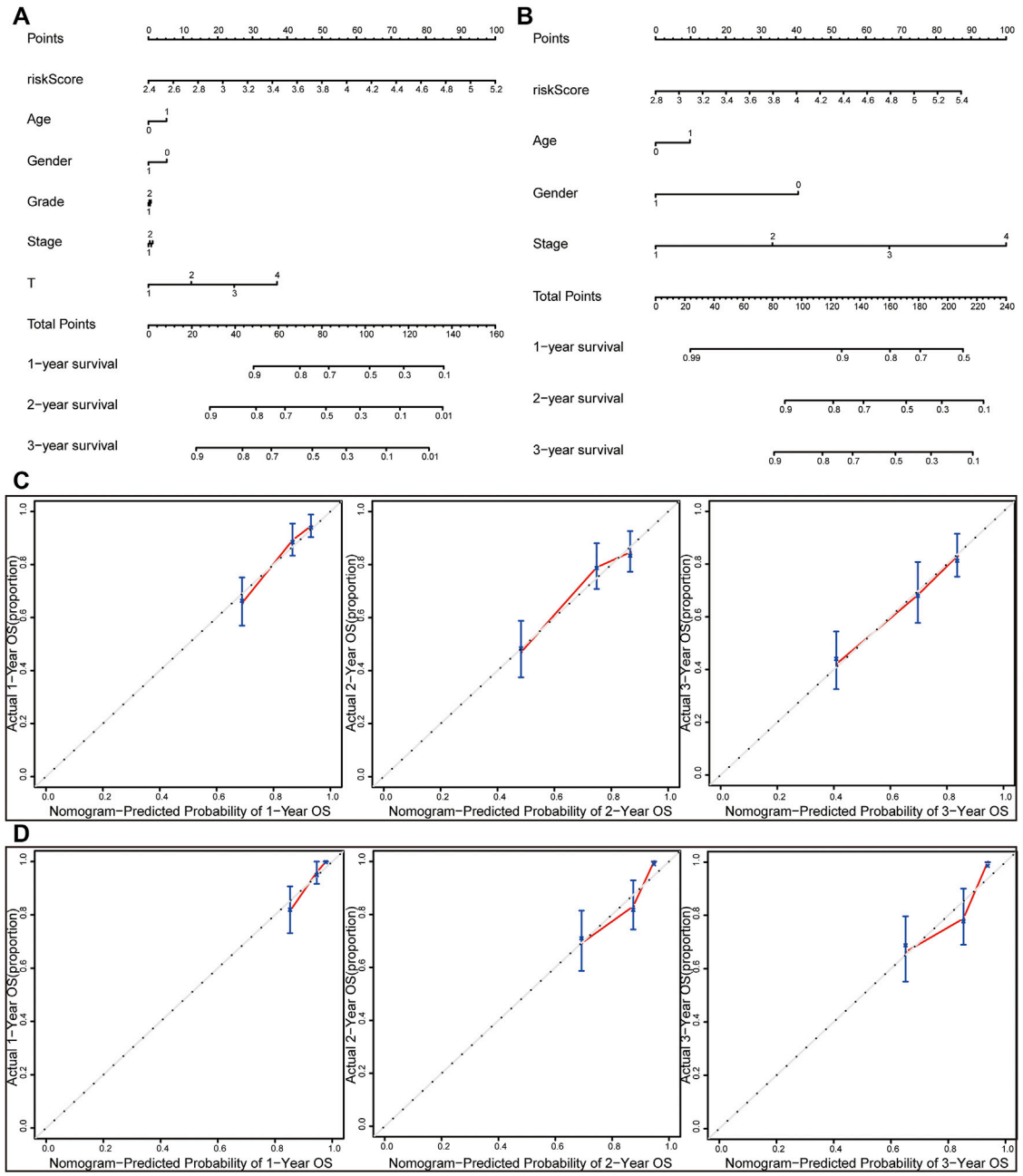


FIGURE 7 | Nomograms and calibration curves for the prognostic signature. Nomograms for predicting the OS of 1-, 2-, and 3-years in the TCGA dataset **(A)** and ICGC dataset **(B)**. Calibration curves of nomograms in the TCGA dataset **(C)** and ICGC dataset **(D)**.

independent prognostic indicator. Then, a nomogram comprising the risk score and clinicopathological data was generated to predict OS and showed superior performance in its validity. To gain more insight into the potential biological mechanism of the prognostic signature, we further used GSEA for the identification of biological functions. As expected, the results revealed that the prognostic signature was enriched in metabolism-related signalling pathways. Meanwhile, we know that glycosylation can modify protein structure and function; likewise, glycosylation affects immune cells with diverse functions and therefore modifies the tumour-immune microenvironment. M0 macrophages and naive B cells were identified as the most relevant TICS in our risk score group.

Coincidentally, the gene signature showed a tumour-promoting effect in our study. POMGNT1 has been examined in depth in glioblastoma (GBM), and it can promote proliferation

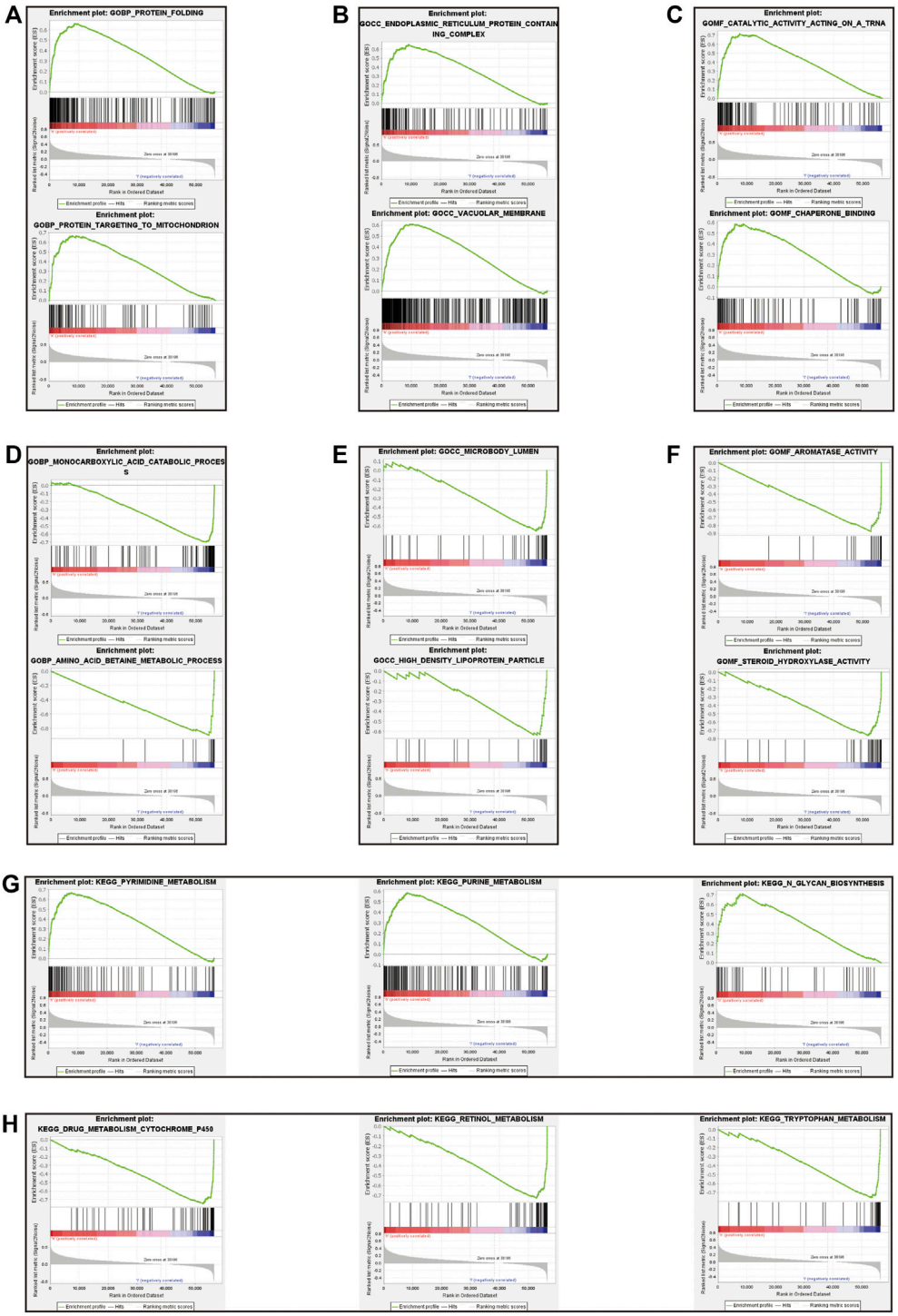


FIGURE 8 | Gene set enrichment analysis between high-risk and low-risk groups. The result of top 3 in GO analysis in the high-risk group **(A–C)**. The result of top 3 in GO analysis in the low-risk group **(D–F)**. The upregulated KEGG pathways of top 3 in the high-risk group **(G)**. The upregulated KEGG pathways of top 3 in the low-risk group **(H)**.

and invasion by regulating EGFR/ERK signalling (Lan, et al., 2015); Furthermore, it induces temozolomide resistance of tumour cells in GBM by regulating the expression of factors in EMT signalling (Liu, et al., 2017b). In addition, it acts as a prognostic and predictive novel marker in GBM, similar to our results (Lan, et al., 2013). DPM1 acts as a core catalytic

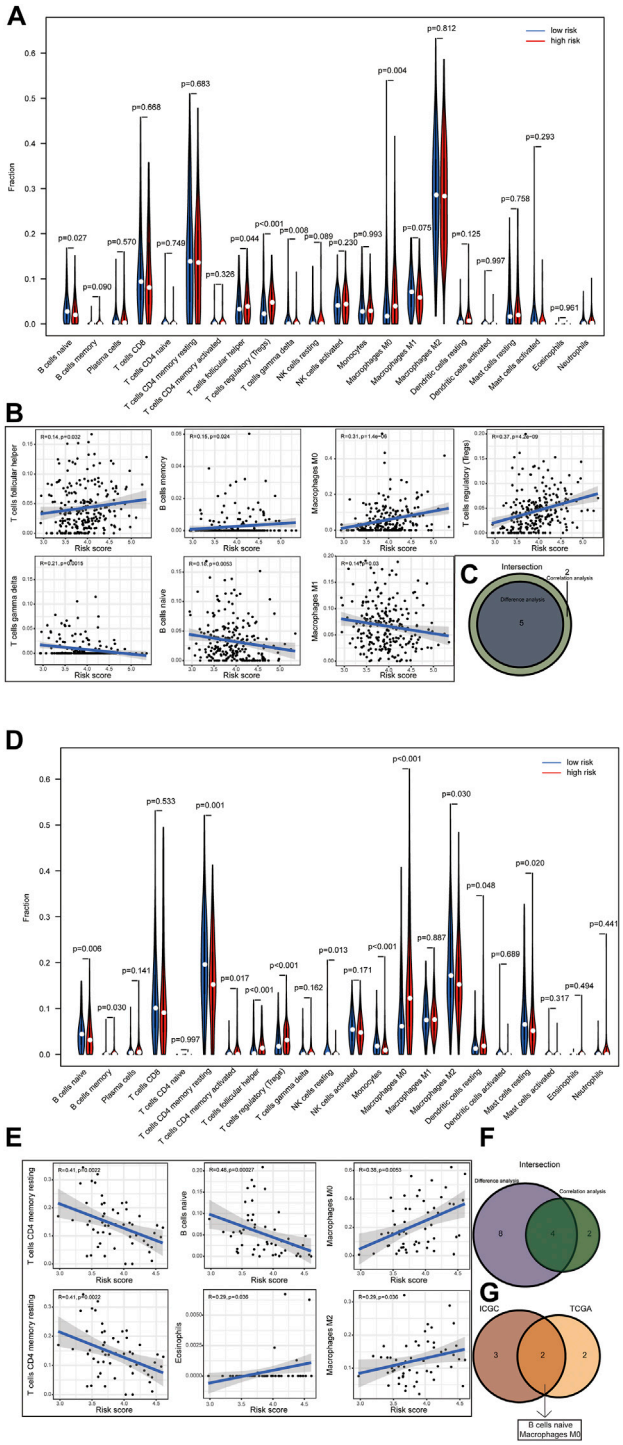


FIGURE 9 | (1–2) Correlation of risk score with tumor-infiltrating immune. Results of the infiltrating level of 22 immune cell types in the ICGC dataset **(A)** and TCGA dataset **(D)**. Correlations of risk scores with immune infiltration level in the ICGC dataset **(B)** and TCGA dataset **(E)** (only significant correlations were plotted). Venn diagram of immune cells by the results of difference analysis and correlation analysis in the ICGC dataset **(C)** and TCGA dataset **(F)**. Result of the overlapping immune cell in the ICGC dataset and TCGA dataset **(G)**.

component of Dolichol phosphate mannose synthase (DPMS) (Tomita, et al., 1998). Li et al. reported that DPM1 serves as a biomarker for HCC patients' prognostic prediction because the level of expression is significantly associated with clinicopathological parameters (Li, et al., 2020). The beta-1,4-1galactosyltransferase (B4GALT) family is a class of key enzymes that have crucial roles in many biological events, and catalyses the biosynthesis of N-acetyllactosamine on N-glycans by transferring UDP-galactose. It has been reported that upregulation of B4GALT2 induces p53-mediated apoptosis in HeLa cells and reveals a relationship with cisplatin-resistance in ovarian cancer cells (Zhou, et al., 2008; Zhao, et al., 2017). On the other hand, B4GALT3 has been evaluated more in-depth in tumour research than the former and it mainly plays a functional role by directly modifying β1-integrin glycosylation (Chen, et al., 2014; Chang, et al., 2013; Sun, et al., 2016). However, B3GALT3 develops different effects in different tumours, research shows that B4GALT3 overexpression can promote tumour growth and invasion in cervical cancer, neuroblastoma and GBM (Chang, et al., 2013; Sun, et al., 2016; Wu, et al., 2020), opposite to the tumour suppressor effects in colorectal and bladder cancer (Chen, et al., 2014; Liu, et al., 2018). In our study, we also found a high level of expression of this gene in HCC. B4GALNT1 (also known as GM2/GD2 synthase) functions as the key enzyme that transfers *N*-acetylgalactosamine (GalNAc) to GM3/GD3, catalysing the biosynthesis of gangliosides GM2/GD2 (Yoshida, et al., 2020). In breast cancer stem cells (CSCs), the upregulation of B4GALNT1 plays key roles in maintaining the CSC phenotype (Liang, et al., 2013). Jiang et al. confirmed that B4GALNT1 promoted the progression and metastasis of lung adenocarcinoma through the JNK/c-Jun/Slug signalling pathway and was involved in the tumour development of melanoma and clear cell renal cell carcinoma (Yoshida, et al., 2020; Jiang, et al., 2021; Yang, et al., 2019). B3GAT3 participates in the biosynthesis of the glycosaminoglycan (GAG) linker region of proteoglycan (PG) (Barré, et al., 2006). Given its important role in tumour metabolism, which was used repeatedly as a candidate gene for constructing prognostic models (Zhao, et al., 2021; Zhao, et al., 2020; Bingxiang, et al., 2021), its function of promoting the process of tumour EMT in HCC was confirmed by experimental verification (Zhang, et al., 2019). Based on the analysis above, the cancer-promoting effects of dysregulated expression are in accordance with our prediction results.

Risk prediction models have been developed as a powerful tool to provide references for clinical decision-making. A large amount of evidence has identified that dysregulation of GT expression plays critical roles in tumorigenesis, affecting the prognosis of HCC. For this reason, we developed a risk model and tried to explore its prognostic value. Although promising prediction results were displayed in our study, there is still room for improvement. First, we took a bioinformatics approach to mine GT data, which should be taken prudently and further validated by experimental studies before it is developed for clinical use. Second, to improve the efficiency of risk

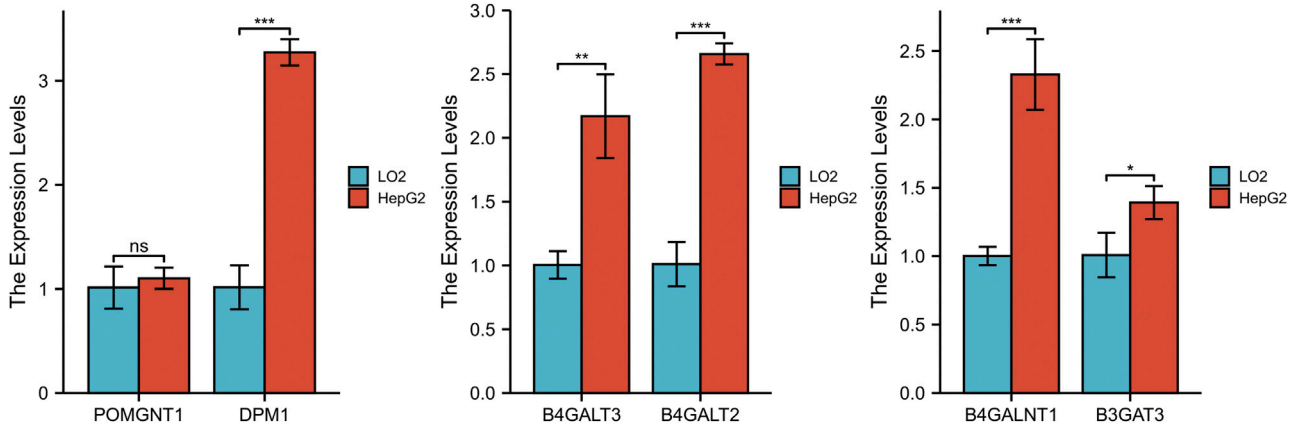


FIGURE 10 | Validation of the mRNA expression levels of the prognostic genes in HCC cell line (HepG2) and normal hepatocyte cell line (LO2) using qRT–PCR.

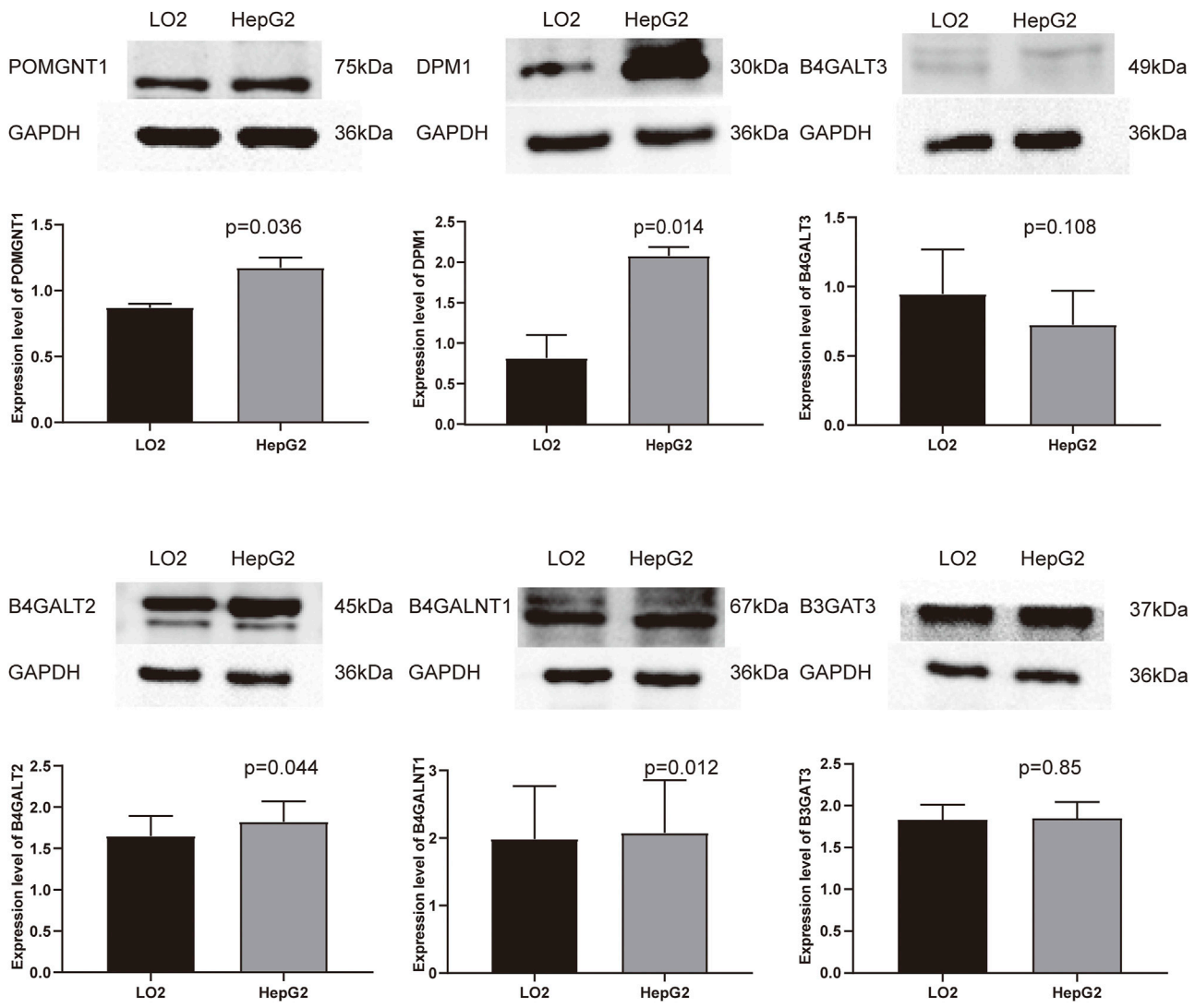


FIGURE 11 | Validation of the protein expression levels of the prognostic genes in HCC cell line (HepG2) and normal hepatocyte cell line (LO2) using Western blot.

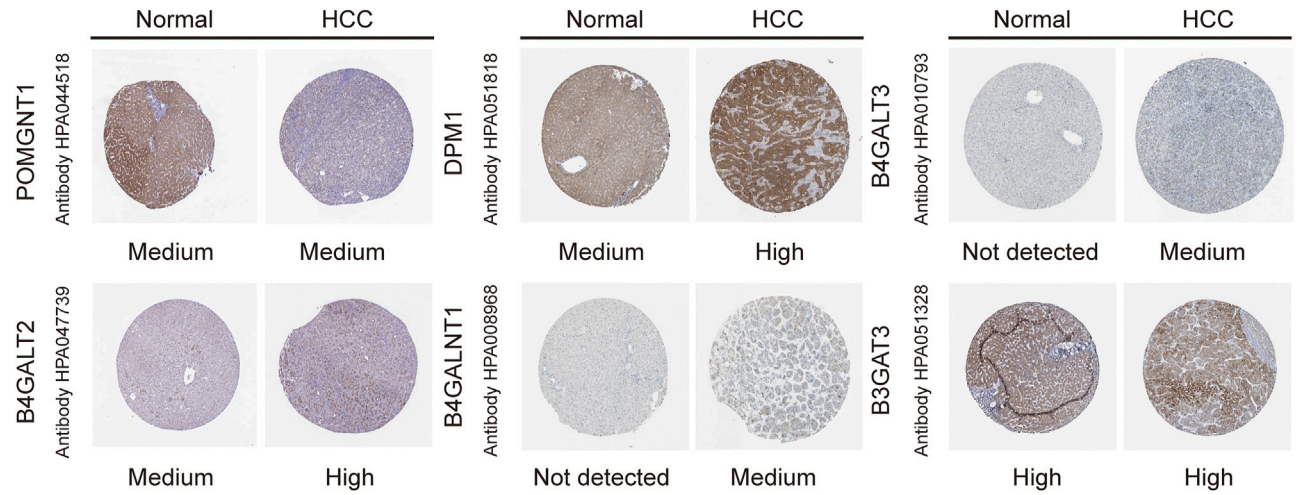


FIGURE 12 | Immunohistochemistry staining of the prognostic genes in HCC and normal liver tissues derived from the HPA database.

prediction in heterogeneity, further studies on larger sample sizes are needed.

CONCLUSION

In summary, a computational risk model combining six GTs was developed to aid in the clinical prediction of HCC prognosis. The model showed good prediction efficiency after verification by the internal testing and the external validation cohort. Furthermore, these prognostic markers were validated by western blot and qRT–PCR. However, further studies should be conducted to explore the clinical value of our current study.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

ZZ, GG, and JZ performed the study design. ZZ and TW extracted the data and analyzed the data. YD and JD performed the data validation. ZZ performed writing the original manuscript draft. GG and JZ contributed to the manuscript review and editing. All authors read and approved the final manuscript.

FUNDING

This research was funded by the Science and Technology Foundation of Jiangxi Education Department (grant number GJJ180027).

ACKNOWLEDGMENTS

We thank the peer reviewers for their input, which improved this manuscript. And we thank all researchers and staff who uploaded and maintained data in TCGA and ICGC databases for their great efforts. Meanwhile, we would like to express our gratitude to the team of director Shan of hepatobiliary Surgery department of the First Affiliated Hospital of Nanchang University for the source of the cells.

ACKNOWLEDGMENTS

All data generated or analyzed during this study are included in this published article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.823728/full#supplementary-material

REFERENCES

Barré, L., Venkatesan, N., Magdalou, J., Netter, P., Fournel-Gigleux, S., Ouzzine, M., et al. (2006). Evidence of Calcium-dependent Pathway in the Regulation of Human β1,3-glucuronosyltransferase-1 (GlcAT-I) Gene Expression: a Key Enzyme in Proteoglycan Synthesis. *FASEB j.* 20 (10), 1692–1694. doi:10.1096/fj.05-5073fje

Bingxiang, X., Panxing, W., Lu, F., Xiuyou, Y., and Chao, D. (2021). A Prognostic Model for Brain Glioma Patients Based on 9 Signature Glycolytic Genes. *Biomed. Res. Int.* 2021, 1–15. doi:10.1155/2021/6680066

Cao, B., Duan, M., Xing, Y., Liu, C., Yang, F., Li, Y., et al. (2019). O-GlcNAc Transferase Activates Stem-like Cell Potential in Hepatocarcinoma through O-GlcNAcylation of Eukaryotic Initiation Factor 4E. *J. Cel Mol Med* 23 (4), 2384–2398. doi:10.1111/jcmm.14043

Chang, H.-H., Chen, C.-H., Chou, C.-H., Liao, Y.-F., Huang, M.-J., Chen, Y.-H., et al. (2013). β-1,4-Galactosyltransferase III Enhances Invasive Phenotypes via β1-Integrin and Predicts Poor Prognosis in Neuroblastoma. *Clin. Cancer Res.* 19 (7), 1705–1716. doi:10.1158/1078-0432.Ccr-12-2367

Chen, C.-H., Wang, S.-H., Liu, C.-H., Wu, Y.-L., Wang, W.-J., Huang, J., et al. (2014). β-1,4-Galactosyltransferase III Suppresses β1 Integrin-Mediated Invasive Phenotypes and Negatively Correlates with Metastasis in Colorectal Cancer. *Carcinogenesis* 35 (6), 1258–1266. doi:10.1093/carcin/bgu007

Cheng, L., Gao, S., Song, X., Dong, W., Zhou, H., Zhao, L., et al. (2016). Comprehensive N-Glycan Profiles of Hepatocellular Carcinoma Reveal Association of Fucosylation with Tumor Progression and Regulation of FUT8 by microRNAs. *Oncotarget* 7 (38), 61199–61214. doi:10.18632/oncotarget.11284

Cheng, L., Luo, S., Jin, C., Ma, H., Zhou, H., and Jia, L. (2013). FUT Family Mediates the Multidrug Resistance of Human Hepatocellular Carcinoma via the PI3K/Akt Signaling Pathway. *Cell Death Dis* 4 (11), e923. doi:10.1038/cddis.2013.450

Dobie, C., and Skropeta, D. (2021). Insights into the Role of Sialylation in Cancer Progression and Metastasis. *Br. J. Cancer* 124 (1), 76–90. doi:10.1038/s41416-020-01126-7

Guo, Q., Guo, B., Wang, Y., Wu, J., Jiang, W., Zhao, S., et al. (2012). Functional Analysis of α1,3/4-fucosyltransferase VI in Human Hepatocellular Carcinoma Cells. *Biochem. Biophysical Res. Commun.* 417 (1), 311–317. doi:10.1016/j.bbrc.2011.11.106

Huang, D. Q., El-Serag, H. B., and Loomba, R. (2021a). Global Epidemiology of NAFLD-Related HCC: Trends, Predictions, Risk Factors and Prevention. *Nat. Rev. Gastroenterol. Hepatol.* 18 (4), 223–238. doi:10.1038/s41575-020-00381-6

Huang, H., Wu, Q., Guo, X., Huang, T., Xie, X., Wang, L., et al. (2021b). O-GlcNAcylation Promotes the Migratory Ability of Hepatocellular Carcinoma Cells via Regulating FOXA2 Stability and Transcriptional Activity. *J. Cel Physiol* 236 (11), 7491–7503. doi:10.1002/jcp.30385

Indellicato, R., and Trinchera, M. (2021). Epigenetic Regulation of Glycosylation. *Adv. Exp. Med. Biol.* 1325, 173–186. doi:10.1007/978-3-030-70115-4_8

Jemal, A., Ward, E. M., Johnson, C. J., Cronin, K. A., Ma, J., Ryerson, A. B., et al. (2017). Annual Report to the Nation on the Status of Cancer, 1975-2014, Featuring Survival. *J. Natl. Cancer Inst.* 109 (9), 109. doi:10.1093/jnci/djx030

Jiang, T., Wu, H., Lin, M., Yin, J., Tan, L., Ruan, Y., et al. (2021). B4GALNT1 Promotes Progression and Metastasis in Lung Adenocarcinoma through JNK/c-Jun/Slug Pathway. *Carcinogenesis* 42 (4), 621–630. doi:10.1093/carcin/bgaa141

Johannssen, T., and Lepenies, B. (2017). Glycan-Based Cell Targeting to Modulate Immune Responses. *Trends Biotechnol.* 35 (4), 334–346. doi:10.1016/j.tibtech.2016.10.002

Kuo, H.-H., Lin, R.-J., Hung, J.-T., Hsieh, C.-B., Hung, T.-H., Lo, F.-Y., et al. (2017). High Expression FUT1 and B3GALT5 Is an Independent Predictor of Postoperative Recurrence and Survival in Hepatocellular Carcinoma. *Sci. Rep.* 7 (1), 10750. doi:10.1038/s41598-017-11136-w

Lan, J., Guo, P., Chen, M., Wu, B., Mao, Q., and Qiu, Y. (2013). O-linked Mannose β-1,2-N-acetylglucosaminyltransferase 1 Correlated with the Malignancy in Glioma. *J. Craniofac. Surg.* 24 (4), 1441–1446. doi:10.1097/SCS.0b013e318295378b

Lan, J., Guo, P., Lin, Y., Mao, Q., Guo, L., Ge, J., et al. (2015). Role of Glycosyltransferase PomGnT1 in Glioblastoma Progression. *Neuro-Oncology* 17 (2), 211–222. doi:10.1093/neuonc/nou151

Li, M., Xia, S., and Shi, P. (2020). DPM1 Expression as a Potential Prognostic Tumor Marker in Hepatocellular Carcinoma. *PeerJ* 8, e10307. doi:10.7717/peerj.10307

Liang, Y.-J., Ding, Y., Levery, S. B., Lobaton, M., Handa, K., and Hakomori, S.-i. (2013). Differential Expression Profiles of Glycosphingolipids in Human Breast Cancer Stem Cells vs. Cancer Non-stem Cells. *Proc. Natl. Acad. Sci.* 110 (13), 4968–4973. doi:10.1073/pnas.1302825110

Liu, H., Chen, D., Bi, J., Han, J., Yang, M., Dong, W., et al. (2018). Circular RNA circUBXN7 Represses Cell Growth and Invasion by Sponging miR-1247-3p to Enhance B4GALT3 Expression in Bladder Cancer. *Aging* 10 (10), 2606–2623. doi:10.18632/aging.101573

Liu, Q., Xue, Y., Chen, Q., Chen, H., Zhang, X., Wang, L., et al. (2017b). PomGnT1 Enhances Temozolomide Resistance by Activating Epithelial-Mesenchymal Transition Signaling in Glioblastoma. *Oncol. Rep.* 38 (5), 2911–2918. doi:10.3892/or.2017.5964

Liu, Y., Liu, H., Yang, L., Wu, Q., Liu, W., Fu, Q., et al. (2017a). Loss of N-Acetylgalactosaminyltransferase-4 Orchestrates Oncogenic MicroRNA-9 in Hepatocellular Carcinoma. *J. Biol. Chem.* 292 (8), 3186–3200. doi:10.1074/jbc.M116.751685

Ma, L., Dong, P., Liu, L., Gao, Q., Duan, M., Zhang, S., et al. (2016). Overexpression of Protein O-Fucosyltransferase 1 Accelerates Hepatocellular Carcinoma Progression via the Notch Signaling Pathway. *Biochem. Biophysical Res. Commun.* 473 (2), 503–510. doi:10.1016/j.bbrc.2016.03.062

Makwana, V., Ryan, P., Patel, B., Dukie, S.-A., and Rudrawar, S. (2019). Essential Role of O-GlcNAcylation in Stabilization of Oncogenic Factors. *Biochim. Biophys. Acta (Bba) - Gen. Subjects* 1863 (8), 1302–1317. doi:10.1016/j.bbagen.2019.04.002

Munkley, J., and Elliott, D. J. (2016). Hallmarks of Glycosylation in Cancer. *Oncotarget* 7 (23), 35478–35489. doi:10.18632/oncotarget.8155

Noda, K., Miyoshi, E., Uozumi, N., Yanagidani, S., Ikeda, Y., Gao, C.-x., et al. (1998). Gene Expression of ?1-6 Fucosyltransferase in Human Hepatoma Tissues: A Possible Implication for Increased Fucosylation of ?-fetoprotein. *Hepatology* 28 (4), 944–952. doi:10.1002/hep.510280408

Pinho, S. S., and Reis, C. A. (2015). Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nat. Rev. Cancer* 15 (9), 540–555. doi:10.1038/nrc3982

Rodrigues, J. G., Duarte, H. O., Reis, C. A., and Gomes, J. (2021). Aberrant Protein Glycosylation in Cancer: Implications in Targeted Therapy. *Biochem. Soc. Trans.* 49 (2), 843–854. doi:10.1042/bst20200763

Schjoldager, K. T., Narimatsu, Y., Joshi, H. J., and Clausen, H. (2020). Global View of Human Protein Glycosylation Pathways and Functions. *Nat. Rev. Mol. Cel Biol* 21 (12), 729–749. doi:10.1038/s41580-020-00294-x

Singal, A. G., Lampertico, P., and Nahon, P. (2020). Epidemiology and Surveillance for Hepatocellular Carcinoma: New Trends. *J. Hepatol.* 72 (2), 250–261. doi:10.1016/j.jhep.2019.08.025

Sun, Y., Yang, X., Liu, M., and Tang, H. (2016). B4GALT3 Up-Regulation by miR-27a Contributes to the Oncogenic Activity in Human Cervical Cancer Cells. *Cancer Lett.* 375 (2), 284–292. doi:10.1016/j.canlet.2016.03.016

Thomas, D., Rathinavel, A. K., and Radhakrishnan, P. (2021). Altered Glycosylation in Cancer: A Promising Target for Biomarkers and Therapeutics. *Biochim. Biophys. Acta (Bba) - Rev. Cancer* 1875 (1), 188464. doi:10.1016/j.bbcan.2020.188464

Tomita, S., Inoue, N., Maeda, Y., Ohishi, K., Takeda, J., and Kinoshita, T. (1998). A Homologue of *Saccharomyces cerevisiae* Dpm1p Is Not Sufficient for Synthesis of Dolichol-Phosphate-Mannose in Mammalian Cells. *J. Biol. Chem.* 273 (15), 9249–9254. doi:10.1074/jbc.273.15.9249

Ueno, D., Kawabe, H., Yamasaki, S., Demura, T., and Kato, K. (2021). Feature Selection for RNA Cleavage Efficiency at Specific Sites Using the LASSO Regression Model in *Arabidopsis thaliana*. *BMC Bioinformatics* 22 (1), 380. doi:10.1186/s12859-021-04291-5

Wang, Q. Y., Guo, P., Duan, L. L., Shen, Z. H., and Chen, H. L. (2005). ?-1,3-Fucosyltransferase-VII Stimulates the Growth of Hepatocarcinoma Cells via the Cyclin-dependent Kinase Inhibitor p27Kip1. *Cmls, Cel. Mol. Life Sci.* 62 (2), 171–178. doi:10.1007/s00018-004-4349-8

Wu, C.-S., Lee, T.-Y., Chou, R.-H., Yen, C.-J., Huang, W.-C., Wu, C.-Y., et al. (2014a). Development of a Highly Sensitive Glycan Microarray for Quantifying AFP-L3 for Early Prediction of Hepatitis B Virus-Related Hepatocellular Carcinoma. *PLoS One* 9 (6), e99959. doi:10.1371/journal.pone.0099959

Wu, C.-S., Yen, C.-J., Chou, R.-H., Chen, J.-N., Huang, W.-C., Wu, C.-Y., et al. (2014b). Downregulation of microRNA-15b by Hepatitis B Virus X Enhances Hepatocellular Carcinoma Proliferationviafucosyltransferase 2-induced Globo H Expression. *Int. J. Cancer* 134 (7), 1638–1647. doi:10.1002/ijc.28501

Wu, H., Shi, X.-L., Zhang, H.-J., Hu, W.-D., Mei, G.-L., Chen, X., et al. (2016). Overexpression of ST3Gal-I Promotes Migration and Invasion of HCCLM3 *In Vitro* and Poor Prognosis in Human Hepatocellular Carcinoma. *Ott* 9, 2227–2236. doi:10.2147/ott.S96510

Wu, T., Li, Y., and Chen, B. (2020). B4GALT3 Promotes Cell Proliferation and Invasion in Glioblastoma. *Neurol. Res.* 42 (6), 463–470. doi:10.1080/01616412.2020.1740465

Yang, H., Li, W., Lv, Y., Fan, Q., Mao, X., Long, T., et al. (2019). Exploring the Mechanism of clear Cell Renal Cell Carcinoma Metastasis and Key Genes Based on Multi-Tool Joint Analysis. *Gene* 720, 144103. doi:10.1016/j.gene.2019.144103

Yoshida, H., Koodie, L., Jacobsen, K., Hanzawa, K., Miyamoto, Y., and Yamamoto, M. (2020). B4GALNT1 Induces Angiogenesis, anchorage independence Growth and Motility, and Promotes Tumorigenesis in Melanoma by Induction of Ganglioside GM2/GD2. *Sci. Rep.* 10 (1), 1199. doi:10.1038/s41598-019-57130-2

Zhang, S., Cao, X., Gao, Q., and Liu, Y. (2017). Protein Glycosylation in Viral Hepatitis-Related HCC: Characterization of Heterogeneity, Biological Roles, and Clinical Implications. *Cancer Lett.* 406, 64–70. doi:10.1016/j.canlet.2017.07.026

Zhang, Y.-l., Ding, C., and Sun, L. (2019). High Expression B3GAT3 Is Related with Poor Prognosis of Liver Cancer. *Open Med. (Wars)* 14, 251–258. doi:10.1515/med-2019-0020

Zhao, R., Qin, W., Qin, R., Han, J., Li, C., Wang, Y., et al. (2017). Lectin Array and Glycogene Expression Analyses of Ovarian Cancer Cell Line A2780 and its Cisplatin-Resistant Derivate Cell Line A2780-Cp. *Clin. Proteom* 14, 20. doi:10.1186/s12014-017-9155-z

Zhao, Y., Tao, Z., and Chen, X. (2020). A Three-Metabolic-Genes Risk Score Model Predicts Overall Survival in Clear Cell Renal Cell Carcinoma Patients. *Front. Oncol.* 10, 570281. doi:10.3389/fonc.2020.570281

Zhao, Y., Zhang, J., Wang, S., Jiang, Q., and Xu, K. (2021). Identification and Validation of a Nine-Gene Amino Acid Metabolism-Related Risk Signature in HCC. *Front. Cel Dev. Biol.* 9, 731790. doi:10.3389/fcell.2021.731790

Zhou, J., Wei, Y., Liu, D., Ge, X., Zhou, F., Yun, X., et al. (2008). Identification of 1,4GalT II as a Target Gene of P53-Mediated HeLa Cell Apoptosis. *J. Biochem.* 143 (4), 547–554. doi:10.1093/jb/mvn003

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 Genetics

MINI REVIEW
published: 31 March 2022
doi: 10.3389/fgene.2022.849040

Therapeutic Targeting Hypoxia-Inducible Factor (HIF-1) in Cancer: Cutting Gordian Knot of Cancer Cell Metabolism

Abhilasha Sharma[1†], *Sonam Sinha*[2] *and Neeta Shrivastava*[3*†]

[1]Department of Life Science, University School of Sciences, Gujarat University, Ahmedabad, India, [2]Kashiv Biosciences, Ahmedabad, India, [3]Shri B.V. Patel Education Trust, Ahmedabad, India

OPEN ACCESS

Edited by:
Jaspreet Kaur Dhanjal,
Indraprastha Institute of Information Technology Delhi, India

Reviewed by:
Alok Kumar,
Kyoto University, Japan
Shantanu Shukla,
Northwestern University, United States
Aftab Alam,
Roswell Park Comprehensive Cancer Center, United States

***Correspondence:**
Neeta Shrivastava
neetashrivastava@hotmail.com

†ORCID:
Abhilasha Sharma
0000-0002-2438-6556
Neeta Shrivastava
0000-0002-7600-6682

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 05 January 2022
Accepted: 09 March 2022
Published: 31 March 2022

Citation:
Sharma A, Sinha S and Shrivastava N (2022) Therapeutic Targeting Hypoxia-Inducible Factor (HIF-1) in Cancer: Cutting Gordian Knot of Cancer Cell Metabolism. Front. Genet. 13:849040. doi: 10.3389/fgene.2022.849040

Metabolic alterations are one of the hallmarks of cancer, which has recently gained great attention. Increased glucose absorption and lactate secretion in cancer cells are characterized by the Warburg effect, which is caused by the metabolic changes in the tumor tissue. Cancer cells switch from oxidative phosphorylation (OXPHOS) to aerobic glycolysis due to changes in glucose degradation mechanisms, a process known as "metabolic reprogramming". As a result, proteins involved in mediating the altered metabolic pathways identified in cancer cells pose novel therapeutic targets. Hypoxic tumor microenvironment (HTM) is anticipated to trigger and promote metabolic alterations, oncogene activation, epithelial-mesenchymal transition, and drug resistance, all of which are hallmarks of aggressive cancer behaviour. Angiogenesis, erythropoiesis, glycolysis regulation, glucose transport, acidosis regulators have all been orchestrated through the activation and stability of a transcription factor termed hypoxia-inducible factor-1 (HIF-1), hence altering crucial Warburg effect activities. Therefore, targeting HIF-1 as a cancer therapy seems like an extremely rational approach as it is directly involved in the shift of cancer tissue. In this mini-review, we present a brief overview of the function of HIF-1 in hypoxic glycolysis with a particular focus on novel therapeutic strategies currently available.

Keywords: genomic alterations, cancer, metabolism, warburg effect, hypoxia-induced tumor microenvironment, metabolic reprogramming, cancer therapies, clinical outcomes

INTRODUCTION

Increased incidence of cancer patients around the globe clearly alarms for more comprehensive research of this life-threatening problem. The initiation of cancer is a multi-step process that includes genomic alterations. Hannah and Weinberg have extensively described the "hallmarks of cancer", one of which is "metabolic reprogramming" that has recently emerged as a core trait of tumors (Hanahan and Weinberg, 2011; Hanahan, 2022). Specifically, the altered glycolytic metabolism pathway results in switching from oxidative phosphorylation (OXPHOS) in the mitochondria to aerobic glycolysis even in the abundance of oxygen in various cancer types. The "Warburg effect", proposed by Otto Warburg over a century ago, was the first to reveal basic metabolic distinctions between differentiated cells and rapidly proliferating tumor cells (Otto, 2016). Warburg effect is the result of the interplay between (normoxic/hypoxic) HIF-1 upregulation, activation of an oncogene (cMyc, Ras), loss of function of tumor suppressors (mutant-p53, mutant-PTEN, micro RNAs and sirtuins with suppressor functions), activation of (PI3K/Akt/mTOR; Ras/Raf/Mek/Erk/cMyc; Jak/

Stat3) or deactivation of (LKb1/AMPk) signalling pathways (Arora et al., 2015; Vaupel and Multhoff, 2021). Although Warburg's and others' findings have had a significant impact on our understanding of tumor biology, they constitute only one aspect of tumor metabolism.

In fact, cancer metabolism alterations span a wide range of metabolic pathways that serve a multitude of functions such as apoptosis, angiogenesis, anti-anoikis, and anchorage-independent expansion in cancer cells and in the tumor microenvironment (TME), in addition to glucose metabolism and energetics (Casero and Pegg, 2009; Platten et al., 2012; Zhang and Du, 2012; Jeon and Hay, 2018). Therefore, targeting the energy metabolism of cancer cells, which takes advantage of the metabolic differences between cancer cells and normal cells opens the doorway to novel therapeutic interventions.

The TME endures biochemical alterations during the growth of the solid tumor, including depletion of glucose, bicarbonate, and oxygen (i.e., hypoxia and anoxia), high amounts of lactate and adenosine, and low pH value (Wang et al., 1995; Ke and Costa, 2006). Hypoxia, a prevalent characteristic of cancer especially solid tumors, is hypothesised to enhance tumor invasiveness and metastasis (Ke and Costa, 2006). Tumor hypoxia has been attributed to a variety of factors. First, angiogenesis inability to keep up with cancer growth, such as the need for the cancer cell mass "outstripping" the ability of blood vessels to carry oxygenated blood. Second, ischemia-induced by arteriovenous shunting or microvessel 'steal' syndromes induced by abnormal vessel arborization and aberrant vascular connections inside malignancies. Lastly, elevated hydrostatic pressure within the tumor, results in compression of the microvasculature (Heldin et al., 2004). Several mechanisms, notably the hypoxia-inducible factor-1 (HIF-1) pathway, which promotes the elevated expression of glycolytic enzymes, can govern the metabolic transition state above at the transcriptional level. As a result, tumor hypoxia and HIFs influence the majority of cancer "hallmarks", including cellular proliferation, apoptosis, metabolism, immunological responses, genomic instability, vascularization, neovascularization, invasion, and metastasis (Wigerup et al., 2016). Moreover, HIFs seem to impact chemo and radiation resistance through multiple pathways. Additionally, HIFs expression has been linked to poor prognosis and treatment relapse in clinical tumor samples (Sørensen and Horsman, 2020). Thus, HIFs appear to be critical therapeutic targets that can be used to enhance current cancer treatment for metastatic and treatment-resistant cancers.

The primary intent of this mini-review is to provide a brief overview of the metabolic processes that are regulated by a hypoxia-inducible factor. In this review, we outline the relevance of HIFs in glycolysis, cancer progression and the epithelial-mesenchymal transition (EMT). A further goal of the review is to overview the currently available therapeutic strategies.

Relevance of HIF-1 Stimulated Glycolysis in Hypoxia

Hypoxia affects metabolic pathways in a variety of ways. For example, by blocking the oxygen-dependent process of mitochondrial OXPHOS, hypoxia reduces ATP synthesis, and thus makes O_2-independent glycolysis a more important energy source (Denko, 2008; Frezza and Gottlieb, 2009). Increased glycolysis generates ATP quickly, but at the price of a substantial amount of glucose, as seen by elevated lactic acid levels. Intra-tumoral acidosis is mediated by the latter, in conjugation with mitochondria's impaired capacity to use protons in ATP synthesis (Zhou et al., 2006). Surprisingly, rather than being anti-cancer, the stress placed on cancer cells appears to promote the formation of more aggressive subclones with a greater ability to penetrate tissues and metastasis (Gatenby and Gillies, 2004; Gatenby et al., 2007). Hypoxia-induced events are mostly determined by the activity of the transcriptional regulators' hypoxia-inducible factor-1α (HIF-1α) and its partner HIF-1β.

HIF-1, a transcription factor, regulates the activation of several genes involved in glucose uptake and metabolism, cell survival/proliferation, angiogenesis, invasion, and metastasis (Semenza et al., 1994; Carmeliet et al., 1998). It is a heterodimer of HIF-1α and a constitutively expressed subunit HIF-1β which also forms a dimer with HIF-2α and regulates gene activation (Wang et al., 1995; Carmeliet et al., 1998). HIF-1α is generally targeted for ubiquitin-mediated destruction by proline hydroxylation and association with the Von Hippel-Lindau (VHL) tumor suppressor complex under normoxic conditions, but it is stabilised when the partial pressure of oxygen is low **(Figure 1)**. Moreover, overexpression of HIF-1α is linked to a poor prognosis in various patients with human malignancies including breast, colon, gastric, lung, skin, ovarian, pancreatic, prostate, and renal cancer (Bos et al., 2001; Dales et al., 2005; Chen et al., 2007; Simiantonaki et al., 2008). Thus, HIF-1α significantly enhances our molecular understanding of cancer progression and metastasis which is discussed in detail in the following sections.

Hypoxic Tumor-Microenvironment: Leading to Cancer Progression and Epithelial-Mesenchymal Transition

Mammalian cancer cells within a Hypoxic tumor microenvironment (HTM) undergo tremendous alterations, eventually intensifying their malignant activity. As a result, emphasis has been laid on identifying processes involved in cancer cell adaptation to the HTM in order to identify targets for potential therapeutic treatments (Liu et al., 2011; Kogita et al., 2014; Yang et al., 2015). Basically, in hypoxia conditions, HIF-1α forms the HIF complex, which functions as a transcription factor in the activation of a wide range of genes, orchestrating major phenotypic alterations and eventually leading to EMT. Following EMT, cells lose their normal morphology and gain mesenchymal traits (Kalluri and Weinberg, 2009; Singh and Settleman, 2010), including the development of stemness (Sutherland, 1988), increased invasiveness, and metastasizing capacities (Vaupel, 2004). All of these alterations have been associated with poor prognosis and chemotherapy resistance in a variety of tumor types (Yang et al., 2008; Chou et al., 2012). EMT is characterised by the loss of cell adhesion protein (for instance E-cadherin) and

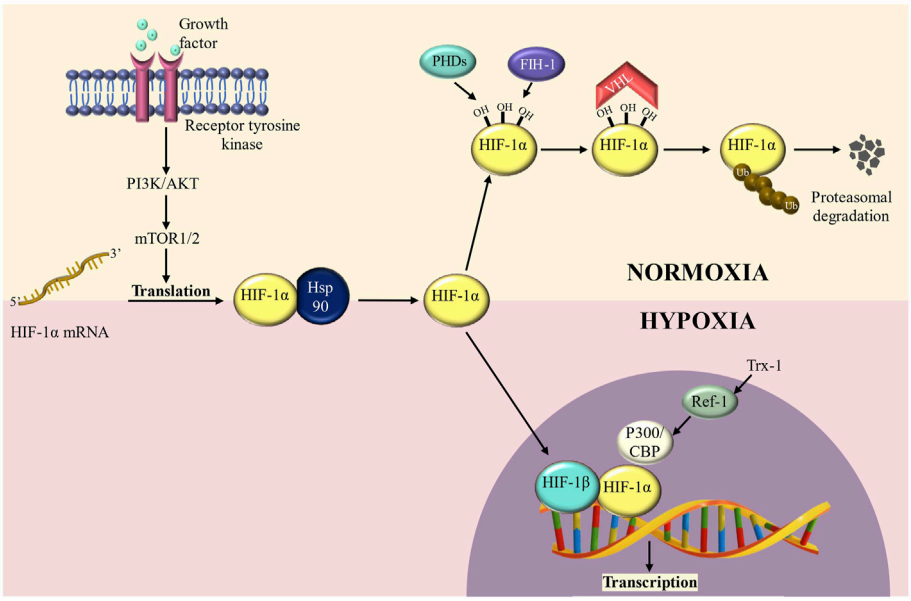


FIGURE 1 | HIF-1α regulation in normoxic and hypoxic conditions. HIF-1α is hydroxylated at conserved residue (Proline 564**)** under normoxic conditions, a process mediated by prolyl-4- hydroxylases (PHDs) and factor inhibiting HIF-1 (FIH-1) enzymes. PHD hydroxylation promotes HIF-1α protein destabilization, whereas FIH-1 hydroxylation inhibits transcriptional activity by preventing interaction with CBP/p300. HIF-degradation is mediated by a ubiquitin-dependent process carried out by the Von Hippel-Lindau (VHL) E3 ubiquitin ligase complex. Under hypoxic circumstances, inactivation of PHDs and FIH-1 causes HIF-stabilization, followed by translocation into the nucleus and dimerization with HIF-1/ARNT to create the HIF transcription factor. During hypoxia, HIFs, in collaboration with the coactivator CBP/p300, promote transcription of a wide range of target genes.

the elevated expression of mesenchymal-specific proteins such as SNAIL, Vimentin, and TWIST. As a matter of fact, this phenotypic shift has been highlighted as a major phase in the intricate process of developing distant metastasis (Chaffer and Weinberg, 2011; Valastyan and Weinberg, 2011).

As represented in **Figure 2**, the HIF-1α complex activates a number of key genes that mediate hypoxia > HIF > EMT axis. This axis has been extensively investigated in many aggressive tumors including lung, triple-negative breast cancer (TNBC), pancreatic ductal adenocarcinoma (PDAC) and renal cell carcinoma (RCC). For instance, autophagy markers (BECN1 and MAP1LC3) are activated in lung and PDAC (Zhu et al., 2014; Zou et al., 2014); overexpression of CAIX, the acidosis modulator has been reported in TNBC and RCC (Tan et al., 2009); further overexpression of epigenetic regulator (DNA methyltransferase, histone acetyltransferases, chromatin-remodelling enzymes, etc) and long-non coding RNA has been reported in gastric cancer, TNBC and PDAC (Krishnamachary et al., 2012; Onishi et al., 2012; Fujikuni et al., 2014; Liu et al., 2014; Wang et al., 2014); the chemokines are overexpressed in gastric cancer and multiple myeloma (Azab et al., 2012; Oh et al., 2012; Tao et al., 2014). Similarly, overexpression of cyclosporin binding protein cyclophilin A (CYPA) in PDAC (Zhang et al., 2014), endothelin in melanoma (Spinella et al., 2014); fascin in PDAC (Zhao et al., 2014); MMPs in PDAC, lung and ovarian cancer cell lines (Quintero-Fabián et al., 2019); protein kinase receptors in gastric, RCC, melanoma cancer (Chuang et al., 2008; Marconi et al., 2013) has been reported. HIF-1α also activates another critical cell signaling pathway i.e., HGF/MET signaling. Several studies suggest that MET, together with its ligand HGF, promotes cancer cell hallmarks including cell proliferation, survival, migration, angiogenesis in multiple mammalian cancer including hepatocellular carcinoma, head and neck cancer etc., (Goyal et al., 2013; Huang et al., 2020; Raj et al., 2022).

Additionally, in a positive feedback mechanism, ILK (Integrin Linked kinase) is activated by HIF-1α and is responsible for elevated HIF-1α expression through the regulatory loop (Matsuoka et al., 2013). Furthermore, E-cadherin, which was previously thought of as a tumor suppressor, was shown to have an unanticipated involvement in regulating genes involved in response to hypoxia and thus posing a potential role in metastatic breast cancer (Chu et al., 2013; Tam et al., 2020).

Moreover, intratumoral hypoxia alters the immune response of tumor in a variety of ways, all of which indicate an immunosuppressive impact (Palazón et al., 2012). HIF-1α, for example, can recruit myeloid-derived suppressor cells, regulatory T-cells, tumour-associated macrophages with immunosuppressive properties, as well as limit cytotoxic T-lymphocyte invasion (Corzo et al., 2010; Doedens et al., 2010; Imtiyaz et al., 2010; Barsoum et al., 2014). Besides that, HIF-1α stimulates the synthesis of the immunological checkpoint protein PD-L1(programmed death ligand-1), which aids in immune suppression and evasion (Noman et al., 2014; Abou Khouzam et al., 2021). As a result, the majority of the data implies that HIFs promote tumor growth through immunosuppression.

Collectively, these recent discoveries have motivated the scientific community to focus its efforts on developing novel drugs that can inhibit HIF-1α or its target genes. Further, we

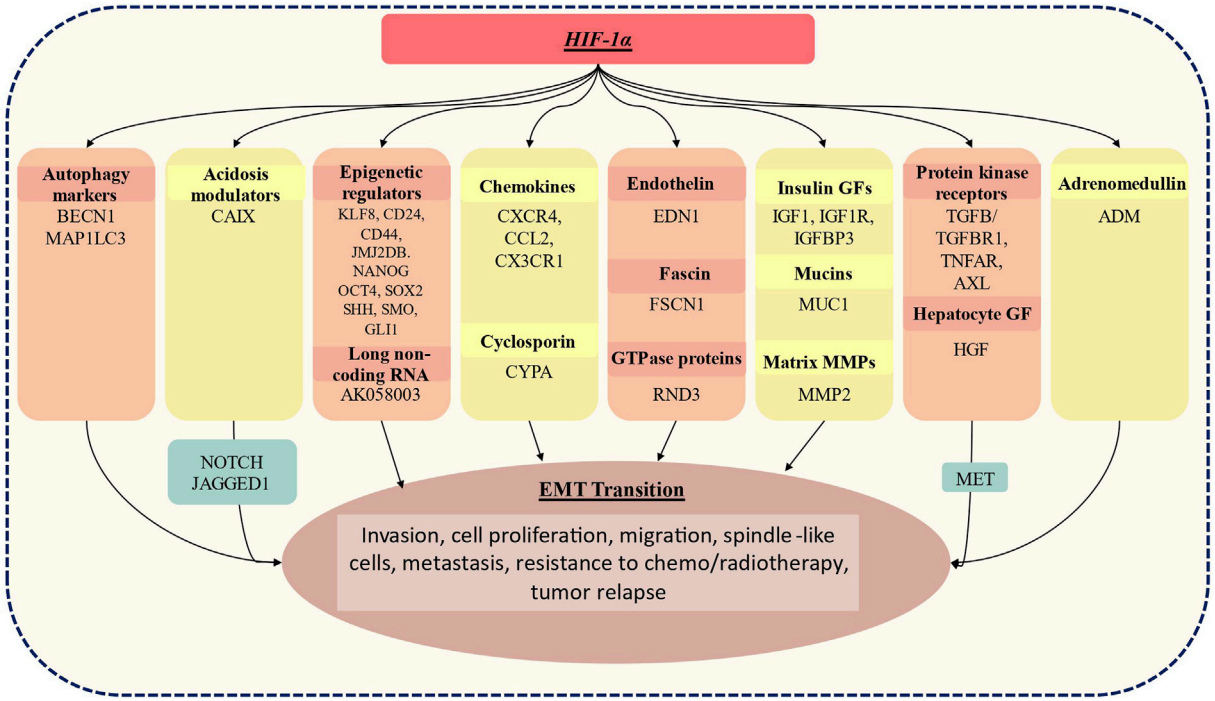


FIGURE 2 | Genes whose expression has been linked to the activation status of HIF-1α, resulting in EMT. HIF-1α induces expression of BECN1, MAP1IC3 which is an autophagy marker; CAIX, acidosis modulators: epigenetic regulators: KLF8, cell surface glycoproteins (CD24, CD44), JMJ2DB which is lysine-specific demethylase jumonji domain, Nanog homeobox (NANOG), Octamer-binding transcription factor 4(OCT4), SRY sex-determining region Y-box (SOX2), sonic hedgehog (SHH), smoothened frizzled class receptor (SMO), GLI family zinc finger 1 (GLI1); AK058003- long non-coding RNA; multiple chemokines: CXCR4, CCL2, CCR7, CX3CR1; cyclosporin bind protein cyclophilin A (CYPA); endothelins: EDN1 (endothelin1; fascins: fascin actin-bundling protein 1(FSCN1); GTPase proteins: Rho family GTPase 3 (RND3): insulin growth factor which includes IGF1, IGF1R, IGFBP3; mucin 1, cell surface-associated (MUC1); matrix metalloproteinase; MMP2, protein kinases receptors including TGFb/TGFBR1, TNFAR, AXL; hepatocyte growth factor (HGF) which is a ligand of MET tyrosine kinase receptor; adrenomedullin (ADM). These activated genes are known to play a crucial role in EMT transition and result in increased invasiveness, cellular proliferation, migration, spindle-like cellular appearance, resistance to chemo/radiotherapy and tumor relapse.

have focused on the compounds that have been developed as HIF-1α inhibitors and are now undergoing clinical trials. These novel compounds may pave the way for more effective therapy and might improve the prognosis of aggressive cancer patients.

Advanced Clinical Trials Targeting the Adaption to Hypoxia Tumor Microenvironment Therapeutic Targets

The ability to specifically target cancer cells while causing minimal harm to normal cells is one of the "Holy Grail" of cancer therapy. The propensity to exploit abnormalities between normal and malignant cells has significantly aided the discovery of novel anti-cancer drugs. Various small compounds discovered have been briefly summarized in the following section, albeit the bulk of them are still in the early stages of clinical trials.

As discussed above, HIF-1α activation has been found to have a significant impact on cancer cell metabolism as it influences the expression of several genes leading to increased glycolysis and impaired mitochondrial function in tumor cells. Several anticancer drugs that modulate the activity or levels of HIF-1α in cells influence HIF-1 without directly targeting it.

Digoxin (DIG) (PubChem CID: 2724385), a cardiac glycoside, has been demonstrated to have an anti-cancer effect *in vitro* and *in vivo* in various solid tumors by inhibiting HIF-1α production (Newman et al., 2008; Zhang et al., 2008; Lin et al., 2009). DIG is now being studied in phase 2 clinical trial (https://clinicaltrials.gov/ct2/show/NCT01763931) as a new HIF-1α inhibitor in breast cancer. This clinical trial will also be valuable in evaluating adverse events, as well as the safety and tolerability of DIG in pre-surgical breast cancer patients using the Common Terminology Criteria for Adverse Events, version 4. Additionally, Ganetespib (PubChem CID: 135564985), (5-[2,4-dihyroxy-5-(1-methylethyl)phenyl]4-(1-methyl-1H-indol-5-yl)-2,4-dihydro-3H-1,2,4-triazol-3-one) have been reported to increase the proteasome-mediated degradation of Hsp90. Hsp90, a chaperone, is implicated in tumor development, angiogenesis, and the generation of cancer stem cells (Pillai and Ramalingam, 2014; White et al., 2016). Its route triggers the activation of multiple oncogenic proteins including HIF-1α. Thus by targeting Hsp90, Ganetespib inhibits HIF-1α in TNBC mouse model (Xiang et al., 2014). Ganetespib is now being studies in a phase 3 trial in patients with advanced non-small cell lung cancer (NSCLC) in conjunction with docetaxel (https://www.clinicaltrials.gov/ct2/show/NCT01798485). This clinical trial

seeks to identify a potential synergism between ganetespib (150 mg/m^2) and docetaxel (75 mg/m^2) in order to suggest a more effective anti-cancer therapy than docetaxel alone.

Among multiple factors that influence hypoxia-induced tumor acidosis, CAIX is a hypoxia-inducible metal enzyme that promotes cancer cell survival/proliferation and invasion *via* HIF activation (Lock et al., 2013). It regulates cellular pH by catalyzing the reversible hydration of carbon dioxide to bicarbonate and protons. It is expressed exclusively on the cell surface of tumor cells, particularly CSCs (cancer stem cells), and is one of the key factors influencing cancer cell survival and metastasis (Lock et al., 2013). Moreover, CAIX is abundantly expressed in pancreatic ductal adenocarcinoma and breast cancer and has been implicated as a biomarker of poor prognosis for metastatic development and survival (Touisni et al., 2011; Lock et al., 2013). Additionally, research has proven a vital role for CAIX expression in the maintenance of the EMT phenotype, "stem cell" function, and hypoxia-induced tumor heterogeneity (Touisni et al., 2011; Ledaki et al., 2015). SLC-0111 (PubChem CID: 310360) is a small molecule that reaches the hypoxic niches and selectively binds and inhibits CAIX. Presently SLC-0111 is in phase I clinical trial (https://clinicaltrials.gov/ct2/show/NCT02215850) and the study focuses on its safety, tolerability, and pharmacokinetics, and efficacy in treating cancers. Similarly, another molecule DTP348 (PubChem CID: 57413968) namely 2-(2-methyl-5-nitro-1H-imidazol-1-yl) ethylsulfamide, is reported to target CIAX (Rami et al., 2013). Presently, this oral dual CAIX inhibitor/radiosensitizer is being researched in phase I clinical trial (https://clinicaltrials.gov/ct2/show/NCT02216669). This clinical study will consider the effects of DTP348 alone and in combination with radiation in patients with solid tumors to establish the appropriate phase II clinical trial dosage, safety, and tolerability.

Interestingly, HGF is the natural ligand of *MET,* a proto-oncogene. The HIF-1α induced HGF/MET pathway activation has been reported to induce EMT transition, resulting in a mesenchymal population that is more tumorigenic and chemoresistant than the preceding ones (Cañadas et al., 2014). Rilotumumab, Crizotinib/axitinib and cabozantinib are designed to effectively target HGF/MET pathway. Rilotumumab (PubChem SID: 135262715), is a human monoclonal antibody that is reported to significantly block the binding of HGF/SF to its MET receptor. Presently, it is being tested in phase 3 clinical trial (https://clinicaltrials.gov/ct2/show/NCT01697072) to evaluate if the treatment with epirubicin, cisplatin, and capecitabine in combination with rilotumumab results in better clinical outcomes in metastatic *MET* positive gastric cancers. Axitinib (PubChem CID: 6450551), with crizotinib (PubChem CID: 11626560), is currently being tested in a phase 1b clinical trial in patients with advanced solid tumors (https://clinicaltrials.gov/ct2/show/NCT01999972) (Kwak et al., 2010; Chen Y. et al., 2015). Moreover, cabozantinib is an oral inhibitor of *MET*, *RET, ROS1, NTRK1,* and AXL. It has been found to shrink tumor cells and significantly reduce cellular proliferation in medullary thyroid and prostate cancer. Cabozantinib (PubChem CID: 46830297), is currently being investigated to determine objective response rate (ORR), overall survival (OS) and progression-free survival (PFS) in advanced non-small cell lung cancer with *RET* fusions and those with *ROS1* or *NTRK1* fusions or elevated *MET* or *AXL* activity (https://clinicaltrials.gov/ct2/show/NCT01639508).

According to the current research, several phytocompounds also have been shown to play a significant role in cancer therapy and have numerous potential targets in tumorigenesis, including HIF-1 (Deng et al., 2019). Baicalein (PubChem CID: 5281605), (5,6,7- trihydroxyflavone), a flavonoid derived from *Scutellaria baicalensis* has been reported to have potent cytotoxic activity against a wide range of cancer (Bie et al., 2017; Dou et al., 2018; Wang et al., 2019). Surprisingly, baicalein when administered leads to the inhibition of hypoxia-induced Akt phosphorylation as a result of increased PTEN accumulation and decreased HIF-1α expression. Thus baicalein is a potential therapeutic sensitiser against gastric cancer since it inhibits glycolysis *via* PTEN/Akt/HIF-1α (Chen F. et al., 2015). Other investigations have corroborated the inhibitory effects of phytochemicals on HIF-1 in control of glucose metabolism. For instance, methylalpinumisoflavone (MF) (PubChem CID: 15596285), a flavonoid isolated from *Lanchocarpus glabrescens,* demonstrates a strong anti-cancer effect on T47D cells by suppressing HIF-1 and targets genes including CDKN1A, VEGF, and GLUT-1 in T47D cells (Liu et al., 2009). Moreover, oroxylin A (PubChem CID: 5320315) treatment has been linked to a reduction in cancer-related glycolysis *via* sirtuin-3 mediated destabilization of HIF-1 in MDA-MB-231 cells (Wei et al., 2015). Furthermore, EGCG (PubChem CID: 65064) is known to decrease the HIF-1α and glycolysis-related enzymes in T47D cells (Wei et al., 2018). Additionally, resveratrol (PubChem CID: 445154) has been shown to reduce the cellular uptake of glucose and induce glycolysis in cancer cell lines. Resveratrol inhibited intracellular reactive oxidative species (ROS) and hence lowered HIF-1 accumulation, decreased GLUT-1 expression, and induced glycolytic flow, according to measurements of cellular absorption of the glucose analogue 18F-fluorodeoxyglucose following resveratrol exposure (Jung et al., 2013).

Further using a combination of anti-cancer therapies is more likely to be successful than using a single drug (Maschek et al., 2004). Another concept has been proposed that takes the use of underlying metabolic variations between malignancies and healthy tissues (Payne, 2007). For instance, many tumors' reliance on glycolysis has been addressed using a variety of glycolytic pathway enzyme inhibitors that are also being evaluated as possible treatment drugs (Maher et al., 2004; Maschek et al., 2004; Xu et al., 2005; Pelicano et al., 2006; Gogvadze et al., 2009; Marín-Hernández et al., 2009; Mathupala et al., 2009). The major targets thus far have been glucose absorption (mediated mostly by GLUT-1), glucose retention (mediated by hexokinase) and lactate generation (catalyzed by lactate dehydrogenase-A). Unfortunately, inhibiting glycolysis has a significant complication; unlike organs that may easily utilise carbon sources other than glucose, the brain, retina, and testes are extremely glucose dependent. As a result, different metabolic targets such as specific glycolytic pathway enzyme isoforms which are transcriptionally overexpressed in response to HIF-1 elevations must be taken into account (Marín-Hernández et al., 2009).

Targeting proteins such as GLUTs, HK1, HKII, PFK-L, ALD-A, ALD-C, PGK1, ENO-α, PYK-M2, LDH-A, PFKFB-3 along with HIF-1α may be more trackable for drug development than HIF-1α itself. Identifying metabolic alterations that are specific to malignancies is inevitably a critical research goal.

CONCLUSION

Metabolic reprogramming is a frequent cancer cell mechanism for dealing with elevated energy demands. The growing interest in cancer metabolism has already resulted in a slew of novel potential therapeutics. In conclusion, several reports have shown that hypoxic cells may adapt to low oxygen levels by changing transcriptional and translational responses to increase glucose absorption and anaerobic catabolism. Since HIF-1 has been proven to be a master regulator of a wide range of proteins and enzymes involved in glucose metabolism and the glycolytic pathway. Thus modulation of the HIF-1 pathway is a promising therapeutic strategy. It is envisaged that a deeper insight into the molecular mechanisms involved in HIF-1 regulation and the Warburg effect in carcinogenesis would unlock new therapeutic interventions. Nonetheless, due to the present generation of agents' limited selectivity and specificity, there are possible challenges and concerns. Additionally, the recent metabolism-based therapeutics have shown some harmful effects on normal cells. Therefore, we propose combining the drugs to target distinct elements of cancer bioenergetics and hypoxia-induced factors in order to develop synergistic cancer treatments. Furthermore, directing these molecules to their targets would limit off-target effects while increasing efficacy.

AUTHOR CONTRIBUTIONS

AS: Conceptualization, Figures, Writing- original draft. SS: Writing- Revise and Editing. NS: Writing- Revise and Editing, Supervision. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

AS express gratitude to the Department of Science & Technology (DST), Ministry of Science and Technology, Government of India for INSPIRE fellowship (Grant no. IF190211).

REFERENCES

Abou Khouzam, R., Brodaczewska, K., Filipiak, A., Zeinelabdin, N. A., Buart, S., Szczylik, C., et al. (2021). Tumor Hypoxia Regulates Immune Escape/Invasion: Influence on Angiogenesis and Potential Impact of Hypoxic Biomarkers on Cancer Therapies. *Front. Immunol.* 11. doi:10.3389/fimmu.2020.613114

Arora, A., Singh, S., Bhatt, A. N., Pandey, S., Sandhir, R., and Dwarakanath, B. S. (2015). Interplay between Metabolism and Oncogenic Process: Role of microRNAs. *Translational oncogenomics* 7, 11–27. doi:10.4137/TOG.S29652

Azab, A. K., Hu, J., Quang, P., Azab, F., Pitsillides, C., Awwad, R., et al. (2012). Hypoxia Promotes Dissemination of Multiple Myeloma through Acquisition of Epithelial to Mesenchymal Transition-like Features. *Blood* 119, 5782–5794. doi:10.1182/BLOOD-2011-09-380410

Barsoum, I. B., Smallwood, C. A., Siemens, D. R., and Graham, C. H. (2014). A Mechanism of Hypoxia-Mediated Escape from Adaptive Immunity in Cancer Cells. *Cancer Res.* 74, 665–674. doi:10.1158/0008-5472.CAN-13-0992

Bie, B., Sun, J., Guo, Y., Li, J., Jiang, W., Yang, J., et al. (2017). Baicalein: A Review of its Anti-cancer Effects and Mechanisms in Hepatocellular Carcinoma. *Biomed. Pharmacother.* 93, 1285–1291. doi:10.1016/j.biopha.2017.07.068

Bos, R., Zhong, H., Hanrahan, C. F., Mommers, E. C. M., Semenza, G. L., Pinedo, H. M., et al. (2001). Levels of Hypoxia-Inducible Factor-1 during Breast Carcinogenesis. *JNCI J. Natl. Cancer Inst.* 93, 309–314. doi:10.1093/jnci/93.4.309

Cañadas, I., Rojo, F., Taus, Á., Arpí, O., Arumí-Uría, M., Pijuan, L., et al. (2014). Targeting Epithelial-To-Mesenchymal Transition with Met Inhibitors Reverts Chemoresistance in Small Cell Lung Cancer. *Clin. Cancer Res.* 20, 938–950. doi:10.1158/1078-0432.CCR-13-1330

Carmeliet, P., Dor, Y., Herbert, J.-M., Fukumura, D., Brusselmans, K., Dewerchin, M., et al. (1998). Role of HIF-1α in Hypoxia-Mediated Apoptosis, Cell Proliferation and Tumour Angiogenesis. *Nature* 394, 485–490. doi:10.1038/28867

Casero, R. A., and Pegg, A. E. (2009). Polyamine Catabolism and Disease. *Biochem. J.* 421, 323–338. doi:10.1042/BJ20090598

Chaffer, C. L., and Weinberg, R. A. (2011). A Perspective on Cancer Cell Metastasis. *Science* 331, 1559–1564. doi:10.1126/SCIENCE.1203543

Chen, F., Zhuang, M., Zhong, C., Peng, J., Wang, X., Li, J., et al. (2015a). Baicalein Reverses Hypoxia-Induced 5-FU Resistance in Gastric Cancer AGS Cells through Suppression of Glycolysis and the PTEN/Akt/HIF-1α Signaling Pathway. *Oncol. Rep.* 33, 457–463. doi:10.3892/or.2014.3550

Chen, H. H. W., Su, W.-C., Lin, P.-W., Guo, H.-R., and Lee, W.-Y. (2007). Hypoxia-inducible Factor-1α Correlates with MET and Metastasis in Node-Negative Breast Cancer. *Breast Cancer Res. Treat.* 103, 167–175. doi:10.1007/s10549-006-9360-3

Chen, Y., Suzuki, A., Tortorici, M. A., Garrett, M., LaBadie, R. R., Umeyama, Y., et al. (2015b). Axitinib Plasma Pharmacokinetics and Ethnic Differences. *Invest. New Drugs* 33, 521–532. doi:10.1007/s10637-015-0214-x

Chou, C.-W., Wang, C.-C., Wu, C.-P., Lin, Y.-J., Lee, Y.-C., Cheng, Y.-W., et al. (2012). Tumor Cycling Hypoxia Induces Chemoresistance in Glioblastoma Multiforme by Upregulating the Expression and Function of ABCB1. *Neuro-Oncology* 14, 1227–1238. doi:10.1093/NEUONC/NOS195

Chu, K., Boley, K. M., Moraes, R., Barsky, S. H., and Robertson, F. M. (2013). The Paradox of E-Cadherin: Role in Response to Hypoxia in the Tumor Microenvironment and Regulation of Energy Metabolism. *Oncotarget* 4, 446–462. doi:10.18632/ONCOTARGET.872

Chuang, M.-J., Sun, K.-H., Tang, S.-J., Deng, M.-W., Wu, Y.-H., Sung, J.-S., et al. (2008). Tumor-derived Tumor Necrosis Factor-Alpha Promotes Progression and Epithelial-Mesenchymal Transition in Renal Cell Carcinoma Cells. *Cancer Sci.* 99, 905–913. doi:10.1111/J.1349-7006.2008.00756.X

Corzo, C. A., Condamine, T., Lu, L., Cotter, M. J., Youn, J.-I., Cheng, P., et al. (2010). HIF-1α Regulates Function and Differentiation of Myeloid-Derived Suppressor Cells in the Tumor Microenvironment. *J. Exp. Med.* 207, 2439–2453. doi:10.1084/jem.20100587

Dales, J.-P., Garcia, S., Meunier-Carpentier, S., Andrac-Meyer, L., Haddad, O., Lavaut, M.-N., et al. (2005). Overexpression of Hypoxia-Inducible Factor HIF-1α Predicts Early Relapse in Breast Cancer: Retrospective Study in a Series of 745 Patients. *Int. J. Cancer* 116, 734–739. doi:10.1002/ijc.20984

Deng, X., Peng, Y., Zhao, J., Lei, X., Zheng, X., Xie, Z., et al. (2020). Anticancer Activity of Natural Flavonoids: Inhibition of HIF-1α Signaling Pathway. *Coc* 23, 2945–2959. doi:10.2174/1385272823666191203122030

Denko, N. C. (2008). Hypoxia, HIF1 and Glucose Metabolism in the Solid Tumour. *Nat. Rev. Cancer* 8, 705–713. doi:10.1038/nrc2468

Doedens, A. L., Stockmann, C., Rubinstein, M. P., Liao, D., Zhang, N., DeNardo, D. G., et al. (2010). Macrophage Expression of Hypoxia-Inducible Factor-1α Suppresses T-Cell Function and Promotes Tumor Progression. *Cancer Res.* 70, 7465–7475. doi:10.1158/0008-5472.CAN-10-1439

Dou, J., Wang, Z., Ma, L., Peng, B., Mao, K., Li, C., et al. (2018). Baicalein and Baicalin Inhibit colon Cancer Using Two Distinct Fashions of Apoptosis and Senescence. *Oncotarget* 9, 20089–20102. doi:10.18632/oncotarget.24015

Frezza, C., and Gottlieb, E. (2009). Mitochondria in Cancer: Not Just Innocent Bystanders. *Semin. Cancer Biol.* 19, 4–11. doi:10.1016/j.semcancer.2008.11.008

Fujikuni, N., Yamamoto, H., Tanabe, K., Naito, Y., Sakamoto, N., Tanaka, Y., et al. (2014). Hypoxia-mediated CD24 Expression Is Correlated with Gastric Cancer Aggressiveness by Promoting Cell Migration and Invasion. *Cancer Sci.* 105, 1411–1420. doi:10.1111/CAS.12522

Gatenby, R. A., and Gillies, R. J. (2004). Why Do Cancers Have High Aerobic Glycolysis? *Nat. Rev. Cancer* 4, 891–899. doi:10.1038/nrc1478

Gatenby, R. A., Smallbone, K., Maini, P. K., Rose, F., Averill, J., Nagle, R. B., et al. (2007). Cellular Adaptations to Hypoxia and Acidosis during Somatic Evolution of Breast Cancer. *Br. J. Cancer* 97, 646–653. doi:10.1038/sj.bjc.6603922

Gogvadze, V., Orrenius, S., and Zhivotovsky, B. (2009). Mitochondria as Targets for Cancer Chemotherapy. *Semin. Cancer Biol.* 19, 57–66. doi:10.1016/j.semcancer.2008.11.007

Goyal, L., Muzumdar, M. D., and Zhu, A. X. (2013). Targeting the HGF/c-MET Pathway in Hepatocellular Carcinoma. *Clin. Cancer Res.* 19, 2310–2318. doi:10.1158/1078-0432.CCR-12-2791

Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 12, 31–46. doi:10.1158/2159-8290.CD-21-1059

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell* 144, 646–674. doi:10.1016/j.cell.2011.02.013

Heldin, C.-H., Rubin, K., Pietras, K., and Östman, A. (2004). High Interstitial Fluid Pressure - an Obstacle in Cancer Therapy. *Nat. Rev. Cancer* 4, 806–813. doi:10.1038/nrc1456

Huang, X., Li, E., Shen, H., Wang, X., Tang, T., Zhang, X., et al. (2020). Targeting the HGF/MET Axis in Cancer Therapy: Challenges in Resistance and Opportunities for Improvement. *Front. Cel Dev. Biol.* 8. doi:10.3389/fcell.2020.00152

Imtiyaz, H. Z., Williams, E. P., Hickey, M. M., Patel, S. A., Durham, A. C., Yuan, L.-J., et al. (2010). Hypoxia-inducible Factor 2α Regulates Macrophage Function in Mouse Models of Acute and Tumor Inflammation. *J. Clin. Invest.* 120, 2699–2714. doi:10.1172/JCI39506

Jeon, S.-M., and Hay, N. (2018). Expanding the Concepts of Cancer Metabolism. *Exp. Mol. Med.* 50, 1–3. doi:10.1038/s12276-018-0070-9

Jung, K.-H., Lee, J. H., Thien Quach, C. H., Paik, J.-Y., Oh, H., Park, J. W., et al. (2013). Resveratrol Suppresses Cancer Cell Glucose Uptake by Targeting Reactive Oxygen Species-Mediated Hypoxia-Inducible Factor-1α Activation. *J. Nucl. Med.* 54, 2161–2167. doi:10.2967/jnumed.112.115436

Kalluri, R., and Weinberg, R. A. (2009). The Basics of Epithelial-Mesenchymal Transition. *J. Clin. Invest.* 119, 1420–1428. doi:10.1172/JCI39104

Ke, Q., and Costa, M. (2006). Hypoxia-inducible Factor-1 (HIF-1). *Mol. Pharmacol.* 70, 1469–1480. doi:10.1124/mol.106.027029

Kogita, A., Togashi, Y., Hayashi, H., Sogabe, S., Terashima, M., de Velasco, M. A., et al. (2014). Hypoxia Induces Resistance to ALK Inhibitors in the H3122 Non-small Cell Lung Cancer Cell Line with an ALK Rearrangement *via* Epithelial-Mesenchymal Transition. *Int. J. Oncol.* 45, 1430–1436. doi:10.3892/IJO.2014.2574

Krishnamachary, B., Penet, M.-F., Nimmagadda, S., Mironchik, Y., Raman, V., Solaiyappan, M., et al. (2012). Hypoxia Regulates CD44 and its Variant Isoforms through HIF-1α in Triple Negative Breast Cancer. *PloS one* 7, e44078. doi:10.1371/JOURNAL.PONE.0044078

Kwak, E. L., Bang, Y.-J., Camidge, D. R., Shaw, A. T., Solomon, B., Maki, R. G., et al. (2010). Anaplastic Lymphoma Kinase Inhibition in Non-small-cell Lung Cancer. *N. Engl. J. Med.* 363, 1693–1703. doi:10.1056/NEJMoa1006448

Ledaki, I., McIntyre, A., Wigfield, S., Buffa, F., McGowan, S., Baban, D., et al. (2015). Carbonic Anhydrase IX Induction Defines a Heterogeneous Cancer Cell Response to Hypoxia and Mediates Stem Cell-like Properties and Sensitivity to HDAC Inhibition. *Oncotarget* 6, 19413–19427. doi:10.18632/oncotarget.4989

Lin, J., Denmeade, S., and Carducci, M. (2009). HIF-1α and Calcium Signaling as Targets for Treatment of Prostate Cancer by Cardiac Glycosides. *Ccdt* 9, 881–887. doi:10.2174/156800909789760249

Liu, N., Wang, Y., Zhou, Y., Pang, H., Zhou, J., Qian, P., et al. (2014). Krüppel-like Factor 8 Involved in Hypoxia Promotes the Invasion and Metastasis of Gastric Cancer *via* Epithelial to Mesenchymal Transition. *Oncol. Rep.* 32, 2397–2404. doi:10.3892/OR.2014.3495

Liu, S., Kumar, S. M., Martin, J. S., Yang, R., and Xu, X. (2011). Snail1 Mediates Hypoxia-Induced Melanoma Progression. *Am. J. Pathol.* 179, 3020–3031. doi:10.1016/J.AJPATH.2011.08.038

Liu, Y., Veena, C. K., Morgan, J. B., Mohammed, K. A., Jekabsons, M. B., Nagle, D. G., et al. (2009). Methylalpinumisoflavone Inhibits Hypoxia-Inducible Factor-1 (HIF-1) Activation by Simultaneously Targeting Multiple Pathways. *J. Biol. Chem.* 284, 5859–5868. doi:10.1074/jbc.M806744200

Lock, F. E., McDonald, P. C., Lou, Y., Serrano, I., Chafe, S. C., Ostlund, C., et al. (2013). Targeting Carbonic Anhydrase IX Depletes Breast Cancer Stem Cells within the Hypoxic Niche. *Oncogene* 32, 5210–5219. doi:10.1038/onc.2012.550

Maher, J. C., Krishan, A., and Lampidis, T. J. (2004). Greater Cell Cycle Inhibition and Cytotoxicity Induced by 2-Deoxy-D-Glucose in Tumor Cells Treated under Hypoxic vs Aerobic Conditions. *Cancer Chemother. Pharmacol.* 53, 116–122. doi:10.1007/s00280-003-0724-7

Marconi, C., Peppicelli, S., Bianchini, F., and Calorini, L. (2013). TNFα Receptor1 Drives Hypoxia-Promoted Invasiveness of Human Melanoma Cells. *Exp. Oncol.* 35, 187–191. Available at: https://pubmed.ncbi.nlm.nih.gov/24084456/ (Accessed December 28, 2021).

Marin-Hernandez, A., Gallardo-Perez, J., Ralph, S., Rodriguez-Enriquez, S., and Moreno-Sanchez, R. (2009). HIF-1α Modulates Energy Metabolism in Cancer Cells by Inducing Over-expression of Specific Glycolytic Isoforms. *Mrmc* 9, 1084–1101. doi:10.2174/138955709788922610

Maschek, G., Savaraj, N., Priebe, W., Braunschweiger, P., Hamilton, K., Tidmarsh, G. F., et al. (2004). 2-deoxy-D-glucose Increases the Efficacy of Adriamycin and Paclitaxel in Human Osteosarcoma and Non-small Cell Lung Cancers *In Vivo*. *Cancer Res.* 64, 31–34. doi:10.1158/0008-5472.can-03-3294

Mathupala, S. P., Ko, Y. H., and Pedersen, P. L. (2009). Hexokinase-2 Bound to Mitochondria: Cancer's Stygian Link to the "Warburg Effect" and a Pivotal Target for Effective Therapy. *Semin. Cancer Biol.* 19, 17–24. doi:10.1016/j.semcancer.2008.11.006

Matsuoka, J., Yashiro, M., Doi, Y., Fuyuhiro, Y., Kato, Y., Shinto, O., et al. (2013). Hypoxia Stimulates the EMT of Gastric Cancer Cells through Autocrine TGFβ Signaling. *PLoS ONE* 8, e62310. doi:10.1371/JOURNAL.PONE.0062310

Newman, R. A., Yang, P., Pawlus, A. D., and Block, K. I. (2008). Cardiac Glycosides as Novel Cancer Therapeutic Agents. *Mol. interventions* 8, 36–49. doi:10.1124/mi.8.1.8

Noman, M. Z., Desantis, G., Janji, B., Hasmim, M., Karray, S., Dessen, P., et al. (2014). PD-L1 Is a Novel Direct Target of HIF-1α, and its Blockade under Hypoxia Enhanced MDSC-Mediated T Cell Activation. *J. Exp. Med.* 211, 781–790. doi:10.1084/jem.20131916

Oh, Y. S., Kim, H. Y., Song, I.-C., Yun, H.-J., Jo, D.-Y., Kim, S., et al. (2012). Hypoxia Induces CXCR4 Expression and Biological Activity in Gastric Cancer Cells through Activation of Hypoxia-Inducible Factor-1α. *Oncol. Rep.* 28, 2239–2246. doi:10.3892/OR.2012.2063

Onishi, H., Morifuji, Y., Kai, M., Suyama, K., Iwasaki, H., and Katano, M. (2012). Hedgehog Inhibitor Decreases Chemosensitivity to 5-fluorouracil and Gemcitabine under Hypoxic Conditions in Pancreatic Cancer. *Cancer Sci.* 103, 1272–1279. doi:10.1111/J.1349-7006.2012.02297.X

Otto, A. M. (2016). Warburg Effect(s)-A Biographical Sketch of Otto Warburg and His Impacts on Tumor Metabolism. *Cancer Metab.* 4, 5. doi:10.1186/s40170-016-0145-9

Palazón, A., Aragonés, J., Morales-Kastresana, A., de Landázuri, M. O., and Melero, I. (2012). Molecular Pathways: Hypoxia Response in Immune Cells Fighting or

Promoting Cancer. *Clin. Cancer Res.* 18, 1207–1213. doi:10.1158/1078-0432.CCR-11-1591

Payne, A. G. (2007). Exploiting Hypoxia in Solid Tumors to Achieve Oncolysis. *Med. hypotheses* 68, 828–831. doi:10.1016/j.mehy.2006.09.013

Pelicano, H., Martin, D. S., Xu, R.-H., and Huang, P. (2006). Glycolysis Inhibition for Anticancer Treatment. *Oncogene* 25, 4633–4646. doi:10.1038/sj.onc.1209597

Pillai, R. N., and Ramalingam, S. S. (2014). Heat Shock Protein 90 Inhibitors in Non-small-cell Lung Cancer. *Curr. Opin. Oncol.* 26, 159–164. doi:10.1097/CCO.0000000000000047

Platten, M., Wick, W., and van den Eynde, B. J. (2012). Tryptophan Catabolism in Cancer: beyond Ido and Tryptophan Depletion. *Cancer Res.* 72, 5435–5440. doi:10.1158/0008-5472.CAN-12-0569

Quintero-Fabián, S., Arreola, R., Becerril-Villanueva, E., Torres-Romero, J. C., Arana-Argáez, V., Lara-Riegos, J., et al. (2019). Role of Matrix Metalloproteinases in Angiogenesis and Cancer. *Front. Oncol.* 9, 1370. doi:10.3389/FONC.2019.01370/BIBTEX

Raj, S., Kesari, K. K., Kumar, A., Rathi, B., Sharma, A., Gupta, P. K., et al. (2022). Molecular Mechanism(s) of Regulation(s) of C-MET/HGF Signaling in Head and Neck Cancer. *Mol. Cancer* 21, 31. doi:10.1186/s12943-022-01503-1

Rami, M., Dubois, L., Parvathaneni, N.-K., Alterio, V., van Kuijk, S. J. A., Monti, S. M., et al. (2013). Hypoxia-targeting Carbonic Anhydrase IX Inhibitors by a New Series of Nitroimidazole-Sulfonamides/sulfamides/sulfamates. *J. Med. Chem.* 56, 8512–8520. doi:10.1021/jm4009532

Semenza, G. L., Roth, P. H., Fang, H. M., and Wang, G. L. (1994). Transcriptional Regulation of Genes Encoding Glycolytic Enzymes by Hypoxia-Inducible Factor 1. *J. Biol. Chem.* 269, 23757–23763. doi:10.1016/s0021-9258(17)31580-6

Simiantonaki, N., Taxeidis, M., Jayasinghe, C., Kurzik-Dumke, U., and Kirkpatrick, C. J. (2008). Hypoxia-inducible Factor 1 Alpha Expression Increases during Colorectal Carcinogenesis and Tumor Progression. *BMC Cancer* 8, 320. doi:10.1186/1471-2407-8-320

Singh, A., and Settleman, J. (2010). EMT, Cancer Stem Cells and Drug Resistance: an Emerging axis of Evil in the War on Cancer. *Oncogene* 29, 4741–4751. doi:10.1038/ONC.2010.215

Sørensen, B. S., and Horsman, M. R. (2020). Tumor Hypoxia: Impact on Radiation Therapy and Molecular Pathways. *Front. Oncol.* 10. doi:10.3389/fonc.2020.00562

Spinella, F., Caprara, V., Cianfrocca, R., Rosanò, L., di Castro, V., Garrafa, E., et al. (2014). The Interplay between Hypoxia, Endothelial and Melanoma Cells Regulates Vascularization and Cell Motility through Endothelin-1 and Vascular Endothelial Growth Factor. *Carcinogenesis* 35, 840–848. doi:10.1093/CARCIN/BGU018

Sutherland, R. M. (1988). Cell and Environment Interactions in Tumor Microregions: the Multicell Spheroid Model. *Science* 240, 177–184. doi:10.1126/SCIENCE.2451290

Tam, S. Y., Wu, V. W. C., and Law, H. K. W. (2020). Hypoxia-Induced Epithelial-Mesenchymal Transition in Cancers: HIF-1α and beyond. *Front. Oncol.* 10. doi:10.3389/fonc.2020.00486

Tan, E. Y., Yan, M., Campo, L., Han, C., Takano, E., Turley, H., et al. (2009). The Key Hypoxia Regulated Gene CAIX Is Upregulated in Basal-like Breast Tumours and Is Associated with Resistance to Chemotherapy. *Br. J. Cancer* 100, 405–411. doi:10.1038/SJ.BJC.6604844

Tao, L.-L., Shi, S. J., Chen, L. B., and Huang, G. C. (2014). Expression of Monocyte Chemotactic protein-1/CCL2 in Gastric Cancer and its Relationship with Tumor Hypoxia. *Wjg* 20, 4421–4427. doi:10.3748/WJG.V20.I15.4421

Touisni, N., Maresca, A., McDonald, P. C., Lou, Y., Scozzafava, A., Dedhar, S., et al. (2011). Glycosyl Coumarin Carbonic Anhydrase IX and XII Inhibitors Strongly Attenuate the Growth of Primary Breast Tumors. *J. Med. Chem.* 54, 8271–8277. doi:10.1021/jm200983e

Valastyan, S., and Weinberg, R. A. (2011). Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell* 147, 275–292. doi:10.1016/J.CELL.2011.09.024

Vaupel, P., and Multhoff, G. (2021). Revisiting the Warburg Effect: Historical Dogma versus Current Understanding. *J. Physiol.* 599, 1745–1757. doi:10.1113/JP278810

Vaupel, P. (2004). The Role of Hypoxia-Induced Factors in Tumor Progression. *The oncologist* 9 (Suppl. 5), 10–17. doi:10.1634/THEONCOLOGIST.9-90005-10

Wang, G. L., Jiang, B. H., Rue, E. A., and Semenza, G. L. (1995). Hypoxia-inducible Factor 1 Is a basic-helix-loop-helix-PAS Heterodimer Regulated by Cellular O2 Tension. *Proc. Natl. Acad. Sci.* 92, 5510–5514. doi:10.1073/pnas.92.12.5510

Wang, M., Qiu, S., and Qin, J. (2019). Baicalein Induced Apoptosis and Autophagy of Undifferentiated Thyroid Cancer Cells by the ERK/PI3K/Akt Pathway. *Am. J. Transl Res.* 11, 3341–3352.

Wang, Y., Liu, X., Zhang, H., Sun, L., Zhou, Y., Jin, H., et al. (2014). Hypoxia-Inducible lncRNA-AK058003 Promotes Gastric Cancer Metastasis by Targeting γ-Synuclein. *Neoplasia* 16, 1094–1106. doi:10.1016/J.NEO.2014.10.008

Wei, L., Zhou, Y., Qiao, C., Ni, T., Li, Z., You, Q., et al. (2015). Oroxylin A Inhibits Glycolysis-dependent Proliferation of Human Breast Cancer *via* Promoting SIRT3-Mediated SOD2 Transcription and HIF1α Destabilization. *Cell Death Dis* 6–e1714. doi:10.1038/cddis.2015.86

Wei, R., Mao, L., Xu, P., Zheng, X., Hackman, R. M., Mackenzie, G. G., et al. (2018). Suppressing Glucose Metabolism with Epigallocatechin-3-Gallate (EGCG) Reduces Breast Cancer Cell Growth in Preclinical Models. *Food Funct.* 9, 5682–5696. doi:10.1039/c8fo01397g

White, P. T., Subramanian, C., Zhu, Q., Zhang, H., Zhao, H., Gallagher, R., et al. (2016). Novel HSP90 Inhibitors Effectively Target Functions of Thyroid Cancer Stem Cell Preventing Migration and Invasion. *Surgery* 159, 142–151. doi:10.1016/j.surg.2015.07.050

Wigerup, C., Påhlman, S., and Bexell, D. (2016). Therapeutic Targeting of Hypoxia and Hypoxia-Inducible Factors in Cancer. *Pharmacol. Ther.* 164, 152–169. doi:10.1016/j.pharmthera.2016.04.009

Xiang, L., Gilkes, D. M., Chaturvedi, P., Luo, W., Hu, H., Takano, N., et al. (2014). Ganetespib Blocks HIF-1 Activity and Inhibits Tumor Growth, Vascularization, Stem Cell Maintenance, Invasion, and Metastasis in Orthotopic Mouse Models of Triple-Negative Breast Cancer. *J. Mol. Med.* 92, 151–164. doi:10.1007/s00109-013-1102-5

Xu, R. H., Pelicano, H., Zhou, Y., Carew, J. S., Feng, L., Bhalla, K. N., et al. (2005). Inhibition of Glycolysis in Cancer Cells: a Novel Strategy to Overcome Drug Resistance Associated with Mitochondrial Respiratory Defect and Hypoxia. *Cancer Res.* 65, 613–621. doi:10.1158/0008-5472.613.65.2

Yang, M.-H., Wu, M.-Z., Chiou, S.-H., Chen, P.-M., Chang, S.-Y., Liu, C.-J., et al. (2008). Direct Regulation of TWIST by HIF-1α Promotes Metastasis. *Nat. Cel Biol* 10, 295–305. doi:10.1038/NCB1691

Yang, Y. J., Na, H. J., Suh, M. J., Ban, M. J., Byeon, H. K., Kim, W. S., et al. (2015). Hypoxia Induces Epithelial-Mesenchymal Transition in Follicular Thyroid Cancer: Involvement of Regulation of Twist by Hypoxia Inducible Factor-1α. *Yonsei Med. J.* 56, 1503. doi:10.3349/YMJ.2015.56.6.1503

Zhang, F., and Du, G. (2012). Dysregulated Lipid Metabolism in Cancer. *Wjbc* 3, 167–174. doi:10.4331/wjbc.v3.i8.167

Zhang, H., Chen, J., Liu, F., Gao, C., Wang, X., Zhao, T., et al. (2014). CypA, a Gene Downstream of HIF-1α, Promotes the Development of PDAC. *PloS one* 9, e92824. doi:10.1371/JOURNAL.PONE.0092824

Zhang, H., Qian, D. Z., Tan, Y. S., Lee, K., Gao, P., Ren, Y. R., et al. (2008). Digoxin and Other Cardiac Glycosides Inhibit HIF-1α Synthesis and Block Tumor Growth. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19579–19586. doi:10.1073/pnas.0809763105

Zhao, X., Gao, S., Ren, H., Sun, W., Zhang, H., Sun, J., et al. (2014). Hypoxia-inducible Factor-1 Promotes Pancreatic Ductal Adenocarcinoma Invasion and Metastasis by Activating Transcription of the Actin-Bundling Protein Fascin. *Cancer Res.* 74, 2455–2464. doi:10.1158/0008-5472.CAN-13-3009

Zhou, J., Schmid, T., Schnitzer, S., and Brüne, B. (2006). Tumor Hypoxia and Cancer Progression. *Cancer Lett.* 237, 10–21. doi:10.1016/j.canlet.2005.05.028

Zhu, H., Wang, D., Zhang, L., Xie, X., Wu, Y., Liu, Y., et al. (2014). Upregulation of Autophagy by Hypoxia-Inducible Factor-1α Promotes EMT and Metastatic Ability of CD133+ Pancreatic Cancer Stem-like Cells during Intermittent Hypoxia. *Oncol. Rep.* 32, 935–942. doi:10.3892/OR.2014.3298

Zou, Y.-m., Hu, G.-y., Zhao, X.-q., Lu, T., Zhu, F., Yu, S.-y., et al. (2014). Hypoxia-induced Autophagy Contributes to Radioresistance *via* C-Jun-Mediated Beclin1 Expression in Lung Cancer Cells. *J. Huazhong Univ. Sci. Technol. [Med. Sci.* 34, 761–767. Journal of Huazhong University of Science and Technology. Medical sciences = Hua zhong ke ji da xue xue bao. Yi xue Ying De wen ban = Huazhong keji daxue xuebao. doi:10.1007/S11596-014-1349-2

Conflict of Interest: SS was employed by the company Kashiv Biosciences.

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

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

GLOSSARY

ATP adenosine triphosphate
AXL AXL receptor tyrosine kinase
BECN1 beclin-1
CAIX carbonic anhydrase 9 precursor
CRC colorectal cancer
DIG digoxin
EGCG epigallocatechin gallate
EMT epithelial-mesenchymal transition
ENO-α alpha-enolase
GLUT glucose transporter
HIF 1α-hypoxia-inducible factor-1α
HIF-1 hypoxia-inducible factor-1
HGF hepatocyte growth factor
HK1 hexokinase-1
hsp90 heat shock protein 90
HTM hypoxic tumor microenvironment
ILK Integrin Linked kinase
LDH lactate dehydrogenase
MAP1LC3 microtubule-associated proteins 1A/1B light chain 3B
MF methylalpinumisoflavone
MET mesenchymal-epithelial transition
NTRK1 neurotrophic receptor tyrosine kinase 1
NSCLC non-small lung cancer
OXPHOS oxidative phosphorylation
PDAC pancreatic ductal adenocarcinoma
PD-L1 programmed death ligand-1
RCC renal cell carcinoma
RNA ribonucleic acid
RET rearranged during transfection
PGK1 phosphoglycerate kinase 1
PFKFB 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3
PFK Phosphofructokinase
PTEN phosphatase and tensin homolog
PYK M2-M2 isoform of pyruvate kinase
ROS1 ROS proto-oncogene 1
TME tumor microenvironment
TNBC triple-negative breast cancer
VEGF vascular endothelial growth factor
VHL Von Hippel-Lindau

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 31 May 2022
doi: 10.3389/fgene.2022.866473

Differential Expression of Genes Regulating Store-operated Calcium Entry in Conjunction With Mitochondrial Dynamics as Potential Biomarkers for Cancer: A Single-Cell RNA Analysis

Mangala Hegde[1,2], *Uzini Devi Daimary*[1,2], *Sandra Jose*[1,2], *Anjana Sajeev*[1,2], *Arunachalam Chinnathambi*[3], *Sulaiman Ali Alharbi*[3], *Mehdi Shakibaei*[4] and *Ajaikumar B. Kunnumakkara*[1,2]*

[1]Cancer Biology Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology-Guwahati, Guwahati, India, [2]DBT-AIST International Center for Translational and Environmental Research, Indian Institute of Technology-Guwahati, Guwahati, India, [3]Department of Botany and Microbiology, College of Science, King Saud University, Riyadh, Saudi Arabia, [4]Musculoskeletal Research Group and Tumor Biology, Faculty of Medicine, Institute of Anatomy, Ludwig-Maximilian-University Munich, Munich, Germany

Regulation of intracellular concentration of calcium levels is crucial for cell signaling, homeostasis, and in the pathology of diseases including cancer. Agonist-induced entry of calcium ions into the non-excitable cells is mediated by store-operated calcium channels (SOCs). This pathway is activated by the release of calcium ions from the endoplasmic reticulum and further regulated by the calcium uptake through mitochondria leading to calcium-dependent inactivation of calcium-release activated calcium channels (CARC). SOCs including stromal interaction molecules (STIM) and ORAI proteins have been implicated in tumor growth, progression, and metastasis. In the present study, we analyzed the mRNA and protein expression of genes mediating SOCs—STIM1, STIM2, ORAI1, ORAI2, ORAI3, TRPC1, TRPC3, TRPC4, TRPC5, TRPC6, TRPC7, TRPV1, TRPV2, TRPM1, and TRPM7 in head and neck squamous cell cancer (HNSC) patients using TCGA and CPTAC analysis. Further, our *in silico* analysis showed a significant correlation between the expression of SOCs and genes involved in the mitochondrial dynamics (MDGs) both at mRNA and protein levels. Protein-protein docking results showed lower binding energy for SOCs with MDGs. Subsequently, we validated these results using gene expression and single-cell RNA sequencing datasets retrieved from Gene Expression Omnibus (GEO). Single-cell gene expression analysis of HNSC tumor tissues revealed that SOCs expression is remarkably associated with the MDGs expression in both cancer and fibroblast cells.

Keywords: head and neck cancer, store-operated calcium channels, mitochondrial dynamics, TCGA database, CPTAC database, gene expression omnibus

OPEN ACCESS

Edited by:
Rajkumar S Kalra, Okinawa Institute of Science and Technology Graduate University, Japan

Reviewed by:
Jeremy Smyth, Uniformed Services University of the Health Sciences, United States
Prashanth Thevkar Nagesh, Harvard Medical School, United States

***Correspondence:**
Ajaikumar B. Kunnumakkara
kunnumakkara@iitg.ac.in
ajai78@gmail.com

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *31 January 2022*
Accepted: *27 April 2022*
Published: *31 May 2022*

Citation:
Hegde M, Daimary UD, Jose S, Sajeev A, Chinnathambi A, Alharbi SA, Shakibaei M and Kunnumakkara AB (2022) Differential Expression of Genes Regulating Store-operated Calcium Entry in Conjunction With Mitochondrial Dynamics as Potential Biomarkers for Cancer: A Single-Cell RNA Analysis. Front. Genet. 13:866473. doi: 10.3389/fgene.2022.866473

INTRODUCTION

Calcium (Ca^{2+}) is a ubiquitous intracellular second messenger that controls a wide range of physiological and pathological processes (Berridge et al., 2003). In metazoans including humans, the store-operated calcium channel entry (SOCE) is the predominant calcium entry mechanism in non-excitable cells (Parekh and Putney, 2005). The concept of SOCE was first described by James Putney in 1986; however, the molecular signatures and functional validation of SOCE were identified more recently (Putney, 1986). SOCs are activated when Ca^{2+} is released from the endoplasmic reticulum, which is necessary for cells to replenish Ca^{2+} after signaling processes (Parekh and Putney, 2005). SOCE is implicated in a wide range of biological processes such as transcriptional regulation of gene expression, exocytosis, cellular metabolism, and cell motility (Lewis, 2007; Parekh and Putney, 2005). Among the SOCs, STIM1, and ORAI1 have been shown to be widely expressed. Additionally, closely associated channel subfamilies like transient receptor potential channels (TRPC), TRPV (vanilloid), and TRPM (melastatin) have also been shown to be involved in SOCE mediated Ca^{2+} influx (Authi, 2007; Ma et al., 2011; Bastián-Eugenio et al., 2019). A plethora of studies showing the involvement of Ca^{2+} in various stages of cancer development and progression led Yang et al. (2009) to investigate the role of STIM1 and ORAI1 in breast cancer cell migration, invasion, and metastasis (Yang et al., 2009). In addition, transcriptomic analysis of glioblastoma tumor tissues showed overexpression of STIM1, ORAI1, and TRPC1 (Scrideli et al., 2008; Alptekin et al., 2015). Moreover, Zhang et al. (2013) demonstrated that SOCE mediated influx of Ca^{2+} regulated the migration and metastasis of nasopharyngeal carcinoma both *in vitro* and in zebrafish models (Zhang et al., 2013). Further, Jiang et al. (2007) revealed that TRPM7 is necessary for the proliferation and growth of FaDu and SCC25 cells *in vitro* by siRNA-mediated knockdown of TRPM7 (Jiang et al., 2007). SOCs have also been proposed as potential therapeutic targets for various inflammatory disorders and cancer (Feske, 2019; Khan et al., 2020; Chang et al., 2021). Recent studies have shown that non-steroidal anti-inflammatory drugs (NSAIDs) including sulindac, salicylate, flurbiprofen, and indomethacin inhibited SOCEs in colon cancer cells (Hernández-Morales et al., 2017; Villalobos et al., 2019). Furthermore, Gualadani et al. (2019) demonstrated the requisite of SOCE for the anti-cancer effect of cisplatin in non-small cell lung carcinoma (Gualdani et al., 2019). Therefore, these studies suggest that SOCs can be used as cancer diagnostic biomarkers and therapeutic targets.

According to the latest GLOBOCAN report on cancer burden worldwide, the prevalence of head and neck cancer is steadily increasing (Sung et al., 2021). Apart from the well-known risk factors such as tobacco, alcohol, and human papillomavirus (HPV), vitamin D insufficiency and defects in calcium signaling have recently been found to play a significant role in the initiation and progression of head and neck cancer (Orell-kotikangas et al., 2012; Singh et al., 2020).

Recently, studies have shown that ORAI, ORAI2, and STIM1 were significantly elevated in tissues from oral squamous cancer patients compared to normal samples. These studies also reported the significant reduction in proliferation, migration, and invasion upon siRNA-mediated knockdown of ORAI1, ORAI2, and STIM1 in oral cancer cell lines *in vitro* (Singh et al., 2020; Wang et al., 2022). In the present study, we analyzed the gene and protein expression of SOCs—STIM1, STIM2, ORAI1, ORAI2, ORAI3, TRPC1, TRPC3, TRPC4, TRPC5, TRPC6, TRPC7, TRPV1, TRPV2, TRPM1, and TRPM7 in HNSC using The Cancer Genome Atlas (TCGA) and Clinical Proteomic Tumor Analysis Consortium (CPTAC) analysis. In addition, mitochondrial regulation of calcium ions has shown to play an important role in SOCs mediated calcium entry, and tumor cell mitochondrial dysfunction is proposed to be responsible for SOCs upregulation (Villalobos et al., 2018). Hence, we further analyzed the genes involved in mitochondrial dynamics (MDGs) and found a significant correlation between the expression of SOCs and MDGs both at mRNA and protein levels. Further, docking results showed lower binding energy for SOCs with MDGs. Subsequently, the validation of these results was carried out using datasets downloaded from gene expression omnibus (GEO). Interestingly, single-cell RNA sequence analysis revealed that gene expression of SOCs is remarkably associated with the MDGs in both cancer and fibroblast cells.

MATERIALS AND METHODS

The methodology of the overall study have been represented in the form of graphical abstract (**Supplementary Figure 1**).

Gene Ontology Analysis

As a preliminary analysis to show the role of SOCs in signaling pathways we conducted functional enrichment analysis (FEA) to annotate gene ontology (GO) including biological processes (BP), cellular components (CC), molecular function (MF), and pathway enrichment analysis using the Kyoto Encyclopedia of Genes and Genomics (KEGG: http://www.genome.jp) (Kanehisa and Goto, 2000). The MF which showed the highest score for SOCs was visualized using gProfiler (Raudvere et al., 2019). The molecular function with significant *p* adj values are shown. These functions are assigned based on either the experiment or Sequence Model (ISM) or Sequence Alignment (ISA) or Sequence Orthology (ISO) or Sequence or structural similarity (ISS) or Genomic context (IGC) or Biological aspect of ancestor (IBA) or Rapid divergence (IRD) or Reviewed computational analysis (RCA) or Electronic annotation (IEA). The disease ontology analysis was conducted to understand the involvement of SOCs in various cancers. The analysis and visualization for disease ontology were performed using the clusterProfiler package developed by Bioconductor for R statistical environment (Yu et al., 2012; Wu et al., 2022). The adjusted *p*-value of less than 0.05 are considered to be significant.

The Cancer Genome Atlas (TCGA) Analysis

In the next step, we carried out the TCGA analysis to understand the clinical relevance of SOCs in head and neck cancers. The HiSeqV2 TCGA level 3 gene expression data was downloaded using TCGA biolinks version 2.15.1 developed for R statistical environment (Mounir et al., 2019). The data contained 546 tumor samples and 44 normal samples. Correlation analysis was conducted using Corrplot and ggplot2 packages (Wickham et al., 2016; Wei et al., 2017) and survival analysis was performed using Kaplan-Meier

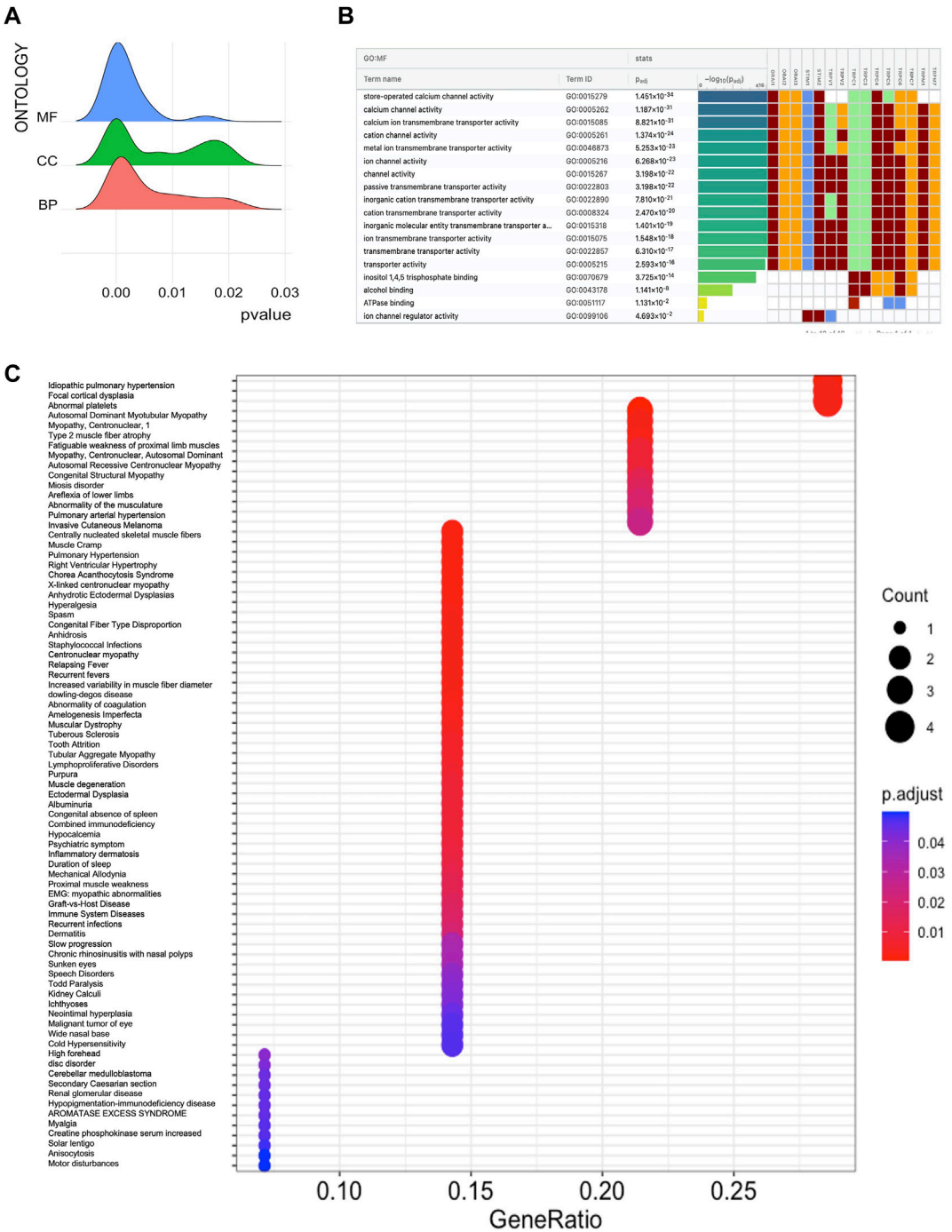


FIGURE 1 | Molecular function and disease ontology of SOCs. **(A)** Ridge plot showing the predominant Gene ontology analysis using ClusterProfile R package. Inset- MF: Molecular function (blue); CC: Cellular component (green); BP: Biological processes (orange) **(B)** Gene ontology analysis using g:Profiler. Molecular function (MF) which showed significant adjusted *p*-value of SOCs is showed. Cumulative hypergeometric padj values are used to dentify the most significant molecular functions. The $-\log_{10}$padj column is colored according to the rank with highest being purple to lowest being yellow. The color for each gene is given as follows: red for inferred from the experiment; brown for Sequence Model (ISM), Sequence Alignment (ISA), Sequence Orthology (ISO) or Sequence or structural similarity (ISS), Genomic context (IGC) or Biological aspect of ancestor (IBA), Rapid divergence (IRD); blue for Reviewed computational analysis (RCA), Electronic annotation (IEA); green for traceable author or non traceable author or inferred by curator **(C)** Disease enrichment analysis of SOCs using ClusterProfile R package. The enriched diseases with *p* value less than 0.05 are shown. The dots represent the enrichment of genes with red color being high enrichment and blue being low enrichment. The size of the dot represent the count of each term in the particular row.

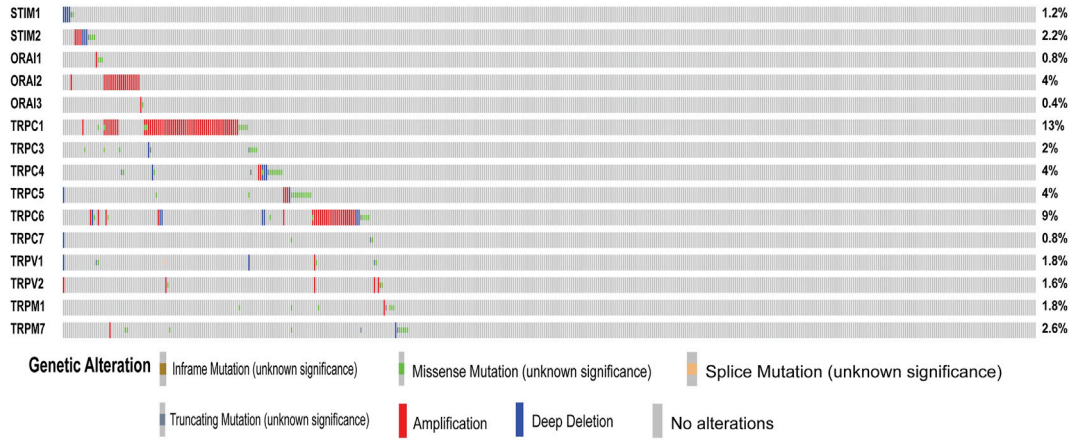


FIGURE 2 | Mutations in SOCs. The mutations in genes associated with SOCs were analysed across the TCGA fire hose (n = 504) datasets using Cbioportal. Oncoprint plot was downloaded from the Cbioportal website. The color for different genetic alterations are mentioned at the bottom. Each row in the plot represents a gene and column represents a tumor sample. The percentage of reported mutation rate across the tumor samples (n = 504) are represented on the right-hand side.

plotter (http://kmplot.com) (Lánczky and Győrffy, 2021; Nagy et al., 2021). The significance and hazard ratio are shown. Further, we conducted mutational analysis to determine the genomic plasticity of SOCs in head neck cancers. Mutations in potential genes encoding SOCs across the TCGA fire hose dataset for HNSC (n = 504) are represented as oncoprint plot downloaded from http://cbioportal.org (Cerami et al., 2012). The row of the plot indicates a gene and column indicates the tumor sample.

Expression Array Data

To further validate the results, we analysed the various expression datasets. Expression profiling dataset GSE171898 contained a total of 323 samples including 208 OSCC tissues from patients treated at Washington University St. Louis and 115 OSCC tissues from patients treated at Vanderbilt University. Illumina Hiseq 3,000 expression profile array data for these samples are available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE171898. The patients were stratified based on the HPV data. The platform used for the GSE17898 expression profiling array was GPL21290 and the spot ID is available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL21290 (Liu et al., 2020).

Single-cell RNA-Seq analysis of head and neck cancer data (GSE103322) were downloaded from the NCBI GEO (Gene Expression Omnibus) database (https://www.ncbi.nlm.nih.gov/geo/). The data consists of an expression profile of 5,902 single cells from 18 patients. The platform used for the GSE103322 expression profiling array was GPL18573 and the spot ID is available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL18573 (Puram et al., 2017).

Store-Operated Calcium Channels (SOCs) in Mitochondrial Dynamics

The reorganization of ER and Ca^{2+} ion transfer have been implicated in the mitochondrial fission and apoptosis of cancer cells (Yedida et al., 2019). Hence, we next analsyed the relevance of SOCs with genes involved in mitochondrial dynamics. The correlation between expression of genes involved in SOCs and genes involved in mitochondrial genes for all HNSC patients, and HPV positive and negative HNSC patients were downloaded using Timer 2.0 (http://timer.comp-genomics.org/) (Li et al., 2020). The data was visualized using the heatmap function of Graphpad Prism software version 9.2.1. Subsequently, the SOC protein structures and proteins involved in mitochondrial dynamics were downloaded from the protein data bank (Berman et al., 2000). Protein structures for SOCs—STIM1:PDB ID-6YEL (Rathner et al., 2021); STIM2:PDB ID-2L5Y (Zheng et al., 2011); ORAI1:PDB ID-4EHQ (Liu et al., 2012); TRPC5:PDB ID-6YSN (Wright et al., 2020) and MDGs-DNM1L:PDB ID-3W6N (Kishida and Sugio, 2013); MFN1:PDB ID-5GOF (Cao et al., 2017); MFN2:PDB ID-6JFK (Li et al., 2019); OPA1:PDB ID-6JTG (Yu et al., 2020) and FIS1:PDB ID-1PC2 (Suzuki et al., 2003) were downloaded. The docking was performed using Clust Pro version 2.0 (https://cluspro.bu.edu/login.php) (Kozakov et al., 2013; Kozakov et al., 2017; Vajda et al., 2017; Desta et al., 2020) and binding energy was calculated using PRODIGY (https://wenmr.science.uu.nl/prodigy/) (Vangone and Bonvin, 2015; Xue et al., 2016; Vangone et al., 2019). Gene regulatory network was analyzed using Geneck (https://lce.biohpc.swmed.edu/geneck/) (Zhang et al., 2019) and protein-protein interaction network was obtained from string database (https://string-db.org/) (Szklarczyk et al., 2019; Szklarczyk et al., 2021).

Clinical Proteomic Tumor Analysis Consortium (CPTAC) Data Analysis

Since, the protein expression is critical for the signaling pathways and physiological responses, we next analysed

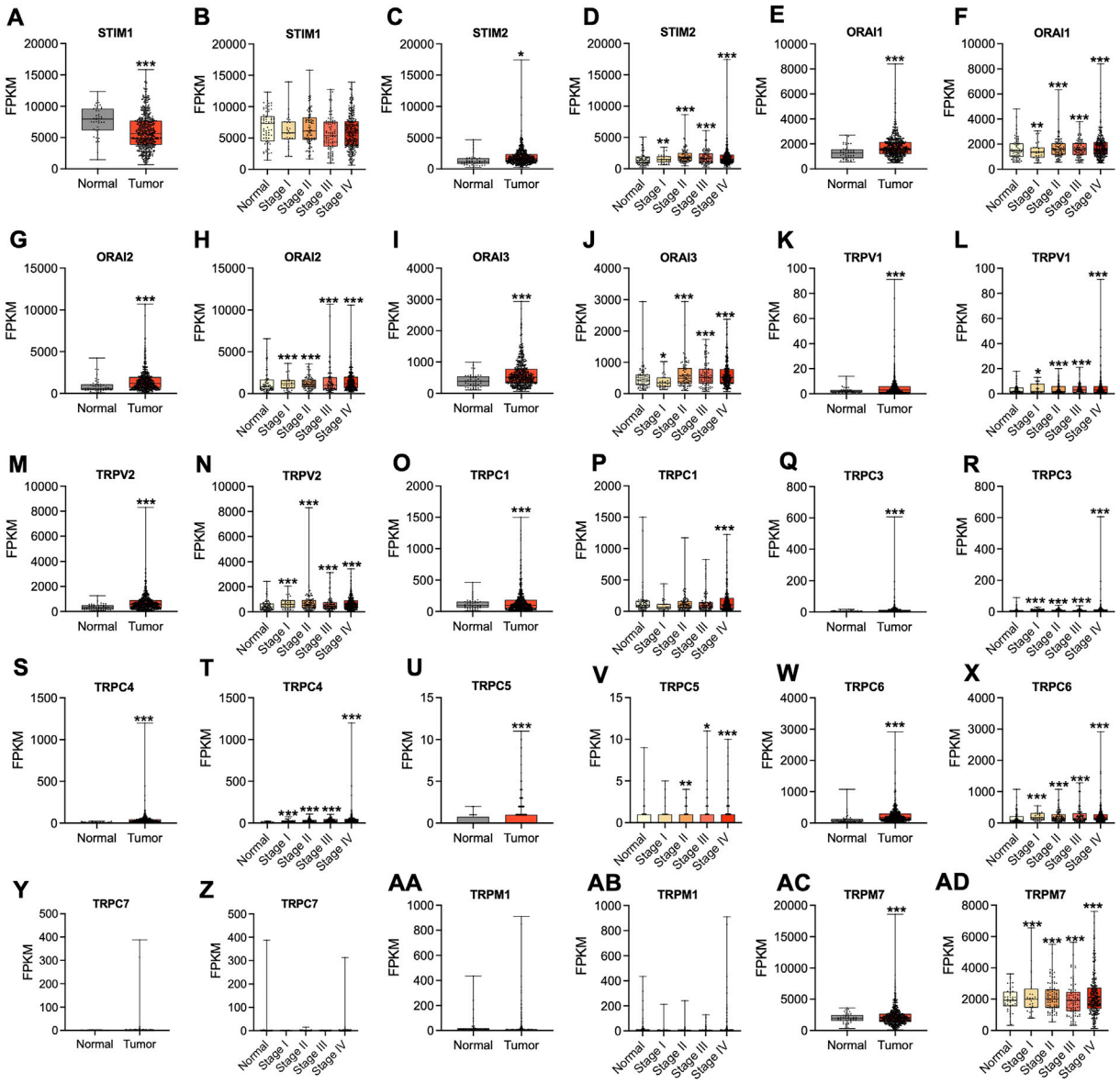


FIGURE 3 | Altered mRNA expression of SOCs in HNSC. The Hiseq2 gene expression and clinical details of TCGA-HNSC data was downloaded from the TCGA biolinks–Bioconductor package for R statistical environment. The expression levels of SOCs (represented as fragments per kilobase per million mapped fragments-FPKM) in normal versus tumor **(A,C,E,G,I,K,M,O,Q,S,U,W, Y, AA, AC)** and different stages- stage I, stage II, stage III, and stage IV of HNSC **(B,D,F,H,J,L,N,P,R,T,V,X, Z, AB, AD)** are plotted. The statistical significance is represented as $*p < 0.05$, $**p < 0.01$, $***p < 0.0001$.

the proteomic and phopshoproteomic levels of SOCs in HNSC. The CPTAC data for proteomics and phosphoproteomics with clinical data for HNSC was downloaded using Python version 3.0 (Huang et al., 2021). The bar plots were plotted using Graphpad Prism version 9.2.1.

Histology Analysis

The histopathology images for SOCs for HNSC were obtained from Human Protein atlas (https://www.proteinatlas.org/) (Uhlén et al., 2015) that is not active.

Correlation Analysis

The correlation analysis for SOCs and MDGs gene expression and protein expression were performed using Corrplot package for the R programming (Wei et al., 2017).

Statistical Analysis

Correlation between different parameters was calculated by Pearson correlation analysis. Differences between the groups were evaluated using nonparametric t-test. p-value < 0.05 was considered to be significant for all the TCGA, CPTAC, and GEO datasets analysis. Statistical analysis was performed using R programming version 3.6.1.

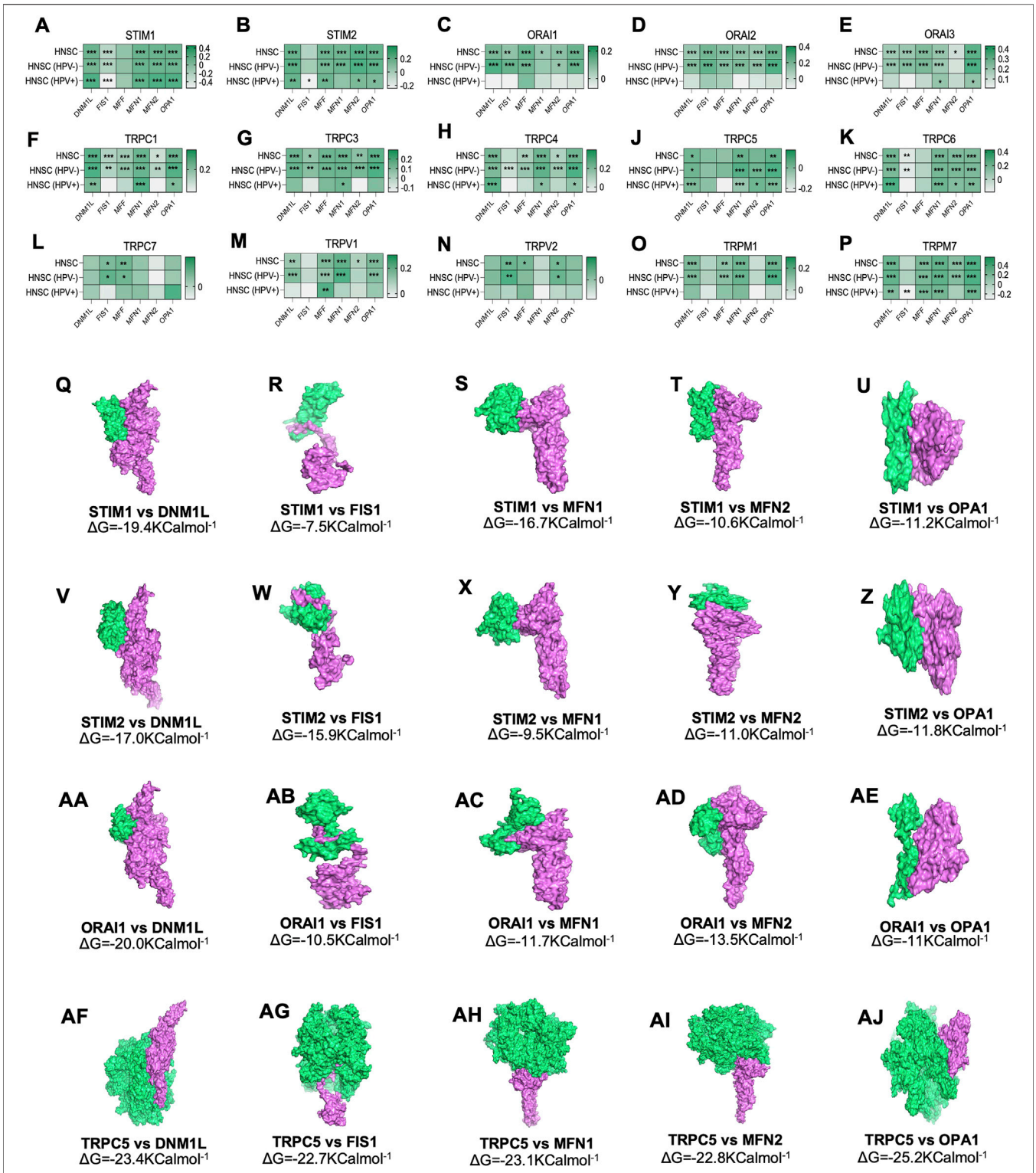
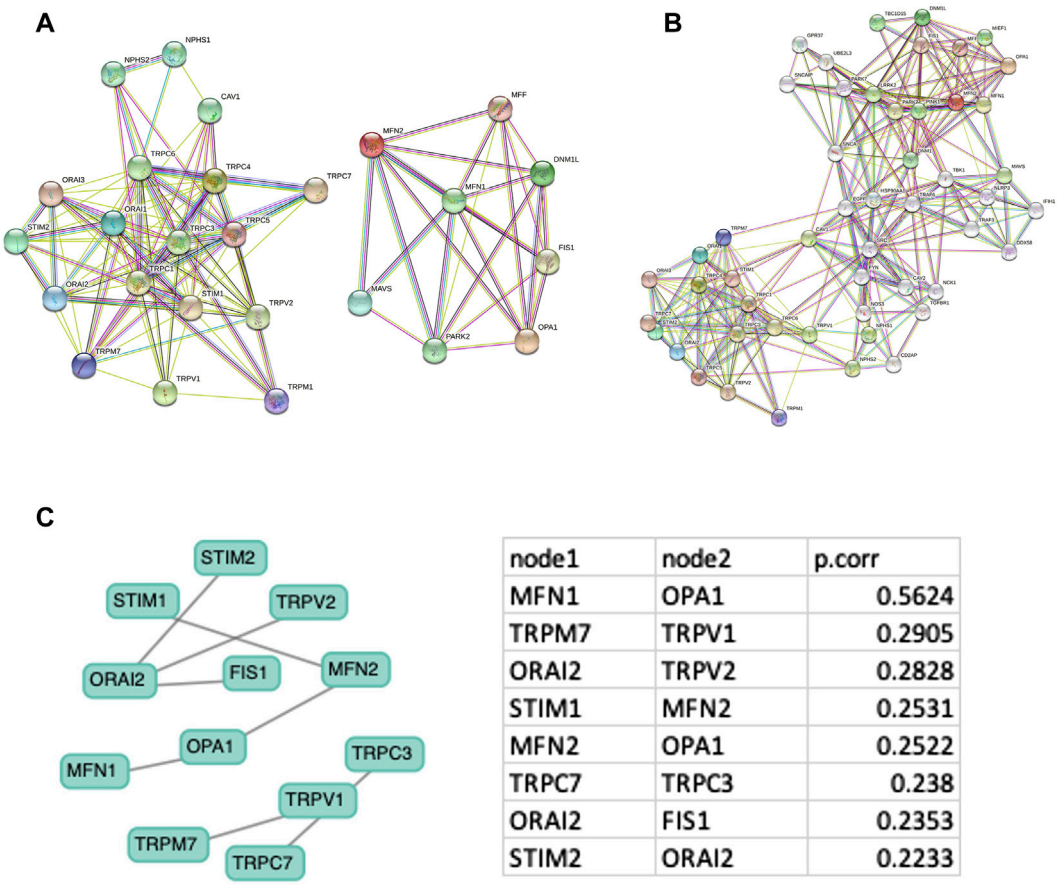


FIGURE 4 | Altered mRNA expression of SOCs are correlated with the expression of MDGs in HNSC. **(A–P)** The Hiseq2 gene expression TCGA-HNSC data was downloaded. The expression levels of SOCs are compared with the expression of MDGs using Timer 2.0 tool. The statistical significance is represented in asterisk. *p < 0.05, **<0.01, ***p < 0.001. **(Q–AJ)** Protein structures for SOCs- STIM1 (PDB ID-6YEL), STIM2 (PDB ID-2L5Y), ORAI1 (PDB ID-4EHQ), TRPC5 (PDB ID-6YSN)- and MDGs-DNM1L (PDB ID-3W6N), MFN1 (PDB ID-5GOF), MFN2 (PDB ID-6JFK), OPA1(PDB ID-6JTG) and FIS1 (PDB ID-1PC2) were downloaded in the. pdb format from protein data bank and docked using ClustPro tool. The protein in green represents SOCs and pink represents MDGs. Prodigy was used to calculate ΔG.



node1	node2	p.corr
MFN1	OPA1	0.5624
TRPM7	TRPV1	0.2905
ORAI2	TRPV2	0.2828
STIM1	MFN2	0.2531
MFN2	OPA1	0.2522
TRPC7	TRPC3	0.238
ORAI2	FIS1	0.2353
STIM2	ORAI2	0.2233

FIGURE 5 | Protein-protein network analysis. Protein-protein interaction network with minimum node **(A)** and increased node **(B)** were downloaded from STRING database **(C)** Gene interaction network is downloaded from GeNeck tool and Pearson correlation for each node is shown.

RESULTS

In the present study, we analyzed the expression of SOCs mRNA, proteins, and phosphoproteins of HNSC by *in silico* approaches. In addition, we have also shown the mutations in these genes to understand the effect of genetic alterations on their expression in HNSC patients. Next, we cohered the gene and protein expression levels of SOCs versus MDGs and correlated their expression with the survival risk of HNSC patients. Further, we performed docking to understand the binding efficiency of SOCE proteins with MDGs. To our knowledge, this is the first comprehensive study that links the SOCs gene expression to MDGs and potential risk of early death in HNSC.

Differential Expression of Store-Operated Calcium Channel Entry in Head and Neck Squamous Cell Cancer

In the first step, GO and KEGG pathway analysis was carried out for fifteen SOCs genes- STIM1, STIM2, ORAI1, ORAI2, ORAI3, TRPC1, TRPC3, TRPC4, TRPC5, TRPC6, TRPC7, TRPV1, TRPV2, TRPM1, and TRPM7 using clusterProfiler to understand their function in MF, CC, and BP. The genes were enriched mainly for MF (**Figure 1A**). The molecular function enrichment using gProfiler identified SOCs genes involvement in inositol binding activity and ATPase binding activity which are crucial for cell signaling and regulation of cellular functions apart from the usual calcium transporter, ion transporter, and transmembrane transporter activities of SOCs (**Figure 1B**). The disease enrichment analysis using ClusterProfiler showed that these genes are involved in pulmonary hypertension, disorders related to muscles, lymphoproliferative disorders, malignant eye tumors, cerebellar medulloblastoma, immune system diseases, and inflammatory disorders (**Figure 1C**). This indicated that SOCs play critical role in tumor development and progression. Further, we analysed for the mutation in SOCs across TCGA fire hose datasets containing 504 samples for HNSC. The mutation analysis showed 1.2% mutation in STIM1 (6/504), 2.2% mutation in STIM2 (11/504), 0.8% mutation in ORAI1 (4/504), 4% mutation in ORAI2 (20/504), 0.4% mutation in ORAI3 (2/504), 13% in TRPC1 (64/504), 2% mutation in TRPC3 (10/504), 4% mutation in TRPC4 (20/504),

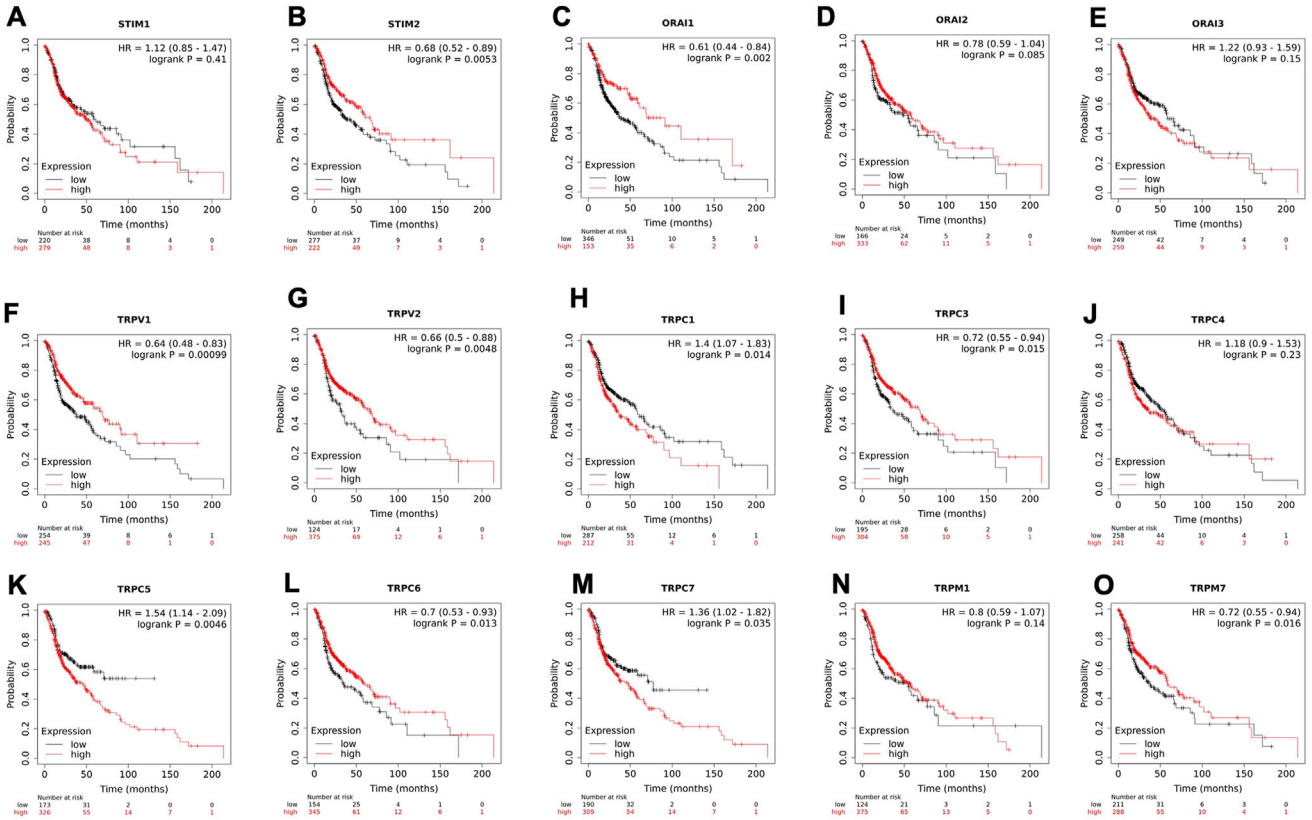


FIGURE 6 | The altered expression of SOCs resulted in poor survival of HNSC patients. **(A-O)** Kaplan-meier plot for all the SOCs expression in HNSC patients was downloaded from KMplotter. The probability of survival for high and low expression are shown. Inset- Hazard ratio (HR) with 95% confidence interval and logrank *p* value are mentioned. The number of patients at risk during the time interval is given below the graph.

4% mutation in TRPC5 (20/504), 9% mutation in TRPC6 (43/504), 0.8% mutation in TRPC7 (4/504), 1.8% mutation in TRPV1 (9/504), 1.6% mutation in TRPV2 (8/504), 1.8% mutation in TRPM1 (9/504), and 2.6% mutation in TRPM7 (13/504). This low percentage of mutations indicated that mutation is solely not responsible for the aberration in the SOCs expression (**Figure 2**). Hence, we further analysed the mRNA expression of SOCs. The mRNA expression of SOCs genes across TCGA-HNSC data was downloaded from the TCGA biolinks R package. The data showed significant upregulation of STIM2 ($p < 0.05$), ORAI1 ($p < 0.001$), ORAI2 ($p < 0.001$), ORAI3 ($p < 0.001$), TRPC1 ($p < 0.001$), TRPC3 ($p < 0.001$), TRPC4 ($p < 0.001$), TRPC5 ($p < 0.001$), TRPC6 ($p < 0.001$), TRPV1 ($p < 0.001$), TRPV2 ($p < 0.001$), and TRPM7 ($p < 0.001$) and downregulation of STIM1 ($p < 0.001$) and TRPM1 ($p > 0.05$) were observed across HNSC samples. The expression level of TRPC7 was found to be low in both controls and HNSC patients and was insignificant. We also conducted the stage-wise expression analysis of these genes in HNSC. The data showed significant upregulation of STIM2 (Normal vs. stage I: $p < 0.01$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), ORAI1 (Normal vs. stage I: $p < 0.01$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), ORAI2 (Normal vs. stage I: $p < 0.001$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), ORAI3 (Normal vs. stage I: $p < 0.05$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), TRPC3 (Normal vs. stage I: $p < 0.001$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), TRPC4 (Normal vs. stage I: $p < 0.001$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), TRPC5 (Normal vs. stage I: $p > 0.05$; Normal vs. stage II: $p < 0.01$; Normal vs. stage III: $p < 0.05$; Normal vs. stage IV: $p < 0.001$), TRPC6 (Normal vs. stage I: $p < 0.001$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), TRPV1 (Normal vs. stage I: $p < 0.05$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), TRPV2 (Normal vs. stage I: $p < 0.001$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$), and TRPM7 (Normal vs. stage I: $p < 0.001$; Normal vs. stage II: $p < 0.001$; Normal vs. stage III: $p < 0.001$; Normal vs. stage IV: $p < 0.001$) across all the stages of HNSC. TRPC1 ($p < 0.001$) level was found to be significantly altered in stage IV disease. STIM1, TRPC7, and TRPM1 were found to be non-significant across the stages of HNSC compared to controls (**Figures 3A-AD**). These results indicated that SOCs are differentially

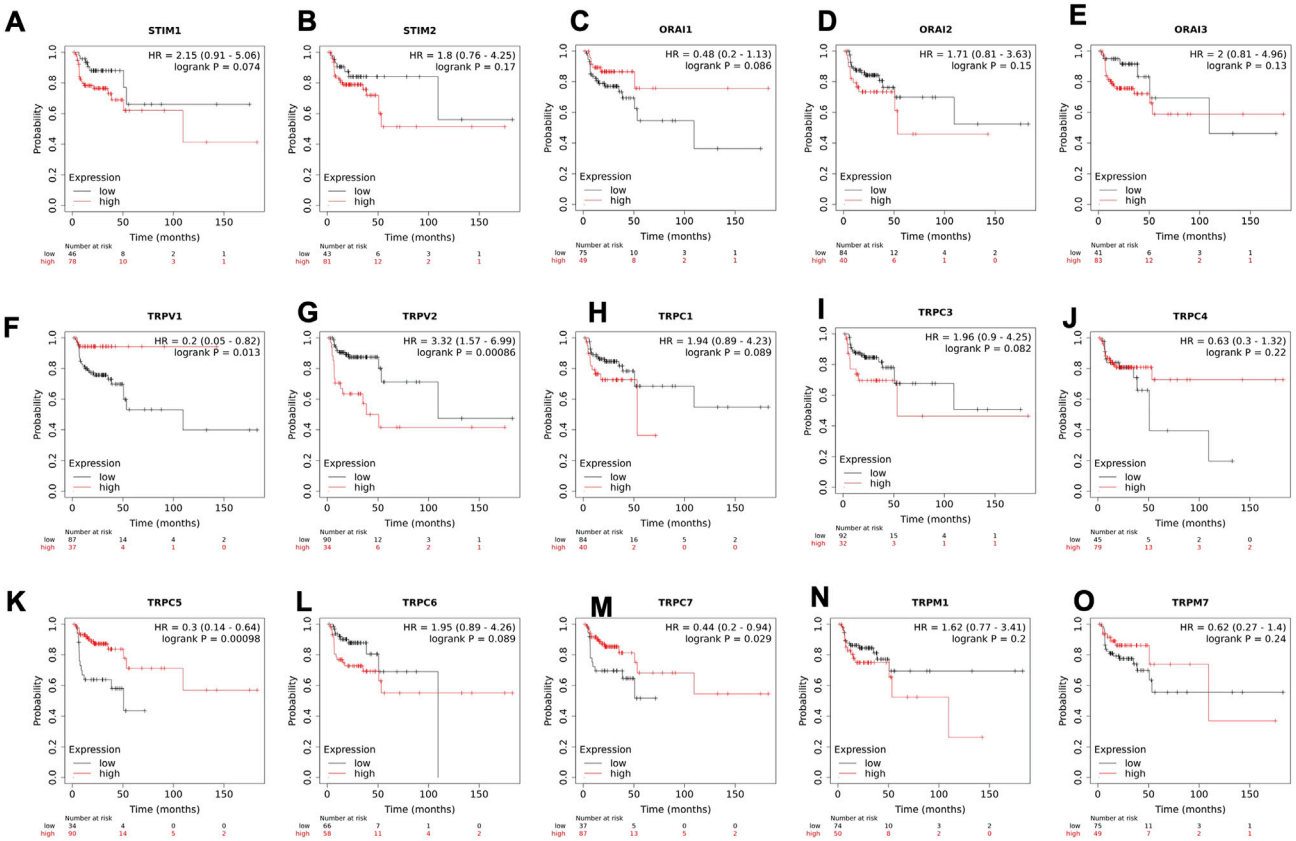


FIGURE 7 | Altered mRNA expression of SOCs lead to poor relapse survival of HNSC patients. **(A-O)** Kaplan-meier plot for relapse free survival of head and neck cancer patients with altered expression of SOCs are shown. The survival probability versus time for high and low expression of SOCs are shown. Inset- Hazard ratio (HR) with 95% confidence interval and logrank *p* value are shown. The number of patients at risk during each time interval for both high and low expression of SOCs is represented below the graph.

expressed in various stages of head and neck cancers and could be a potential biomarker of HNSC.

Correlation of Store-Operated Calcium Channels Gene Expression and Genes Involved in Mitochondrial Dynamics (MDGs)

The expression of SOCs genes and mitochondrial dynamics regulatory genes- DNM1L, FIS1, MFF, MFN1, MFN2, OPA1- across TCGA-HNSC data were conducted using Timer 2.0 webtool. The expression of SOCs genes was found to be remarkably associated with the expression of DNM1L, FIS1, MFN1, MFN2, and OPA1. Similarly, expression of SOCs genes with MDGs in HPV positive and HPV negative HNSC patient samples are also shown (**Figures 4A–P**). Next, we downloaded available protein structures for SOCs (STIM1, STIM2, ORAI1, TRPC5) and mitochondrial fission and fusion regulatory genes (DNM1L, FIS1, MFF, MFN1, MFN2, and OPA1) from the protein data bank. Protein-protein docking using ClustPro and Prodigy showed high negative binding energy for these proteins—STIM1 vs. DNM1L:ΔG = −19.4KCalmol^{-1}, STIM1 vs. FIS1:ΔG = −7.5KCalmol^{-1}, STIM1 vs. MFN1:ΔG = −16.7KCalmol^{-1}, STIM1 vs. MFN2:ΔG = −10.6KCalmol^{-1}, STIM1 vs. OPA1:ΔG = −11.2KCalmol^{-1}; STIM2 vs. DNM1L:ΔG = −17.0KCalmol^{-1}; STIM2 vs. FIS1:ΔG = −15.9KCalmol^{-1}, STIM2 vs. MFN1:ΔG = −9.5KCalmol^{-1}, STIM2 vs. MFN2:ΔG = −11.0KCalmol^{-1}; STIM2 vs. OPA1:ΔG = −11.8KCalmol^{-1}; ORAI1 vs. DNM1L:ΔG = −20.0KCalmol^{-1}, ORAI1 vs. FIS1:ΔG = −10.5KCalmol^{-1}, ORAI1 vs. MFN1:ΔG = −11.7KCalmol^{-1}, ORAI1 vs. MFN2:ΔG = −13.5KCalmol^{-1}, ORAI vs. OPA1:ΔG = −11KCalmol^{-1}; TRPC5 vs. DNM1L:ΔG = −23.4KCalmol^{-1}, TRPC5 vs. FIS1:ΔG = −22.7KCalmol^{-1}, TRPC5 vs. MFN1:ΔG = −23.1KCalmol^{-1}, TRPC5 vs. MFN2:ΔG = −22.8KCalmol^{-1}, TRPC5 vs. OPA1:ΔG = −25.2KCalmol^{-1} indicating higher chances of binding of SOCs with proteins involved in mitochondrial dynamics (**Figures 4Q–AJ**). However, STRING analysis showed no known direct link between SOCs and MDGs (**Figures 5A–C**).

Survival Analysis

To understand the clinical relevance of these SOCs genes, the correlation between gene expression versus overall survival and

FIGURE 8 | Survival analysis among Head neck cancer patients expressing altered levels of SOCs in conjunction with MDGs. The survival data from TCGA was downloaded using TCGA biolinks package built for R statistical environment. The survival data for expression levels of each SOCs in conjunction with MDGs is plotted.

relapse-free survival of HNSC patients were explored. Kaplan-Meier plotter was used to analyze overall survival analysis across TCGA-HNSC samples. The expression of STIM2, ORAI1, TRPV1, TRPV2, TRPC1, TRPC3, TRPC5, TRPC6, TRPC7, and TRPM7 were found to be significantly ($p < 0.05$) associated with HNSC patient survival (**Figures 6A–O**). The effect of expression of SOCs along with MDGs on survival of HNSC was visualized. Similar survival analysis was performed for relapse-free survival across TCGA data sets and are shown in **Figures 7A–O**. Next, the survival data were downloaded using TCGA biolinks via R programming. We analysed the survival probability for patients expressing SOCs in conjunction with the MDGs and found that SOCs along with MDGs are potential diagnostic and prognostic markers of HNSC (**Figure 8**).

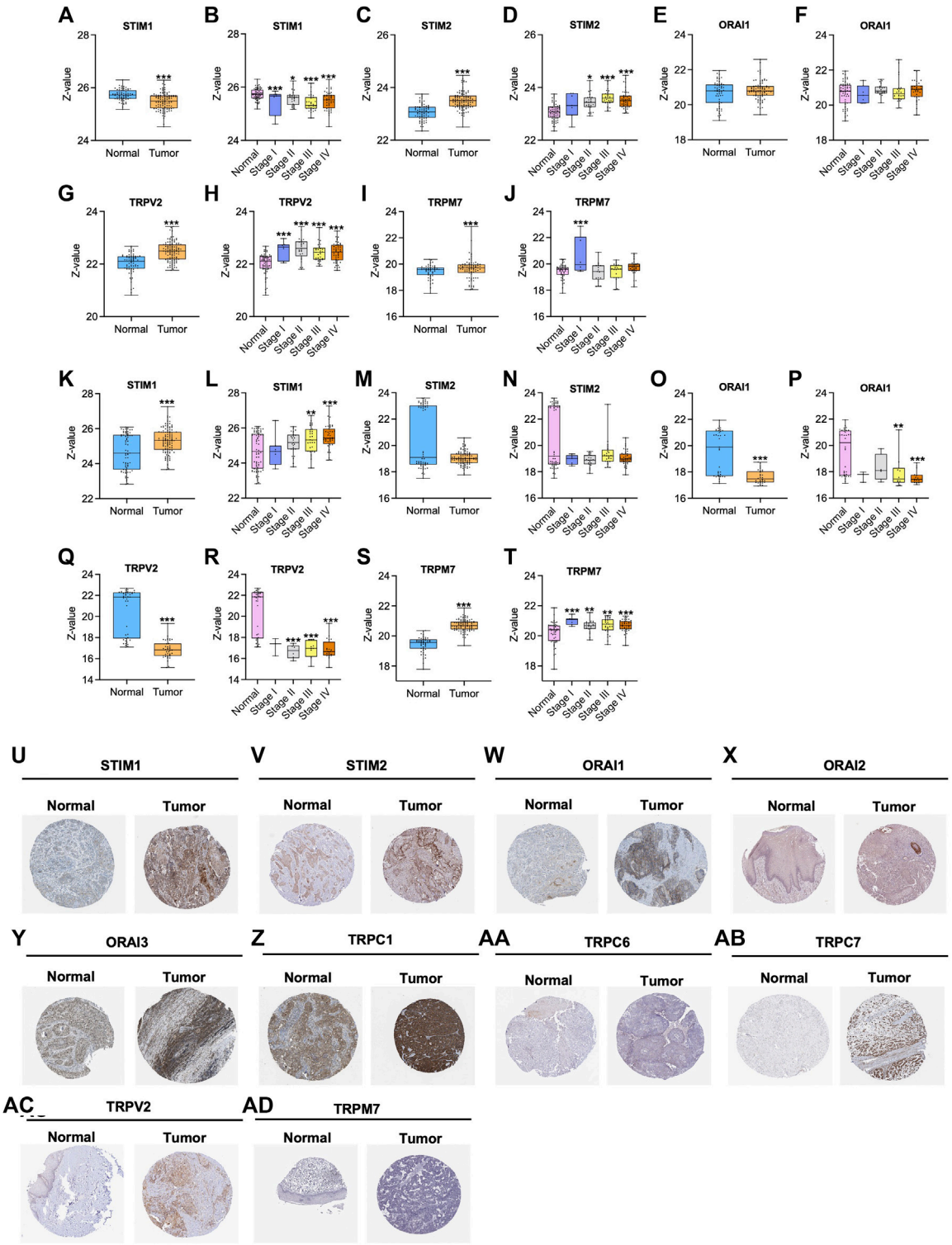


FIGURE 9 | Altered protein expression of SOCs in HNSC. **(A–J)** The proteomics and **(K–T)** phosphoproteomics expression was downloaded from CPTAC-HNSC database using Python 3.0. The protein and phosphoprotein expression of SOCs were plotted. The phosphoprotein sites which are considered to plot this graph include: S257 for STIM1; S261 for STIM2; T295 for ORAI1; S751 for TRPV2; S1477 for TRPM7. Asterisk represents *$p < 0.05$, **<0.01, ***$p < 0.001$ the statistical significance. **(U–AD)** Histopathology slides of tissue microarray were downloaded from Human Protein Atlas and the expression of the SOCs are presented.

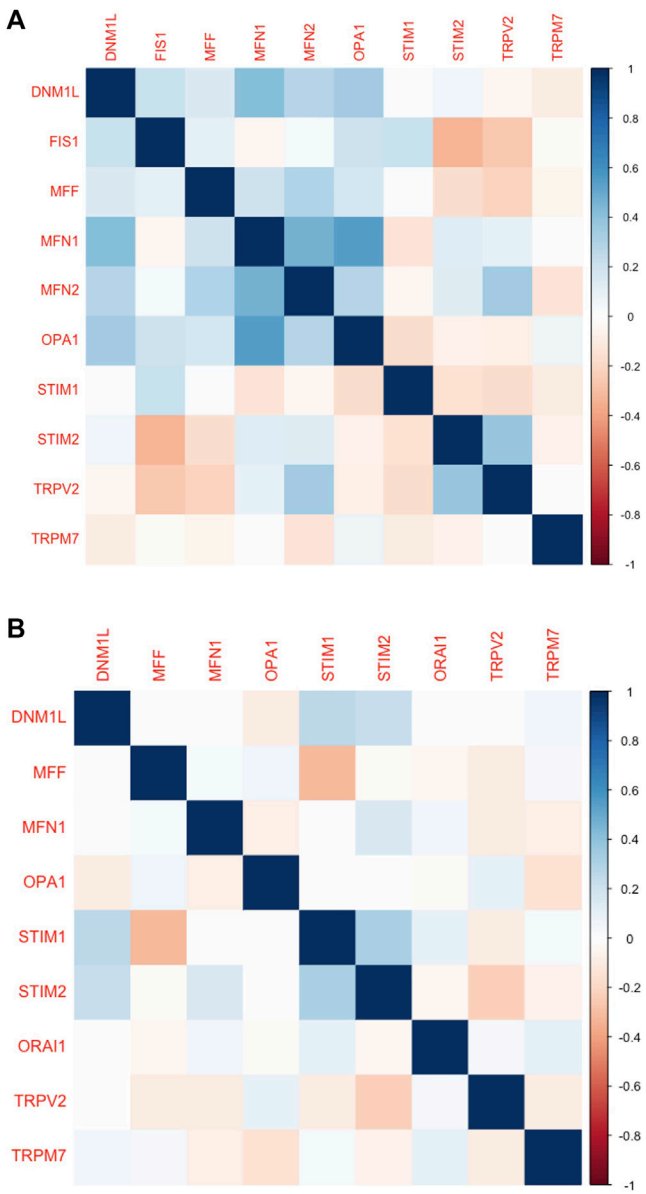


FIGURE 10 | Protein levels of SOCs correlated with the MDGs in HNSC. **(A)** The proteomics and **(B)** phospho-proteomics expression CPTAC-HNSC data was downloaded using Python 3.0. The correlation was calculated and plotted using Corr plot package for R programming. The markings on right hand side indicates the color code for correlation coefficient.

Expression of Store-Operated Calcium Channels and Their Correlation With Mitochondrial Dynamics Among Clinical Proteomic Tumor Analysis Consortium (CPTAC) and Gene Expression Omnibus (GEO) Datasets

Furthermore, we analysed the protein expression of SOCs and MDGs in HNSC. The proteomics, phosphoproteomics, and clinical data were obtained for CPTAC-HNSC using Python version 3.0. The protein expression was available only for STIM1, STIM2, ORAI1, TRPV2, and TRPM7 in CPTAC dataset. This might be due to the spatio-temporal expression of proteins which is unrelated to the mRNA expression of genes. The available data were analyzed and visualized. The protein (**Figures 9A–J**) and phosphoprotein levels (**Figures 9K–T**) showed significant alteration of STIM1, STIM2, ORAI1, TRPV2, and TRPM7 in HNSC patients both overall and stage-wise compared to control samples. The histopathology slides downloaded from the

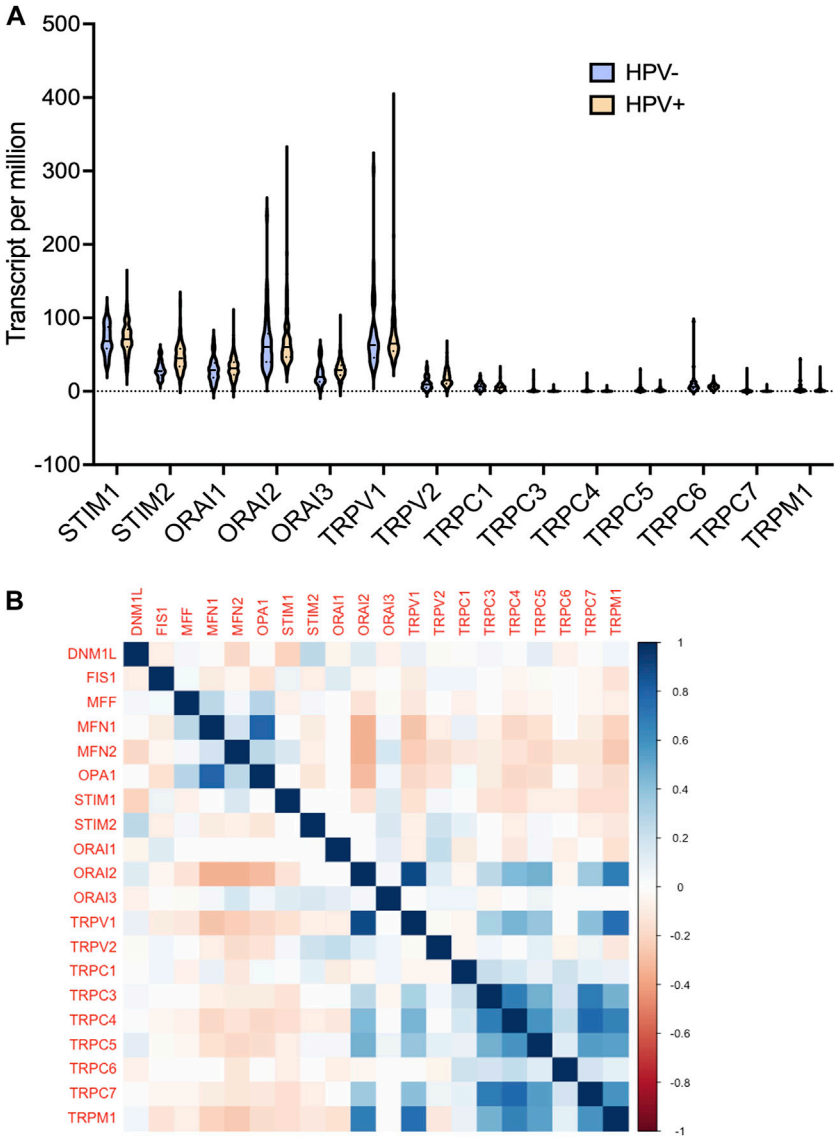


FIGURE 11 | mRNA expression levels of SOCs were correlated with the expression of MDGs in both HPV positive and negative HNSCs. The expression data from GSE17898 was downloaded from GEO datasets consisting total of 323 samples **(A)** The levels of SOCs are plotted using Graphpad prism software and **(B)** correlation analysis was performed and plotted using Corr Plot package for R statistical environment. The gradation on the right side denote the correlation coefficient.

human protein atlas also showed similar results (**Figures 9U–AD**). Additionally, the correlation of SOCs proteins and phosphoproteins with MD proteins were analyzed using Corr Plot package in R statistical environment (**Figures 10A,B**). In addition, the dataset GSE17898 consisting of normalized expression data from a total of 323 HNSC samples were downloaded and analyzed. The data showed remarkable alteration of SOCs in both HPV positive and HPV negative samples (**Figure 11A**). The correlation analysis showed a significant association among SOCs and MDGs expression which is in accordance with our TCGA and CPTAC analysis (**Figure 11B**).

Single-Cell Gene Expression Analysis of Store-Operated Calcium Channels and Their Correlation With Mitochondrial Dynamics

The tumor microenvironment consists of heterogenous population of cells contributing separately to the proliferation, development, metastasis, and therapeutic resistance (Da Silva-Diz et al., 2018; Stanta and Bonin, 2018; El-Sayes et al., 2021). Hence, we further analyzed single cell dataset for HNSC (GSE130922) downloaded from NCBI GEO datasets. Our analysis showed heterogeneous expression of SOCs in different

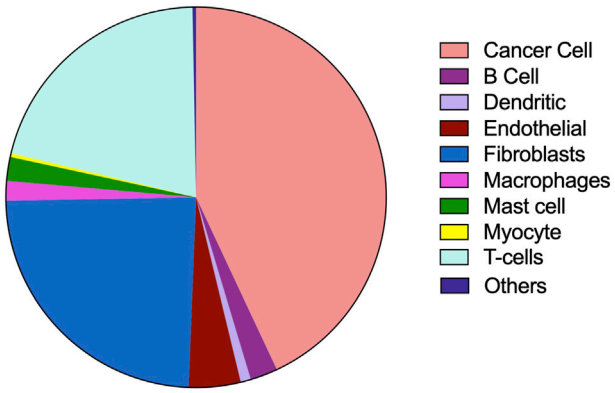


FIGURE 12 | Composition of different cell types in tumor microenvironment analysed using the dataset GSE103322. The percentage composition of cancer cells, B cells, dendritic cells, endothelial cells, fibroblasts, macrophages, mast cells, myocytes, T cells and other cells are visualized using pie-chart representation.

cells of HNSC tumor tissues (Puram et al., 2017). The data was stratified across 10 different cell types- cancer cells and B cells, dendritic cells, endothelial cells, fibroblasts, macrophages, mast cells, monocytes, T-cells, and others among non-cancer stromal cells. The percentage population of these cells in the dataset are shown in **Figure 12**. STIM1, STIM2, ORAI1, ORAI2, ORAI3, TRPV1, TRPV2, and TRPM7 were found to be differentially expressed across all cell types in HNSC tumors (**Figures 13A-O**). Among the immune cells, STIM1 is expressed in all the immune cells with the highest expression in mast cells and lowest in dendritic and B cells. STIM2 and ORAI2 are expressed nearly equally in all the immune cells. ORAI1 levels were found to be nil in B cells. ORAI3 is found to be least expressed in dendritic and T cells. TRPC1, TRPC4, TRPC6, and TRPC7 are almost completely absent in all immune cells types whereas TRPC3 is expressed in T cells and dendritic cells, and TRPC5 is expressed only in T cells. TRPV1, TRPM1, and TRPM7 are almost equally expressed in all types of immune cells. TRPV2 is found to be expressed highly in macrophages. However, expression of ORAI2, TRPC5, and TRPV2 are high in B cells, T cells, and macrophages respectively compared to parenchymal cells. However, other SOCs are enriched in parenchymal cells including cancer cells compared to immune cells. In addition, SOCs were found to be highly expressed in cancer cells and fibroblasts among the parenchymal cells (**Figures 13A-O**). Further, a significant correlation of SOCs with MDGs were observed in cancer cells (**Figure 14**) and fibroblasts (**Figure 15**). The correlation coefficient and *p*-value are represented in the **Table 1**.

These single-cell analysis revealed the comprehensive role of SOCs together with MDGs in the tumor microenvironment.

Overall, our *in silico* approach depicted that SOCs might be involved in the regulation of mitochondrial function in HNSC and the expression of SOCs along with the MDGs can be a predictive marker of HNSC and might have prognostic value in these patients.

DISCUSSION

Dysregulated intracellular Ca^{2+} signaling in cancer cells is shown to be remarkably associated with cancer cell growth, proliferation, angiogenesis, and metastasis (Barr et al., 2008; Chen et al., 2011; Wang et al., 2012; Bergmeier et al., 2013; Motiani et al., 2013; Wang et al., 2022). Dysregulation of SOCE and Ca^{2+} imbalance was reported in Sjögren's syndrome and in head and neck cancers treated with radiation (Cheng et al., 2012; Ambudkar, 2018). Elevated serum calcium levels is a proposed diagnostic marker for head and neck malignancy (Bradley and Hoskin, 2006). Recently, Newton et al. (2020) demonstrated that monoclonal antibody against PD-L1 enhances the functionality of T cells by modulating calcium release-activated calcium channels (Newton et al., 2020). siRNA-mediated knockdown of ORAI1 and STIM1 in Ca9-22 and OECM-1 oral cancer cell lines showed reduced proliferation, migration, and invasion of these cells (Wang et al., 2022). In the current study, we revealed that SOCs might be involved in regular mitochondrial function, and alteration in these might be a predictive and prognostic marker.

Substantial evidence has been provided in several different types of cells that SOCs are involved in cell interaction and secretory Ca-ATPase-2 pathway (Bergmeier et al., 2013). In addition, recent evidence suggests that the formation of a complex of these proteins with phosphatase calcineurin dephosphorylates cytoplasmic NFAT and induces nuclear translocation. Nuclear NFAT transcriptionally activates the expression of several genes including NANOG, OCT4, SOX2, and FGF19 which are involved in cancer cell stemness (Wang et al., 2021). In the current study, we conducted the gene ontology analysis, as a preliminary analysis to show the involvement of SOCs in the regulation of IP3 and ATPase pathways. Due to the involvement of SOCs in signaling pathways apart from their regular transport activities we hypothesized that SOCs might be involved even in the regulation of mitochondrial activities. Furthermore, disease ontology analysis was conducted to show the involvement of SOCs in several cancers.

Earlier studies have shown that the entry of Ca^{2+} ions through store-operated channels begins with the stimulation of plasma membrane receptors to phospholipase C and synthesis of inositol triphosphate. Activated SOCE aid in refilling Ca^{2+} stores for further stimulation (Putney, 1986; Parekh and Putney, 2005). In the current study, disease ontology analysis showed their involvement in inflammatory diseases, eye tumors, and medulloblastomas. This is in accordance with earlier *in vitro* studies (Wang et al., 2022). Subsequently, the TCGA-HNSC mRNA expression of SOCs showed that STIMs, ORAIs, TRPCs, TRPVs, and TRPMs are significantly altered in HNSC patients. The survival analysis clearly demonstrated that alteration in SOCs mRNA expression remarkably decreases the survival rate of these patients. Further, TCGA-HNSC based correlation analysis using Timer 2.0 tool showed a significant correlation in the expression of MDGs with SOCs. Mitochondrial regulation of SOCE is due to the

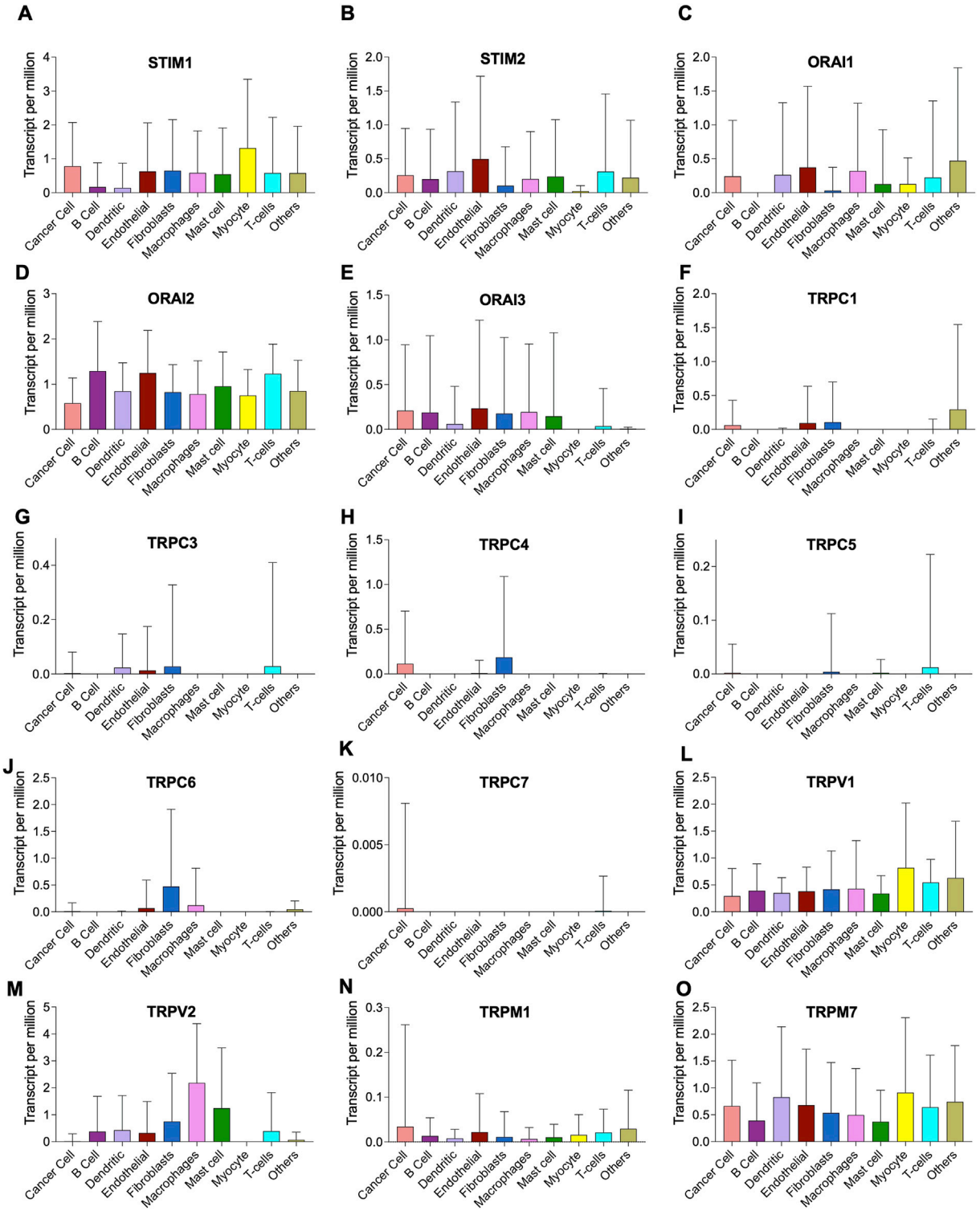


FIGURE 13 | Heterogenous expression of SOCs in tumor microenvironment. **(A-O)** The expression data from GSE103322 comprising the data of 5,902 cells from 18 head and neck cancer tissues was downloaded from NCBI GEO website. The data comprised of different cell populations including cancer cells, B cells, dendritic cells, endothelial cells, fibroblasts, macrophages, mast cells, myocytes, T cells and other cells The expression levels of each SOCs in these different cell types are plotted using Graphpad prism software version 9.2.1.

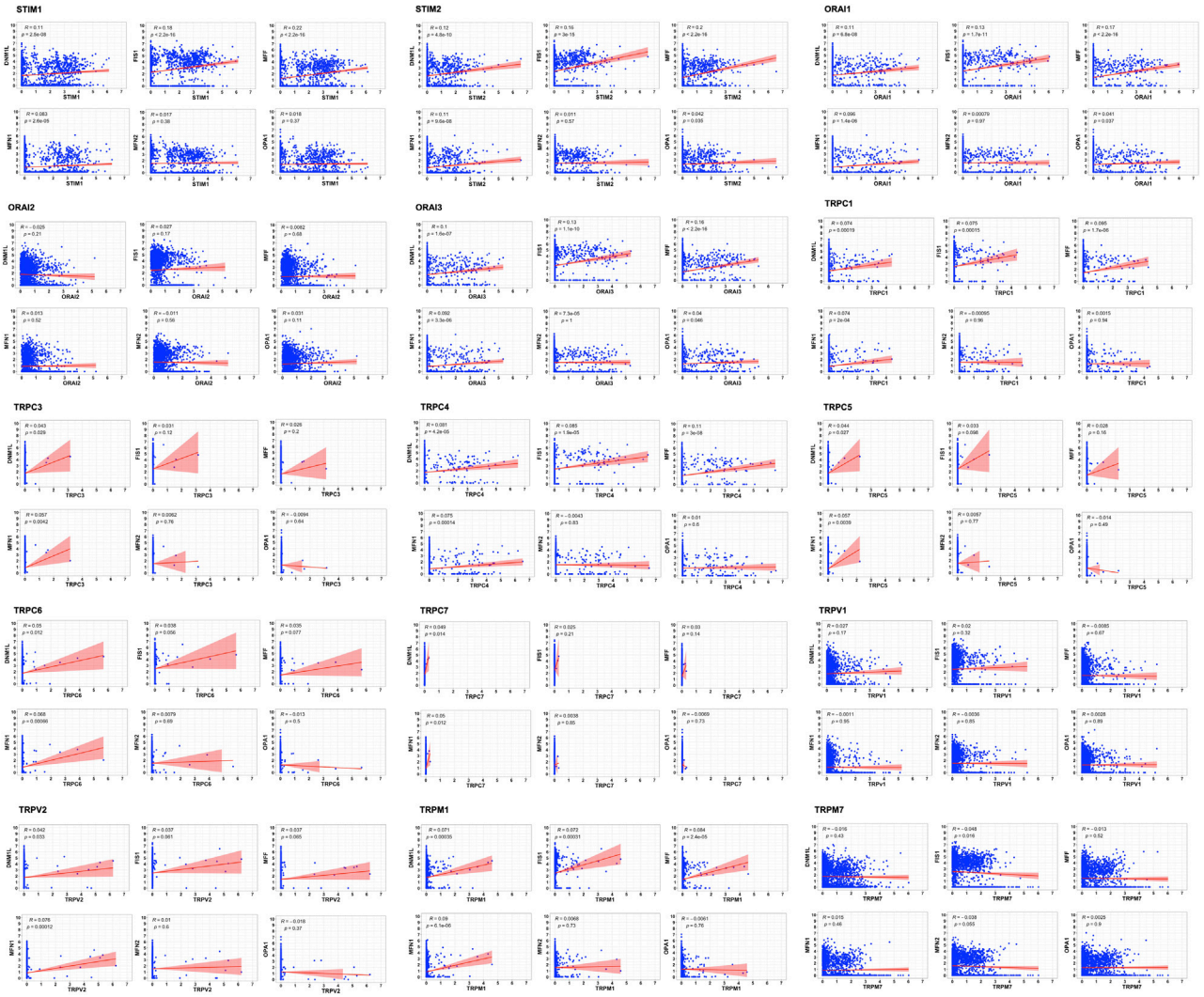


FIGURE 14 | Correlation of SOCs expression with MDGs in cancer cells across dataset GSE103322. The correlation between each SOCs expression (STIM1, STIM2, ORAI1, ORAI2, ORAI3, TRPC1, TRPC3, TRPC4, TRPC5, TRPC6, TRPC7, TRPV1, TRPV2, TRPM1, and TRPM7) versus MDGs expression (DNM1L, FIS1, MFF, MFN1, MFN2, and OPA1) among the cancer cells are plotted using Corr plot package of R statistical environment. The linear regression line per group is represented as red colored line. Inset- R value (correlation coefficient) and *p* value are mentioned.

ability to rapidly uptake the Ca^{2+} thus modulating the inositol phosphate mediated signaling. Hoth et al. (1997) showed that mitochondrial uncouplers inhibited Ca^{2+} exit from mitochondria leading to the prevention of sustained entry of Ca^{2+} into T-cells (Hoth et al., 1997). Further, Gilabert and Parekh revealed respiring mitochondria are required for Ca^{2+} homeostasis by CRAC channels (Gilabert and Parekh, 2000). In addition, FDA-approved drugs, leflunomide and teriflunomide were shown to be inhibitors of SOCs at clinically-relevant doses in neuroblastoma cells (Rahman and Rahman, 2017). Miret-Casals et al. (2018) demonstrated that leflunomide and teriflunomide induces mitochondrial fusion through the activation of MFN2 in cervical cancer cell lines (Miret-Casals et al., 2018). Leflunomide has also been shown to promote mitochondrial fusion via downregulating total and phospho DNM1L and inducing MFN2 leading to growth retardation in pancreatic adenocarcinoma cells (Yu et al., 2019). Recently, Yedida et al. (2019) demonstrated that apogossypol (a small molecule inhibitor of pan-Bcl2) mediated endoplasmic reticulum (ER) reorganization results in Ca^{2+} transfer between ER and mitochondria leading to inhibition of mitochondrial fission and apoptosis of HeLa cells (Yedida et al., 2019). This study also showed that leflunomide, a potent inhibitor of SOCs, inhibits apogossypol-mediated ER reorganization and antagonizes its protective effect against apoptosis (Yedida et al., 2019). In accordance with these studies, our docking results showed higher negative binding energy for SOCs with MDGs. To our knowledge, there are no *in vitro* or *in vivo*

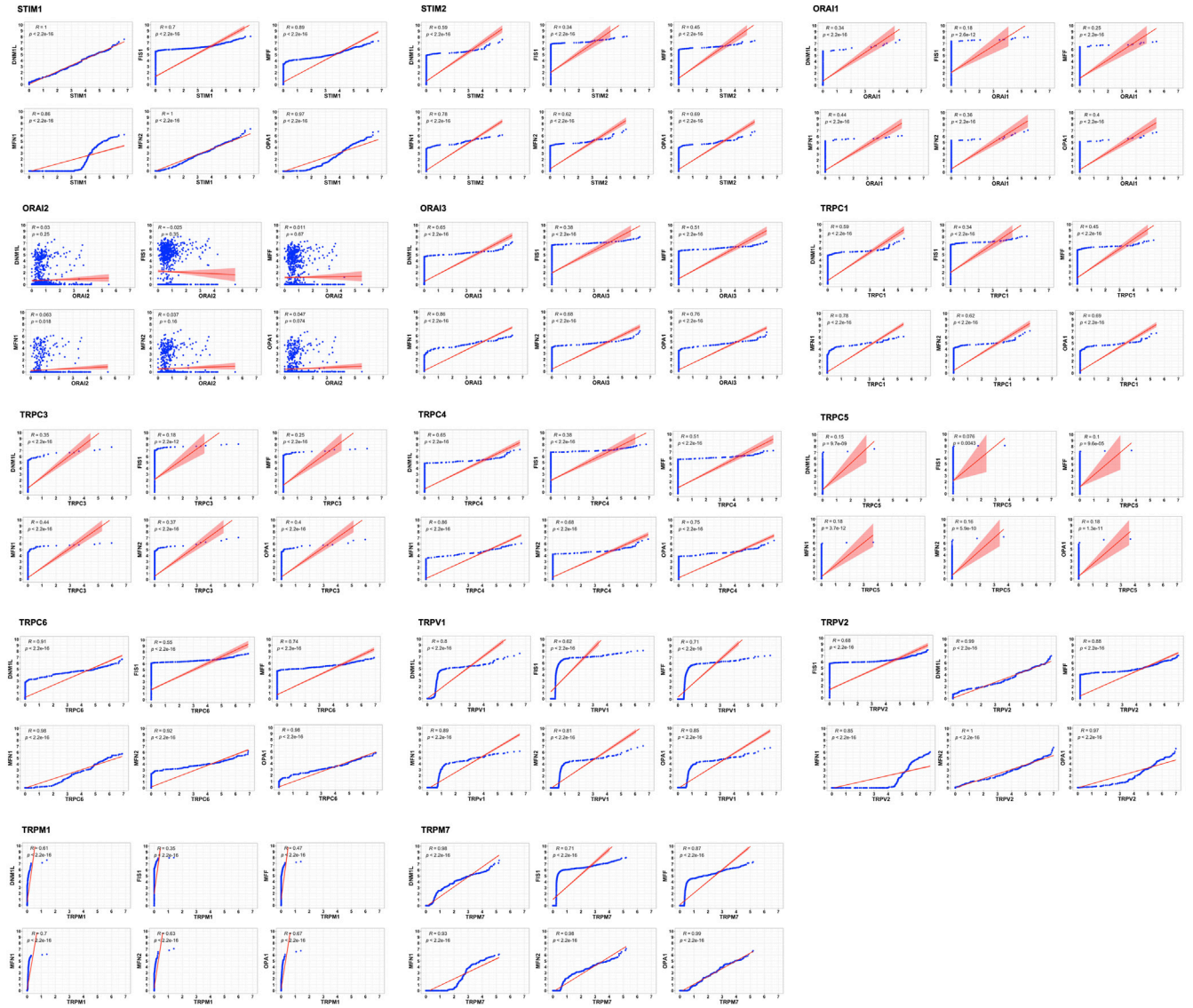


FIGURE 15 | Correlation of SOCs expression with MDGs in fibroblasts across dataset GSE103322. The expression of each SOCs (STIM1, STIM2, ORAI1, ORAI2, ORAI3, TRPC1, TRPC3, TRPC4, TRPC5, TRPC6, TRPV1, TRPV2, TRPM1, and TRPM7) versus expression of each MDGs (DNM1L, FIS1, MFF, MFN1, MFN2, and OPA1) were compared among the fibroblasts and the correlation plots are plotted using Corr plot package of R statistical environment. TRPC7 is not expressed in the fibroblasts and is omitted. Linear regression line is shown in red color. Inset- R value (correlation coefficient) and *p* value are shown.

studies showing the direct or indirect binding of SOCs with mitochondrial proteins. The STRING database also showed no known direct relation among these proteins. In the current study, we analyzed phosphoproteomic data for head and neck cancer from CPTAC. Mertins et al., 2016 integrated proteomics and phosphoproteomics data of CPTAC to identify distinct profiles in 77 genomically annotated breast tumors. In this study, the authors also revealed changes in phosphoproteomics of CDK12, PAK1, RIPK2, and TLK2. This study also proposed the phosphoproteomic changes in these proteins can be utilized as druggable kinases beyond HER2 (Mertins et al., 2016). Hence, we conducted the protein and phosphoprotein analysis for clinical samples to show SOC proteins are also valuable biomarkers alongside the mRNA expression in head and neck cancers.

Among the post-translational modifications, phosphorylation of the protein is central to signaling mechanisms and is critical for various physiological responses. There are about 50,000 known phosphorylation sites that do not currently have any ascribed functions (Mayya and Han, 2009). However, quantitative phosphoproteomics has been an effective tool to identify functional phosphorylation sites and putative substrates of kinases. Immunoprecipitation followed by phosphoproteomics using mass spectrometry demonstrated

TABLE 1 | The correlation coefficient and *p*-value for SOCs and MDGs in cancer cells and fibroblasts.

Cell type	SOCs	MDGs	Correlation coefficient	*p*-Value
Cancer cell	STIM1	DNM1L	0.11	2.5×10^{-8}
		FIS1	0.18	$<2.2 \times 10^{-16}$
		MFF	0.22	$<2.2 \times 10^{-16}$
		MFN1	0.083	2.6×10^{-5}
		MFN2	0.017	0.38
		OPA1	0.018	0.37
	STIM2	DNM1L	0.12	4.8×10^{-10}
		FIS1	0.16	3×10^{-15}
		MFF	0.2	$<2.2 \times 10^{-16}$
		MFN1	0.11	9.6×10^{-8}
		MFN2	0.011	0.57
		OPA1	0.042	0.035
	ORAI1	DNM1L	0.11	6.8×10^{-8}
		FIS1	0.13	1.7×10^{-11}
		MFF	0.17	$<2.2 \times 10^{-16}$
		MFN1	0.096	1.4×10^{-6}
		MFN2	0.00079	0.97
		OPA1	0.041	0.037
	ORAI2	DNM1L	–0.025	0.21
		FIS1	0.027	0.17
		MFF	0.0082	0.68
		MFN1	0.013	0.52
		MFN2	-0.011	0.56
		OPA1	0.031	0.11
	ORAI3	DNM1L	0.1	1.6×10^{-7}
		FIS1	0.13	1.1×10^{-10}
		MFF	0.16	2.2×10^{-16}
		MFN1	0.092	3.3×10^{-6}
		MFN2	7.3×10^{-5}	1
		OPA1	0.04	0.046
	TRPC1	DNM1L	0.074	0.00019
		FIS1	0.075	0.00015
		MFF	0.095	1.7×10^{-6}
		MFN1	0.074	2×10^{-4}
		MFN2	–0.00095	0.96
		OPA1	0.0015	0.94
	TRPC3	DNM1L	0.043	0.029
		FIS1	0.031	0.12
		MFF	0.026	0.2
		MFN1	0.057	0.0042
		MFN2	0.0062	0.76
		OPA1	–0.0094	0.64
	TRPC4	DNM1L	0.081	4.2×10^{-5}
		FIS1	0.085	1.9×10^{-5}
		MFF	0.11	3×10^{-8}
		MFN1	0.075	0.00014
		MFN2	–0.0043	0.83
		OPA1	0.01	0.6
	TRPC5	DNM1L	0.044	0.027
		FIS1	0.033	0.098
		MFF	0.028	1.6
		MFN1	0.057	0.0039
		MFN2	0.0057	0.77
		OPA1	–0.014	0.49
	TRPC6	DNM1L	0.05	0.012
		FIS1	0.038	0.056
		MFF	0.035	0.077
		MFN1	0.068	0.00066
		MFN2	0.0079	0.69
		OPA1	–0.013	0.5
	TRPC7	DNM1L	0.049	0.14
		FIS1	0.025	0.21
		MFF	0.03	0.14
		MFN1	0.05	0.012
		MFN2	0.0038	0.85
		OPA1	–0.0069	0.73
	TRPV1	DNM1L	0.027	0.17
		FIS1	0.02	0.32
		MFF	–0.0085	0.67
		MFN1	–0.0011	0.95
		MFN2	–0.0036	0.85
		OPA1	0.0028	0.89
	TRPV2	DNM1L	0.042	0.033
		FIS1	0.031	0.06
		MFF	0.037	0.065
		MFN1	0.076	0.00012
		MFN2	0.01	0.6
		OPA1	–0.018	0.37
	TRPM1	DNM1L	0.071	0.00035
		FIS1	0.072	0.00031
		MFF	0.084	2.4×10^{-5}
		MFN1	0.09	6.1×10^{-6}
		MFN2	–0.0068	0.73
		OPA1	–0.0061	0.76
	TRPM7	DNM1L	–0.016	0.43
		FIS1	–0.048	0.016
		MFF	–0.013	0.52
		MFN1	0.015	0.46
		MFN2	–0.038	0.055
		OPA1	0.0025	0.9
Fibroblasts	STIM1	DNM1L	1	$<2.2 \times 10^{-16}$
		FIS1	0.7	$<2.2 \times 10^{-16}$
		MFF	0.89	$<2.2 \times 10^{-16}$
		MFN1	0.86	$<2.2 \times 10^{-16}$
		MFN2	1	$<2.2 \times 10^{-16}$
		OPA1	0.97	$<2.2 \times 10^{-16}$
	STIM2	DNM1L	0.59	$<2.2 \times 10^{-16}$
		FIS1	0.34	$<2.2 \times 10^{-16}$
		MFF	0.45	$<2.2 \times 10^{-16}$
		MFN1	0.78	$<2.2 \times 10^{-16}$
		MFN2	0.62	$<2.2 \times 10^{-16}$
		OPA1	0.69	$<2.2 \times 10^{-16}$
	ORAI1	DNM1L	0.34	$<2.2 \times 10^{-16}$
		FIS1	0.18	$=2.2 \times 10^{-16}$
		MFF	0.25	$<2.2 \times 10^{-16}$
		MFN1	0.44	$<2.2 \times 10^{-16}$
		MFN2	0.36	$<2.2 \times 10^{-16}$
		OPA1	0.4	$<2.2 \times 10^{-16}$
	ORAI2	DNM1L	0.03	0.25
		FIS1	-0.025	0.35
		MFF	0.011	0.67
		MFN1	0.063	0.018
		MFN2	0.037	0.16
		OPA1	0.047	0.074
	ORAI3	DNM1L	0.65	$<2.2 \times 10^{-16}$
		FIS1	0.38	$<2.2 \times 10^{-16}$
		MFF	0.51	$<2.2 \times 10^{-16}$
		MFN1	0.86	$<2.2 \times 10^{-16}$
		MFN2	0.68	$<2.2 \times 10^{-16}$
		OPA1	0.76	$<2.2 \times 10^{-16}$
	TRPC1	DNM1L	0.59	$<2.2 \times 10^{-16}$
		FIS1	0.34	$<2.2 \times 10^{-16}$
		MFF	0.45	$<2.2 \times 10^{-16}$
		MFN1	0.78	$<2.2 \times 10^{-16}$
		MFN2	0.62	$<2.2 \times 10^{-16}$

(Continued in next column)

(Continued on following page)

TABLE 1 | (*Continued*) The correlation coefficient and *p*-value for SOCs and MDGs in cancer cells and fibroblasts.

Cell type	SOCs	MDGs	Correlation coefficient	*p*-Value
		OPA1	0.69	$<2.2 \times 10^{-16}$
	TRPC3	DNM1L	0.35	$<2.2 \times 10^{-16}$
		FIS1	0.18	$<2.2 \times 10^{-16}$
		MFF	0.25	$<2.2 \times 10^{-16}$
		MFN1	0.44	$<2.2 \times 10^{-16}$
		MFN2	0.37	$<2.2 \times 10^{-16}$
		OPA1	0.4	$<2.2 \times 10^{-16}$
	TRPC4	DNM1L	0.65	$<2.2 \times 10^{-16}$
		FIS1	0.38	$<2.2 \times 10^{-16}$
		MFF	0.51	$<2.2 \times 10^{-16}$
		MFN1	0.86	$<2.2 \times 10^{-16}$
		MFN2	0.68	$<2.2 \times 10^{-16}$
		OPA1	0.75	$<2.2 \times 10^{-16}$
	TRPC5	DNM1L	0.15	9.7×10^{-9}
		FIS1	0.076	0.0043
		MFF	0.1	$<2.2 \times 10^{-16}$
		MFN1	0.18	3.7×10^{-12}
		MFN2	0.16	5.9×10^{-10}
		OPA1	0.18	1.3×10^{-11}
	TRPC6	DNM1L	0.91	$<2.2 \times 10^{-16}$
		FIS1	0.55	$<2.2 \times 10^{-16}$
		MFF	0.74	$<2.2 \times 10^{-16}$
		MFN1	0.98	$<2.2 \times 10^{-16}$
		MFN2	0.92	$<2.2 \times 10^{-16}$
		OPA1	0.98	$<2.2 \times 10^{-16}$
	TRPV1	DNM1L	0.8	$<2.2 \times 10^{-16}$
		FIS1	0.62	$<2.2 \times 10^{-16}$
		MFF	0.71	$<2.2 \times 10^{-16}$
		MFN1	0.89	$<2.2 \times 10^{-16}$
		MFN2	0.81	$<2.2 \times 10^{-16}$
		OPA1	0.85	$<2.2 \times 10^{-16}$
	TRPV2	DNM1L	0.68	$<2.2 \times 10^{-16}$
		FIS1	0.99	$<2.2 \times 10^{-16}$
		MFF	0.88	$<2.2 \times 10^{-16}$
		MFN1	0.85	$<2.2 \times 10^{-16}$
		MFN2	1	$<2.2 \times 10^{-16}$
		OPA1	0.97	$<2.2 \times 10^{-16}$
	TRPM1	DNM1L	0.61	$<2.2 \times 10^{-16}$
		FIS1	0.35	$<2.2 \times 10^{-16}$
		MFF	0.47	$<2.2 \times 10^{-16}$
		MFN1	0.7	$<2.2 \times 10^{-16}$
		MFN2	0.63	$<2.2 \times 10^{-16}$
		OPA1	0.67	$<2.2 \times 10^{-16}$
	TRPM7	DNM1L	0.98	$<2.2 \times 10^{-16}$
		FIS1	0.71	$<2.2 \times 10^{-16}$
		MFF	0.87	$<2.2 \times 10^{-16}$
		MFN1	0.93	$<2.2 \times 10^{-16}$
		MFN2	0.98	$<2.2 \times 10^{-16}$
		OPA1	0.99	$<2.2 \times 10^{-16}$

that RTKs including ALK, ROS fusion proteins, PDGFRα, and DDR were found to be highly phosphorylated in non-small cell lung carcinoma cell lines and tumor samples (Rikova et al., 2007). Recently, phosphotyrosine directed mass spectrometry analysis conducted by Van Linde et al. (2022) showed the complex kinase activities in glioblastoma. The study suggested the potential of phosphoproteomic analysis for the identification of targets for the treatment modalities (Van Linde et al., 2022). In the current study, analysis of the protein and phosphoprotein expression of SOCs were crucial to show that they might be involved in the regulation of mitochondrial dynamic changes. Also, high levels of phosphoprotein expression of SOCs indicate their activity in head and neck cancers. Further, the CPTAC-HNSC dataset revealed a significant correlation in the expression of SOCs with MDGs. The survival correlation analysis indicated that the patients expressing altered levels of SOCs along with MDGs possess less survival probability compared with those who are expressing normal levels of either SOCs or MDGs. This indicated that the evaluation of SOCs combined with MDGs might be a potential biomarker in HNSCs.

Furthermore, single-cell analysis showed heterogeneity in the expression of SOCs across the ten different cell types. The co-culture of mouse embryonic fibroblasts with MDA-MB-231 by Yang and Huang (2005) showed the critical role of Ca^{2+} ions influx in fibroblast cells migration (Yang and Huang, 2005). In addition, Davis et al. (2012) demonstrated that altered function of ORAI1 and TRPC1 led to EGF-induced EMT changes in triple-negative breast cancer cell lines (Davis et al., 2012). More recently, Zhang et al. (2020) revealed serum- and glucocorticoid-inducible kinase 1 (SGK1) regulates osteoclastogenesis via controlling ORAI1 leading to bone metastasis of breast cancer both in *in vitro* and *in vivo* models (Zhang et al., 2020). In accordance with these results, single-cell analysis described that the expression of SOCs were found to be strongly correlated with the expression of MDGs in cancer cells and fibroblasts indicating the role of SOCs in conjunction with mitochondrial dysfunction in driving the cancer cell hallmarks. Additionally, our analysis showed the critical role of tumor heterogeneity in the progression of cancer and highlights the importance of developing targeted therapy for the microenvironment niche. However, further studies need to be conducted to validate the role of differential expression of SOC in the different cell populations of the tumor tissue.

Taken together, these results indicated that SOCs and MDGs combined alteration might be a potential diagnostic and prognostic marker in HNSC.

CONCLUSION

Our *in silico* analysis shows the altered mRNA and protein expression of SOCs in head and neck cancer and suggests their role as possible biomarkers. We also showed a strong correlation in the expression of MDGs with SOCs in TCGA-HNSC, CPTAC-HNSC, GSE171898, and GSE103322 datasets. We showed for the first time that the SOCs binds to MDGs with very high efficiency. Mechanistic studies need to be conducted further to validate their role in mitochondrial dysfunction and in HNSC development. Based on our *in silico* studies we propose that the expression of SOCs along with MDGs might serve as a better early diagnostic and prognostic marker in HNSC patients. However, further studies need to be conducted to evaluate their potential use in clinical diagnosis and management.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: NCBI Gene Expression Omnibus.

AUTHOR CONTRIBUTIONS

MH: Conceptualization, Methodology, Data Curation, Formal analysis, Investigation, Manuscript preparation, Visualization, UD: Confirmation of data, Editing, Visualization, SJ: Overall review and editing, AS; Overall review and editing AC: Methodology, Software, Resources, SA: Methodology, Resources, Revision, MS: Methodology, Resources, Revision, AK: Conceptualization, Supervision, Proof reading, Project administration, Funding acquisition.

FUNDING

This project was supported by BT/556/NE/U-Excel/2016 grant awarded to AK by Department of Biotechnology (DBT), Government of India. This work was also supported by the Researchers Supporting Project number (RSP-2022/5) King Saud University, Riyadh, Saudi Arabia.

ACKNOWLEDGMENTS

MH acknowledges Science and Engineering Board (SERB)-National Post-Doctoral Fellowship (N-PDF) (PDF/2021/004053) for the financial support. UD acknowledges Prime Minister's Research Fellowship (PMRF) program, Ministry of Education (MoE), Govt. of India for providing her the prestigious fellowship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.866473/full#supplementary-material

Supplementary Figure 1 | Graphical Abstract.

REFERENCE

Alptekin, M., Eroglu, S., Tutar, E., Sencan, S., Geyik, M. A., Ulasli, M., et al. (2015). Gene Expressions of TRP Channels in Glioblastoma Multiforme and Relation with Survival. *Tumor Biol.* 36, 9209–9213. doi:10.1007/s13277-015-3577-x

Ambudkar, I. (2018). Calcium Signaling Defects Underlying Salivary Gland Dysfunction. *Biochimica Biophysica Acta (BBA) - Mol. Cell. Res.* 1865, 1771–1777. doi:10.1016/j.bbamcr.2018.07.002

Authi, K. S. (2007). TRP Channels in Platelet Function. *Handb. Exp. Pharmacol.* 179, 425–443. doi:10.1007/978-3-540-34891-7_25

Barr, V. A., Bernot, K. M., Srikanth, S., Gwack, Y., Balagopalan, L., Regan, C. K., et al. (2008). Dynamic Movement of the Calcium Sensor STIM1 and the Calcium Channel Orai1 in Activated T-Cells: Puncta and Distal Caps. *MBoC* 19, 2802–2817. doi:10.1091/mbc.e08-02-0146

Bastián-Eugenio, C. E., Bohórquez-Hernández, A., Pacheco, J., Sampieri, A., Asanov, A., Ocelotl-Oviedo, J. P., et al. (2019). Heterologous Calcium-dependent Inactivation of Orai1 by Neighboring TRPV1 Channels Modulates Cell Migration and Wound Healing. *Commun. Biol.* 2, 88. doi:10.1038/s42003-019-0338-1

Bergmeier, W., Weidinger, C., Zee, I., and Feske, S. (2013). Emerging Roles of Store-Operated Ca2+entry through STIM and ORAI Proteins in Immunity, Hemostasis and Cancer. *Channels* 7, 379–391. doi:10.4161/chan.24302

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242. doi:10.1093/nar/28.1.235

Berridge, M. J., Bootman, M. D., and Roderick, H. L. (2003). Calcium Signalling: Dynamics, Homeostasis and Remodelling. *Nat. Rev. Mol. Cell. Biol.* 4, 517–529. doi:10.1038/nrm1155

Bradley, P. J., and Hoskin, D. (2006). Hyercalcaemia in Head and Neck Squamous Cell Carcinoma. *Curr. Opin. Otolaryngol. Head. Neck Surg.* 14, 51–54. doi:10.1097/01.moo.0000193176.54450.c4

Cao, Y.-L., Meng, S., Chen, Y., Feng, J.-X., Gu, D.-D., Yu, B., et al. (2017). MFN1 Structures Reveal Nucleotide-Triggered Dimerization Critical for Mitochondrial Fusion. *Nature* 542, 372–376. doi:10.1038/nature21077

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discov.* 2, 401–404. doi:10.1158/2159-8290.cd-12-0095

Chang, Y., Roy, S., and Pan, Z. (2021). Store-Operated Calcium Channels as Drug Target in Gastroesophageal Cancers. *Front. Pharmacol.* 12, 668730. doi:10.3389/fphar.2021.668730

Chen, Y.-F., Chiu, W.-T., Chen, Y.-T., Lin, P.-Y., Huang, H.-J., Chou, C.-Y., et al. (2011). Calcium Store Sensor Stromal-Interaction Molecule 1-dependent Signaling Plays an Important Role in Cervical Cancer Growth, Migration, and Angiogenesis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15225–15230. doi:10.1073/pnas.1103315108

Cheng, K. T., Alevizos, I., Liu, X., Swaim, W. D., Yin, H., Feske, S., et al. (2012). STIM1 and STIM2 Protein Deficiency in T Lymphocytes Underlies Development of the Exocrine Gland Autoimmune Disease, Sjögren's Syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14544–14549. doi:10.1073/pnas.1207354109

Da Silva-Diz, V., Lorenzo-Sanz, L., Bernat-Peguera, A., Lopez-Cerda, M., and Muñoz, P. (2018). Cancer Cell Plasticity: Impact on Tumor Progression and Therapy Response. *Seminars Cancer Biol.* 53, 48–58. doi:10.1016/j.semcancer.2018.08.009

Davis, F. M., Peters, A. A., Grice, D. M., Cabot, P. J., Parat, M.-O., Roberts-Thomson, S. J., et al. (2012). Non-stimulated, Agonist-Stimulated and Store-Operated Ca2+ Influx in MDA-MB-468 Breast Cancer Cells and the Effect of EGF-Induced EMT on Calcium Entry. *PLoS One* 7, e36923. doi:10.1371/journal.pone.0036923

Desta, I. T., Porter, K. A., Xia, B., Kozakov, D., and Vajda, S. (2020). Performance and its Limits in Rigid Body Protein-Protein Docking. *Structure* 28, 1071–1081. doi:10.1016/j.str.2020.06.006

El-Sayes, N., Vito, A., and Mossman, K. (2021). Tumor Heterogeneity: A Great Barrier in the Age of Cancer Immunotherapy. *Cancers (Basel)* 13, 806. doi:10.3390/cancers13040806

Feske, S. (2019). CRAC Channels and Disease - from Human CRAC Channelopathies and Animal Models to Novel Drugs. *Cell. Calcium* 80, 112–116. doi:10.1016/j.ceca.2019.03.004

Gilabert, J. A., and Parekh, A. B. (2000). Respiring Mitochondria Determine the Pattern of Activation and Inactivation of the Store-Operated Ca(2+) Current I(CRAC). *Embo J.* 19, 6401–6407. doi:10.1093/emboj/19.23.6401

Gualdani, R., De Clippele, M., Ratbi, I., Gailly, P., and Tajeddine, N. (2019). Store-Operated Calcium Entry Contributes to Cisplatin-Induced Cell Death in Non-small Cell Lung Carcinoma. *Cancers (Basel)* 11, 430. doi:10.3390/cancers11030430

Hernández-Morales, M., Sobradillo, D., Valero, R. A., Muñoz, E., Ubierna, D., Moyer, M. P., et al. (2017). Mitochondria Sustain Store-Operated Currents in Colon Cancer Cells but Not in Normal Colonic Cells: Reversal by Non-steroidal Anti-inflammatory Drugs. *Oncotarget* 8, 55332–55352. doi:10.18632/oncotarget.19430

Hoth, M., Fanger, C. M., and Lewis, R. S. (1997). Mitochondrial Regulation of Store-Operated Calcium Signaling in T Lymphocytes. *J. Cell. Biol.* 137, 633–648. doi:10.1083/jcb.137.3.633

Huang, C., Chen, L., Savage, S. R., Eguez, R. V., Dou, Y., Li, Y., et al. (2021). Proteogenomic Insights into the Biology and Treatment of HPV-Negative Head and Neck Squamous Cell Carcinoma. *Cancer Cell.* 39, 361–379. doi:10.1016/j.ccell.2020.12.007

Jiang, J., Li, M.-H., Inoue, K., Chu, X.-P., Seeds, J., and Xiong, Z.-G. (2007). Transient Receptor Potential Melastatin 7-like Current in Human Head and Neck Carcinoma Cells: Role in Cell Proliferation. *Cancer Res.* 67, 10929–10938. doi:10.1158/0008-5472.can-07-1121

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30. doi:10.1093/nar/28.1.27

Khan, H. Y., Mpilla, G. B., Sexton, R., Viswanadha, S., Penmetsa, K. V., Aboukameel, A., et al. (2020). Calcium Release-Activated Calcium (CRAC) Channel Inhibition Suppresses Pancreatic Ductal Adenocarcinoma Cell Proliferation and Patient-Derived Tumor Growth. *Cancers (Basel)* 12, 750. doi:10.3390/cancers12030750

Kishida, H., and Sugio, S. (2013). Crystal Structure of GTPase Domain Fused with Minimal Stalks from Human Dynamin-1-like Protein (Dlp1) in Complex with Several Nucleotide Analogues. *Curr. Top. Pept. Protein Res.* 14, 67–77.

Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., et al. (2013). How Good Is Automated Protein Docking? *Proteins* 81, 2159–2166. doi:10.1002/prot.24403

Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., et al. (2017). The ClusPro Web Server for Protein-Protein Docking. *Nat. Protoc.* 12, 255–278. doi:10.1038/nprot.2016.169

Lánczky, A., and Győrffy, B. (2021). Web-Based Survival Analysis Tool Tailored for Medical Research (KMplot): Development and Implementation. *J. Med. Internet Res.* 23, e27633. doi:10.2196/27633

Lewis, R. S. (2007). The Molecular Choreography of a Store-Operated Calcium Channel. *Nature* 446, 284–287. doi:10.1038/nature05637

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for Analysis of Tumor-Infiltrating Immune Cells. *Nucleic Acids Res.* 48, W509–w514. doi:10.1093/nar/gkaa407

Li, Y.-J., Cao, Y.-L., Feng, J.-X., Qi, Y., Meng, S., Yang, J.-F., et al. (2019). Structural Insights of Human Mitofusin-2 into Mitochondrial Fusion and CMT2A Onset. *Nat. Commun.* 10, 4914. doi:10.1038/s41467-019-12912-0

Liu, X., Liu, P., Chernock, R. D., Kuhs, K. A. L., Lewis, J. S., Jr., Luo, J., et al. (2020). A Prognostic Gene Expression Signature for Oropharyngeal Squamous Cell Carcinoma. *EBioMedicine* 61, 102805. doi:10.1016/j.ebiom.2020.102805

Liu, Y., Zheng, X., Mueller, G. A., Sobhany, M., Derose, E. F., Zhang, Y., et al. (2012). Crystal Structure of Calmodulin Binding Domain of Orai1 in Complex with Ca2+Calmodulin Displays a Unique Binding Mode. *J. Biol. Chem.* 287, 43030–43041. doi:10.1074/jbc.m112.380964

Ma, X., Cheng, K.-T., Wong, C.-O., O'Neil, R. G., Birnbaumer, L., Ambudkar, I. S., et al. (2011). Heteromeric TRPV4-C1 Channels Contribute to Store-Operated Ca(2+) Entry in Vascular Endothelial Cells. *Cell. Calcium* 50, 502–509. doi:10.1016/j.ceca.2011.08.006

Mayya, V., and Han, D. K. (2009). Phosphoproteomics by Mass Spectrometry: Insights, Implications, Applications and Limitations. *Expert Rev. Proteomics* 6, 605–618. doi:10.1586/epr.09.84

Mertins, P., Mani, D. R., Ruggles, K. V., Gillette, M. A., Clauser, K. R., Wang, P., et al. (2016). Proteogenomics Connects Somatic Mutations To Signalling In Breast Cancer. *Nature* 534, 55–62. doi:10.1038/nature18003

Miret-Casals, L., Sebastián, D., Brea, J., Rico-Leo, E. M., Palacín, M., Fernández-Salguero, P. M., et al. (2018). Identification of New Activators of Mitochondrial Fusion Reveals a Link between Mitochondrial Morphology and Pyrimidine Metabolism. *Cell. Chem. Biol.* 25, 268–278. doi:10.1016/j.chembiol.2017.12.001

Motiani, R. K., Hyzinski-García, M. C., Zhang, X., Henkel, M. M., Abdullaev, I. F., Kuo, Y.-H., et al. (2013). STIM1 and Orai1 Mediate CRAC Channel Activity and Are Essential for Human Glioblastoma Invasion. *Pflugers Arch. - Eur. J. Physiol.* 465, 1249–1260. doi:10.1007/s00424-013-1254-8

Nagy, Á., Munkácsy, G., and Győrffy, B. (2021). Pancancer Survival Analysis Of Cancer Hallmark Genes. *Sci. Rep.* 11, 6047. doi:10.1038/s41598-021-84787-5

Newton, H. S., Gawali, V. S., Chimote, A. A., Lehn, M. A., Palackdharry, S. M., Hinrichs, B. H., et al. (2020). PD1 Blockade Enhances K^+ Channel Activity, Ca^{2+} Signaling, and Migratory Ability in Cytotoxic T Lymphocytes of Patients with Head and Neck Cancer. *J. Immunother. Cancer* 8, e000844. doi:10.1136/jitc-2020-000844

Oell-Kotikangas, H., Schwab, U., Österlund, P., Saarilahti, K., Mäkitie, O., and Mäkitie, A. A. (2012). High Prevalence of Vitamin D Insufficiency in Patients with Head And Neck Cancer at Diagnosis. *Head Neck* 34, 1450–1455. doi:10.1002/hed.21954

Mounir, M., Lucchetta, M., Silva, T. C., Olsen, C., Bontempi, G., Chen, X., et al. (2019). New Functionalities in the TCGAbiolinks Package for the Study and Integration of Cancer Data from GDC and GTEx. *PLoS Comput. Biol.* 15, e1006701. doi:10.1371/journal.pcbi.1006701

Parekh, A. B., and Putney, J. W., Jr. (2005). Store-operated Calcium Channels. *Physiol. Rev.* 85, 757–810. doi:10.1152/physrev.00057.2003

Puram, S. V., Tirosh, I., Parikh, A. S., Patel, A. P., Yizhak, K., Gillespie, S., et al. (2017). Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell.* 171, 1611–1624. doi:10.1016/j.cell.2017.10.044

Putney, J. W., Jr. (1986). A Model for Receptor-Regulated Calcium Entry. *Cell. Calcium* 7, 1–12. doi:10.1016/0143-4160(86)90026-6

Rahman, S., and Rahman, T. (2017). Unveiling Some FDA-Approved Drugs as Inhibitors of the Store-Operated Ca^{2+} Entry Pathway. *Sci. Rep.* 7, 12881. doi:10.1038/s41598-017-13343-x

Rathner, P., Fahrner, M., Cerofolini, L., Grabmayr, H., Horvath, F., Krobath, H., et al. (2021). Interhelical Interactions within the STIM1 CC1 Domain Modulate CRAC Channel Activation. *Nat. Chem. Biol.* 17, 196–204. doi:10.1038/s41589-020-00672-8

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). g:Profiler: a Web Server for Functional Enrichment Analysis and Conversions of Gene Lists (2019 Update). *Nucleic Acids Res.* 47, W191–w198. doi:10.1093/nar/gkz369

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., et al. (2007). Global Survey of Phosphotyrosine Signaling Identifies Oncogenic Kinases in Lung Cancer. *Cell.* 131, 1190–1203. doi:10.1016/j.cell.2007.11.025

Scrideli, C. A., Carlotti, C. G., Jr., Okamoto, O. K., Andrade, V. S., Cortez, M. A. A., Motta, F. J. N., et al. (2008). Gene Expression Profile Analysis of Primary Glioblastomas and Non-neoplastic Brain Tissue: Identification of Potential Target Genes by Oligonucleotide Microarray and Real-Time Quantitative PCR. *J. Neurooncol* 88, 281–291. doi:10.1007/s11060-008-9579-4

Singh, A. K., Roy, N. K., Bordoloi, D., Padmavathi, G., Banik, K., Khwairakpam, A. D., et al. (2020). Orai-1 and Orai-2 Regulate Oral Cancer Cell Migration and Colonisation by Suppressing Akt/mTOR/NF-Kb Signalling. *Life Sci.* 261, 118372. doi:10.1016/j.lfs.2020.118372

Stanta, G., and Bonin, S. (2018). Overview on Clinical Relevance of Intra-tumor Heterogeneity. *Front. Med.* 5, 85. doi:10.3389/fmed.2018.00085

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Suzuki, M., Jeong, S.-Y., Karbowski, M., Youle, R. J., and Tjandra, N. (2003). The Solution Structure of Human Mitochondria Fission Protein Fis1 Reveals a Novel TPR-like Helix Bundle. *J. Mol. Biol.* 334, 445–458. doi:10.1016/j.jmb.2003.09.064

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING V11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-wide Experimental Datasets. *Nucleic Acids Res.* 47, D607–d613. doi:10.1093/nar/gky1131

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING Database in 2021: Customizable Protein-Protein Networks, and Functional Characterization of User-Uploaded Gene/measurement Sets. *Nucleic Acids Res.* 49, D605–d612. doi:10.1093/nar/gkaa1074

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-Based Map of the Human Proteome. *Science* 347, 1260419. doi:10.1126/science.1260419

Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., et al. (2017). New Additions to the C Lus P Ro Server Motivated by CAPRI. *Proteins* 85, 435–444. doi:10.1002/prot.25219

Van Linde, M. E., Labots, M., Brahm, C. G., Hovinga, K. E., De Witt Hamer, P. C., Honeywell, R. J., et al. (2022). Tumor Drug Concentration and Phosphoproteomic Profiles after Two Weeks of Treatment with Sunitinib in Patients with Newly-Diagnosed Glioblastoma. *Clin. Cancer Res* 28, 1595–1602. doi:10.1158/1078-0432.ccr-21-1933

Vangone, A., and Bonvin, A. M. (2015). Contacts-based Prediction of Binding Affinity in Protein-Protein Complexes. *Elife* 4, e07454. doi:10.7554/eLife.07454

Vangone, A., Schaarschmidt, J., Koukos, P., Geng, C., Citro, N., Trellet, M. E., et al. (2019). Large-scale Prediction of Binding Affinity in Protein-Small Ligand Complexes: the PRODIGY-LIG Web Server. *Bioinformatics* 35, 1585–1587. doi:10.1093/bioinformatics/bty816

Villalobos, C., Gutiérrez, L. G., Hernández-Morales, M., Del Bosque, D., and Núñez, L. (2018). Mitochondrial Control of Store-Operated Ca(2+) Channels in Cancer: Pharmacological Implications. *Pharmacol. Res.* 135, 136–143. doi:10.1016/j.phrs.2018.08.001

Villalobos, C., Hernández-Morales, M., Gutiérrez, L. G., and Núñez, L. (2019). TRPC1 and ORAI1 Channels in Colon Cancer. *Cell. Calcium* 81, 59–66. doi:10.1016/j.ceca.2019.06.003

Wang, J.-Y., Chen, B.-K., Wang, Y.-S., Tsai, Y.-T., Chen, W.-C., Chang, W.-C., et al. (2012). Involvement of Store-Operated Calcium Signaling in EGF-Mediated COX-2 Gene Activation in Cancer Cells. *Cell. Signal.* 24, 162–169. doi:10.1016/j.cellsig.2011.08.017

Wang, J., Zhao, H., Zheng, L., Zhou, Y., Wu, L., Xu, Y., et al. (2021). FGF19/SOCE/NFATc2 Signaling Circuit Facilitates the Self-Renewal of Liver Cancer Stem Cells. *Theranostics* 11, 5045–5060. doi:10.7150/thno.56369

Wang, Y.-Y., Wang, W.-C., Su, C.-W., Hsu, C.-W., Yuan, S.-S., and Chen, Y.-K. (2022). Expression of Orai1 and STIM1 in Human Oral Squamous Cell Carcinogenesis. *J. Dent. Sci.* 17, 78–88. doi:10.1016/j.jds.2021.07.004

Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., and Zemla, J. (2017). Package 'corrplot. *Statistician* 56, e24.

Wickham, H., Chang, W., and Wickham, M. H. (2016). Package 'ggplot2'. Create Elegant Data Visualisations Using the Grammar of Graphics. Version 2, 1–189.

Wright, D. J., Simmons, K. J., Johnson, R. M., Beech, D. J., Muench, S. P., and Bon, R. S. (2020). Human TRPC5 Structures Reveal Interaction of a Xanthine-Based TRPC1/4/5 Inhibitor with a Conserved Lipid Binding Site. *Commun. Biol.* 3, 704. doi:10.1038/s42003-020-01437-8

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A Universal Enrichment Tool for Interpreting Omics Data. *Innovation (N Y)* 2, 100141. doi:10.1016/j.xinn.2021.100141

Xue, L. C., Rodrigues, J. P., Kastritis, P. L., Bonvin, A. M., and Vangone, A. (2016). PRODIGY: a Web Server for Predicting the Binding Affinity of Protein-Protein Complexes. *Bioinformatics* 32, 3676–3678. doi:10.1093/bioinformatics/btw514

Yang, S., and Huang, X.-Y. (2005). Ca2+ Influx through L-type Ca2+ Channels Controls the Trailing Tail Contraction in Growth Factor-Induced Fibroblast Cell Migration. *J. Biol. Chem.* 280, 27130–27137. doi:10.1074/jbc.m501625200

Yang, S., Zhang, J. J., and Huang, X.-Y. (2009). Orai1 and STIM1 Are Critical for Breast Tumor Cell Migration and Metastasis. *Cancer Cell.* 15, 124–134. doi:10.1016/j.ccr.2008.12.019

Yedida, G., Milani, M., Cohen, G. M., and Varadarajan, S. (2019). Apogossypol-mediated Reorganisation of the Endoplasmic Reticulum Antagonises Mitochondrial Fission and Apoptosis. *Cell. Death Dis.* 10, 521. doi:10.1038/s41419-019-1759-y

Yu, C., Zhao, J., Yan, L., Qi, Y., Guo, X., Lou, Z., et al. (2020). Structural Insights into G Domain Dimerization and Pathogenic Mutation of OPA1. *J. Cell. Biol.* 219, e201907098. doi:10.1083/jcb.201907098

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, 284–287. doi:10.1089/omi.2011.0118

Yu, M., Nguyen, N. D., Huang, Y., Lin, D., Fujimoto, T. N., Molkentine, J. M., et al. (2019). Mitochondrial Fusion Exploits a Therapeutic Vulnerability of Pancreatic Cancer. *JCI Insight* 5, e126915. doi:10.1172/jci.insight.126915

Zhang, J., Wei, J., Kanada, M., Yan, L., Zhang, Z., Watanabe, H., et al. (2013). Inhibition of Store-Operated Ca2+ Entry Suppresses EGF-Induced Migration and Eliminates Extravasation from Vasculature in Nasopharyngeal Carcinoma Cell. *Cancer Lett.* 336, 390–397. doi:10.1016/j.canlet.2013.03.026

Zhang, M., Li, Q., Yu, D., Yao, B., Guo, W., Xie, Y., et al. (2019). GeNeCK: a Web Server for Gene Network Construction and Visualization. *BMC Bioinforma.* 20, 12. doi:10.1186/s12859-018-2560-0

Zhang, Z., Xu, Q., Song, C., Mi, B., Zhang, H., Kang, H., et al. (2020). Serum- and Glucocorticoid-Inducible Kinase 1 Is Essential for Osteoclastogenesis and Promotes Breast Cancer Bone Metastasis. *Mol. Cancer Ther.* 19, 650–660. doi:10.1158/1535-7163.mct-18-0783

Zheng, L., Stathopulos, P. B., Schindl, R., Li, G.-Y., Romanin, C., and Ikura, M. (2011). Auto-inhibitory Role of the EF-SAM Domain of STIM Proteins in Store-Operated Calcium Entry. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1337–1342. doi:10.1073/pnas.1015125108

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 | Frontiers in Genetics

ORIGINAL RESEARCH
published: 01 June 2022
doi: 10.3389/fgene.2022.909546

Cellular Senescence-Related Genes: Predicting Prognosis in Gastric Cancer

*Longfei Dai†, Xu Wang†, Tao Bai, Jianjun Liu, Bo Chen and Wenqi Yang**

Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei, China

Our study aimed to explore the effect of cellular senescence and to find potential therapeutic strategies for gastric cancer. Cellular senescence-related genes were acquired from the CellAge database, while gastric cancer data were obtained from GEO and TCGA databases. SMARCA4 had the highest mutation frequency (6%), and it was linked to higher overall survival (OS) and progression-free survival (PFS). The gastric cancer data in TCGA database served as a training set to construct a prognostic risk score signature, and GEO data were used as a testing set to validate the accuracy of the signature. Patients with the low-risk score group had a longer survival time, while the high-risk score group is the opposite. Patients with low-risk scores had higher immune infiltration and active immune-related pathways. The results of drug sensitivity analysis and the TIDE algorithm showed that the low-risk score group was more susceptible to chemotherapy and immunotherapy. Most patients with mutation genes had a lower risk score than the wild type. Therefore, the risk score signature with cellular senescence-related genes can predict gastric cancer prognosis and identify gastric cancer patients who are sensitive to chemotherapy and immunotherapy.

Keywords: cellular senescence, GC, prognosis, chemotherapy, immunotherapy

OPEN ACCESS

Edited by:
Dhanendra Tomar,
Wake Forest School of Medicine,
United States

Reviewed by:
Gaurav Kumar,
Medical College of Wisconsin,
United States
Xinwei Han,
Zhengzhou University, China
Milton Roy,
Johns Hopkins Medicine,
United States

****Correspondence:***
Wenqi Yang
doctorwenqiyang@163.com

†These authors have contributed equally to this work

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

Received: *31 March 2022*
Accepted: *25 April 2022*
Published: *01 June 2022*

Citation:
Dai L, Wang X, Bai T, Liu J, Chen B and Yang W (2022) Cellular Senescence-Related Genes: Predicting Prognosis in Gastric Cancer. Front. Genet. 13:909546. doi: 10.3389/fgene.2022.909546

INTRODUCTION

Nowadays, cancer is the primary cause of threat to human health (Bray et al., 2021). It ranked fifth in incidence and fourth in mortality worldwide, while the number of GC diagnosed in 2020 was more than 1 million and the number of deaths was more than 700,000 (Sung et al., 2021). There are various treatments for GC, such as surgery, radiotherapy, targeted therapy, and immunotherapy (Joshi and Badgwell, 2021). Currently, early-stage GC is mainly treated by surgical resection (Eusebi et al., 2020). It was found that early-stage GC treated by surgery has a 5-year survival probability above 60%, but late-stage GC is only between 18% and 50% (Sexton et al., 2020). Moreover, the appearance of resistance to chemotherapy drugs has greatly reduced the effectiveness of chemotherapy (Zhang et al., 2022). Therefore, a new therapeutic strategy is urgently needed to improve this situation.

Cellular senescence is an irreversible way of cell proliferation cessation. It not only stops the cell division cycle but also activates the senescence-associated secretory phenotype (SASP), which affects the cellular metabolism (Birch and Gil, 2020). Cellular senescence is a Jekyll and Hyde phenomenon, that is, both beneficial in inhibiting the division of DNA-damaged cells to form tumors and deleterious due to the promotion of cancer cell invasion and distant metastasis, especially in cells with stronger SASP (Coppé et al., 2008; Demirci et al., 2021; Yasuda et al., 2021). Studying the effect of cellular senescence in GC could help develop a new approach to cancer therapy (Zhou et al., 2022). Therefore, the study of cellular senescence in GC is crucial.

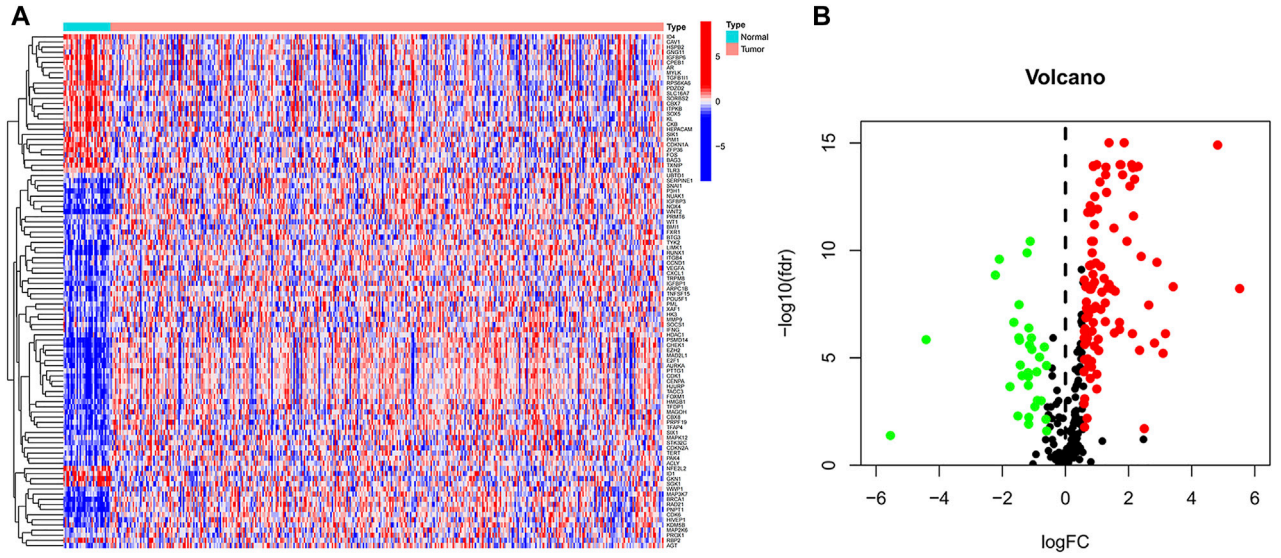


FIGURE 1 | Cellular senescence-related differential genes between tumor samples and normal samples. **(A)** Heat map. **(B)** Volcano plot. Green, downregulated; red, upregulated.

Machine-learning-derived signatures are useful in predicting cancer prognosis and guiding immunotherapy (Liu Z et al., 2021; Liu et al., 2022a; Liu et al., 2022b). In the study, we constructed a cellular senescence prognostic risk score signature by analyzing the role of cellular senescence in GC. The signature can independently predict GC patients' prognosis and effectively differentiate patients who are more sensitive to chemotherapy and immunotherapy. The findings of this study may provide new strategies for exploring the therapy of GC.

METHODS

Acquisition of Gastric Cancer Samples and Cellular Senescence-Related Genes

The process diagram is shown in **Supplementary Figure S1**. We acquired transcriptome data, clinical information, and mutation information of GC from TCGA databases. The gene symbol ID was translated to gene name in transcriptome data. Tumor mutation load (TMB) was calculated. TMB refers to how many bases per million bases are mutated. The platform file (GPL6947) and probe matrix file (GSE84437) were extracted from GEO. The correspondence between the probe matrix and gene names was found according to the platform file annotation information. The probe matrix was converted to a gene matrix to obtain the expression of each gene. Cellular senescence-related genes were downloaded from CellAge. A total of 279 cellular senescence-related genes were included in this study (**Supplementary Table S1**).

Identification of Prognostic Differential Genes

Differential analysis was conducted by the "limma" package to select differentially expressed genes (DEGs) in normal samples and tumor samples. DEGs were visualized by drawing heat maps and volcano maps. Next, we also extracted the expression of DEGs. Expression data and survival data were merged. Prognostic-associated genes were identified based on univariate Cox analysis. The waterfall plot of prognostic genes was plotted by the "maftools" package to obtain the mutation frequency of each gene.

Constructing and Validating a Prognostic Signature

TCGA data were used as a training set to construct the prognostic model, and GEO data served as a testing set to validate the model accuracy. Formula: riskscore = $\sum i1(Coefi*ExpGenei)$. "Coef," regression coefficient; "ExpGene," gene expression. The risk score was acquired for each sample based on the model formula. Training and testing sets were separated into two groups of high and low risk according to the median risk score. Principal component analysis (PCA) was performed to demonstrate the accuracy of distinguishing the two groups based on the signature. The survival difference in the two groups was compared by Kaplan–Meier analysis. The predictive accuracy of the signature was evaluated by plotting ROC curves using the "survivalROC" package. The signature was explored as an independent prognostic factor by univariate and multivariate Cox analyses. The "ggpubr" package was employed to investigate the differences in risk scores among clinical features. Immunotyping analysis was conducted to explore whether risk scores were different among different immunotypes.

Development of a Nomogram

By using "regplot" and "rms" packages, nomogram and calibration curves were developed. Total points were obtained based on summing the scores of the clinical characteristics in the nomogram to predict patients' survival. We also used the

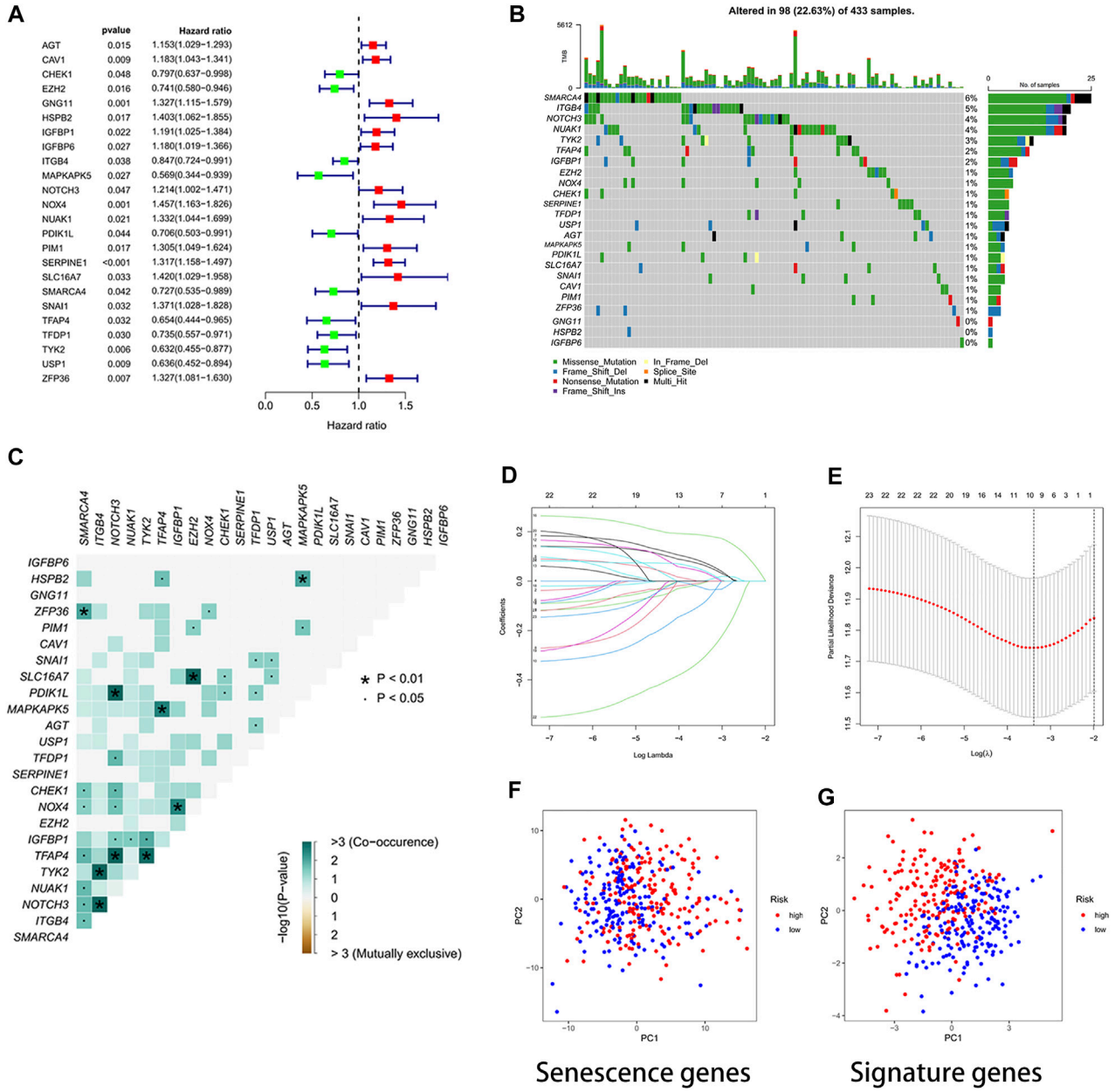


FIGURE 2 | Developing a prognostic signature. **(A)** Forest plot. The 24 cellular senescence-related genes associated with GC prognosis. **(B)** Waterfall plot. Mutation frequency of 24 cellular senescence-associated genes. A total of 98 sample mutations have occurred in 433 gastric cancer samples. **(C)** Mutation co-occurrence and exclusion analysis. Green, co-occurrence; purple, exclusion. **(D)** LASSO regression coefficients. **(E)** Identified genes were used to construct a signature. **(F)** PCA diagram. The high- and low-risk groups were distinguished by cellular senescence-associated genes. Groups marked in blue represent low-risk patients, and groups marked in red represent high-risk patients. **(G)** PCA plot. The risk score signature genes distinguished high- and low-risk groups of patients with high accuracy.

"timeROC" package to draw ROC curves to compare the accuracy of the nomogram and clinical characteristics in predicting survival. Then, we confirmed whether the nomogram could be used as an independent predictor of prognosis based on univariate and multivariate Cox analyses. C-index curves were constructed using the "survcomp" package.

Exploring the Association Between Risk Scores and Immunotherapy

Immune cell infiltration analysis was undertaken to acquire the immune cell content of each sample (**Supplementary Table S2**). "reshape2" and "ggpubr" packages were performed to observe immune cell differences and immune-related functional

differences between different risk groups. The "GSVA" package was applied to explore the functional or pathway differences between different risk groups. We acquired the reference gene set "c2.cp.kegg.v7.1.symbols" from the Molecular Signature Database (https://www.gsea-msigdb.org/gsea/msigdb). The samples were categorized into mutation and wild type based on the gene mutation status. The difference in risk scores between mutation and wild type was observed by plotting box plots with the "ggpubr" package. Drug sensitivity analysis was conducted by the pRRophetic package (https://www.cancerrxgene.org/) to investigate IC_{50} (the half-maximal inhibitory concentration) differences between high- and low-risk groups. We also used the TIDE (http://tide.dfci.harvard.edu/) algorithm to predict the response of different risk groups to immunotherapy.

Enrichment Analysis of Differentially Expressed Genes

The expression differences of cellular senescence-related genes in tumor and normal samples were further analyzed by the "limma" package. We also conducted GO and KEGG enrichment analyses for DEGs with the "clusterProfiler" package.

Construction of the Protein–Protein Interaction Network

A PPI network (interaction score >0.70) was constructed using the STRING database (https://string-db.org/). PPI network data were further processed using Cytoscape software (https://cytoscape.org/). The plugin cytoHubba was applied to explore the hub genes of DEGs. "limma" and "beeswarm" packages were used to investigate the differentially expressed hub genes in normal and tumor tissues. The samples were categorized into high- and low-expression groups based on the median expression values of the hub genes. Survival differences between the two groups were investigated by Kaplan–Meier analysis. Finally, we also explored the differences in the gene expression in immune infiltration and different clinical features.

Statistical Analysis

R 4.1.2 and Strawberry-Perl-5.32.1.1 were employed in this study. *p*-values less than 0.05 were regarded as statistically significant. Survival differences between different groups were investigated by performing a Kaplan–Meier analysis. The independent predictors of GC were identified by univariate and multivariate Cox analyses. The accuracy of the signature and nomogram in predicting survival was explored by ROC analysis.

RESULTS

Identification of Cellular Senescence-Related Differential Genes

In TCGA data, we identified 135 differential genes by comparing the difference in the expression of cellular senescence-related genes in tumor and normal tissue samples (FDR <0.05 and logFC = 0.585). The heat map (**Figure 1A**) and volcano map (**Figure 1B**) visualized the aforementioned results. There were 32 genes significantly hyper-expressed and 103 genes significantly down-expressed in the tumor tissue samples.

Construction of a Prognostic Signature

The 24 cellular senescence DEGs associated with GC prognosis were identified through univariate Cox analysis, such as SMARCA4 (**Figure 2A**). **Figure 2B** shows the somatic mutations of 24 genes with a mutation frequency of 22.63% (98 out of 433 GC samples showed mutations in cellular senescence-related genes). Of these, SMARCA4 had the highest mutation frequency (6%), while GNG11 and IGFBP6 were not mutated (0%). We also found a significant difference between the high- and low-expression of SMARCA4, and patients with high expression of SMARCA4 were associated with higher overall survival (OS) and progression-free survival (PFS) (**Supplementary Figure S2**). Interestingly, there was a mutation co-occurrence relationship between SMARCA4 and ZFP36, ITGB4 and TYK2, ITGB4 and NOTCH3, NOTCH3 and PDIK1L, NOTCH3 and TFAP4, TFAP4 and MAPKAPK5, TFAP4 and TYK2, EZH2 and SLC16A7, IGFBP1 and NOX4, and HSPB2 and MAPKAPK5 (**Figure 2C**). Next, we further identified 24 cellular senescence DEGs associated with gastric cancer prognosis by using the least absolute shrinkage and selection operator (LASSO) Cox regression analysis. A total of 11 genes (AGT, CHEK1, GNG11, IGFBP1, MAPKAPK5, NOX4, SERPINE1, TFDP1, TYK2, USP1, and ZFP36) were identified (**Figures 2D,E**). Meanwhile, we developed a prognostic risk score signature based on the 11 genes mentioned earlier in the training set (**Supplementary Table S3**). Formula: risk score = (0.0617318899456387) × AGT + (−0.004416679732613) × CHEK1 + (0.00146582976956934) × GNG11 + (0.027549739978882) × IGFBP1 + (−0.0823743685561005) × MAPKAPK5 + (0.0337754127670565) × NOX4 + (0.184215619451523) × SERPINE1 + (−0.00197579186740112) × TFDP1+ (−0.303214268137102) × TYK2 + (−0.0314400326211636) × USP1 + (0.0400501256934474) × ZFP36 (**Supplementary Table S3**). We found that the signature could accurately distinguish low-risk and high-risk samples in GC by PCA (**Figures 2F,G**).

Validation of Signature Genes in the HPA Database

To investigate the protein expression of the signature genes in normal and gastric cancer tissues, we downloaded immunohistochemical images of gastric cancer tissues and normal tissues from the Human Protein Atlas database (https://www.proteinatlas.org/). We found that MAPKAPK5 and USP1 proteins were highly expressed in gastric cancer tissues, while the ZFP36 protein was lowly expressed in tumor tissues (**Supplementary Figure S3**).

Predicting Survival With the Risk Score Signature

Through survival curves, we observed longer overall survival (OS) and progression-free survival (PFS) in the low-risk subgroup of the training set (**Figures 3A,C**). The

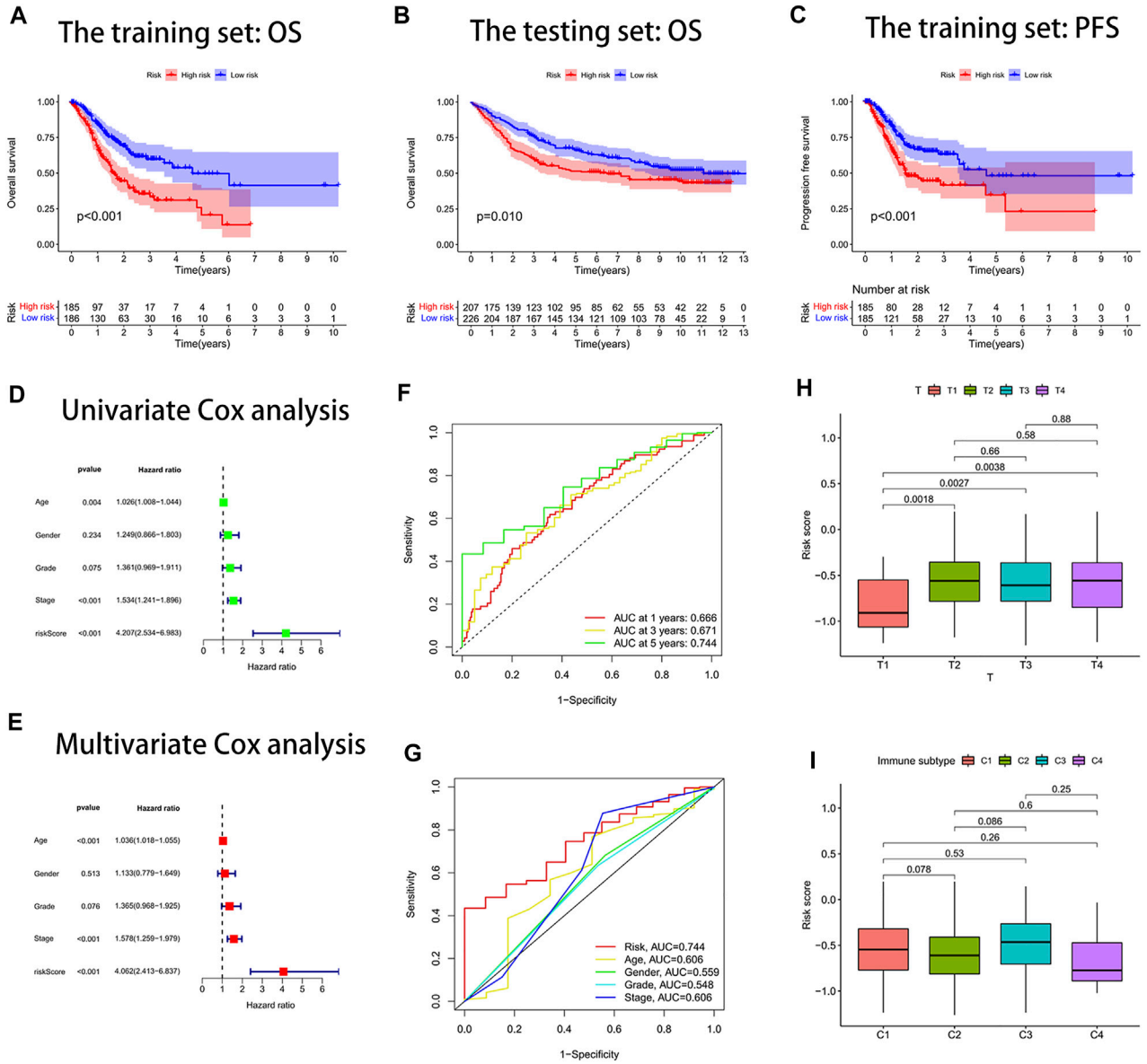


FIGURE 3 | Risk score signature predicted prognosis for gastric cancer patients. **(A)** Overall survival (OS) curves of the high- and low-risk groups in the training set. **(B)** Overall survival (OS) curves of the high- and low-risk groups in the testing set. **(C)** Progression-free survival (PFS) curves of the high- and low-risk groups in the training set. **(D)** Univariate Cox analysis. **(E)** Multivariate Cox analysis. **(F)** Area under the ROC curve (AUC) for the risk score signature that predicted 1-year, 3-year, and 5-year overall survival. **(G)** ROC curves. The area under the ROC curve (AUC) for the risk score was the highest at 0.744. **(H)** Box plot of the difference in risk score for patients with different T-stages. **(I)** Box line plot of the difference in the risk score for patients with different immune subtypes.

aforementioned results were confirmed in the testing set (**Figure 3B**). The results of univariate and multivariate Cox analyses indicated that the risk score signature could be used as an independent prognostic factor for gastric cancer patients independently of other clinical characteristics (**Figures 3D,E**). The signature was very accurate in predicting survival in patients with gastric cancer, with an area under the ROC curve (AUC) of more than 0.60 for predicting 1-, 3-, and 5-year survival (**Figure 3F**). We found the largest area under the ROC curve for the risk score (AUC = 0.744), which indicated that the signature predicted survival better than other clinical characteristics (**Figure 3G**). We further investigated whether there were differences in risk scores across clinical characteristics (age, gender, grade, stage, and TNM stage). We found an increased risk for patients after the T1 stage and no significant change in risk for patients after the T2 stage (**Figure 3H**). In contrast, there were no significant differences in risk scores for other clinical characteristics (**Supplementary Figure S4**). Interestingly, we also found no difference in risk scores for immune subtypes (**Figure 3I**).

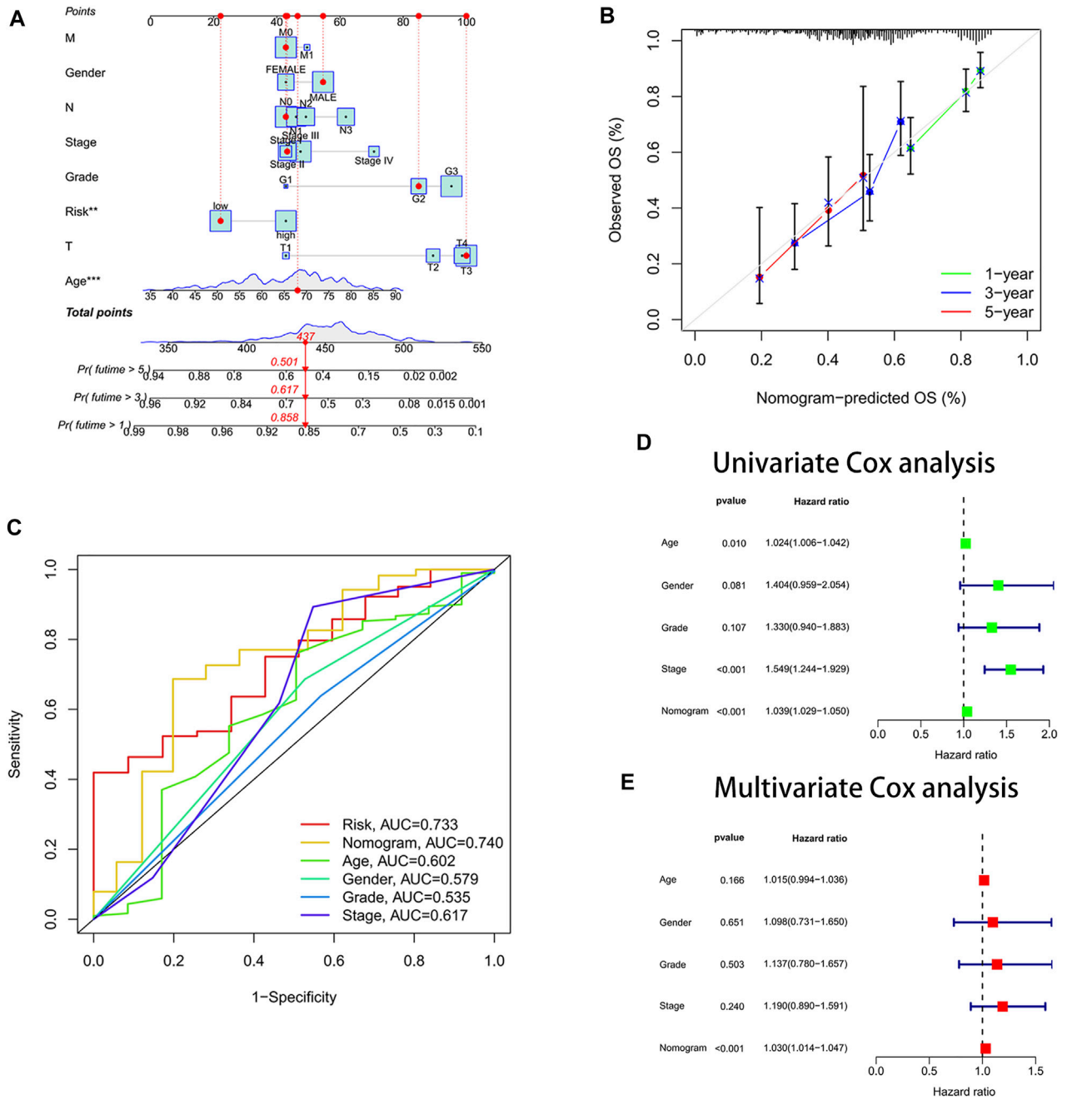


FIGURE 4 | Constructed a nomogram for predicting survival. **(A)** Nomogram. **(B)** Calibration curves. The y-axis is the actual survival rate, and the x-axis is the predicted survival rate. **(C)** Area under the ROC curve (AUC) of the nomogram. **(D)** Univariate Cox analysis. **(E)** Multivariate Cox analysis.

Development of a Nomogram

We drew a nomogram to predict patients' survival (**Figure 4A**). When patients' total point was 437, the predicted survival rate at 1-year was more than 0.858, the predicted survival rate at 3-year was more than 0.617, and the predicted survival rate at 5-year was more than 0.501. We found that the actual survival rate and predicted survival rate were almost in agreement by observing the calibration curve (**Figure 4B**). It validated the high accuracy of the nomogram in predicting the survival rate of gastric cancer patients. In addition, we also found the largest area under the ROC curve for the nomogram (AUC = 0.740) (**Figure 4C**). It implied that the nomogram predicted patients' survival better than other clinical characteristics. The nomogram was confirmed to be an indicator of independent prognosis by the results of univariate and multivariate Cox analyses (**Figures 4D,E**). We randomly selected four prognostic signature articles of gastric cancer in the latest 3 years from the PubMed website (https://pubmed.ncbi.nlm.nih.gov/), including Dai's signature (ITGAV,

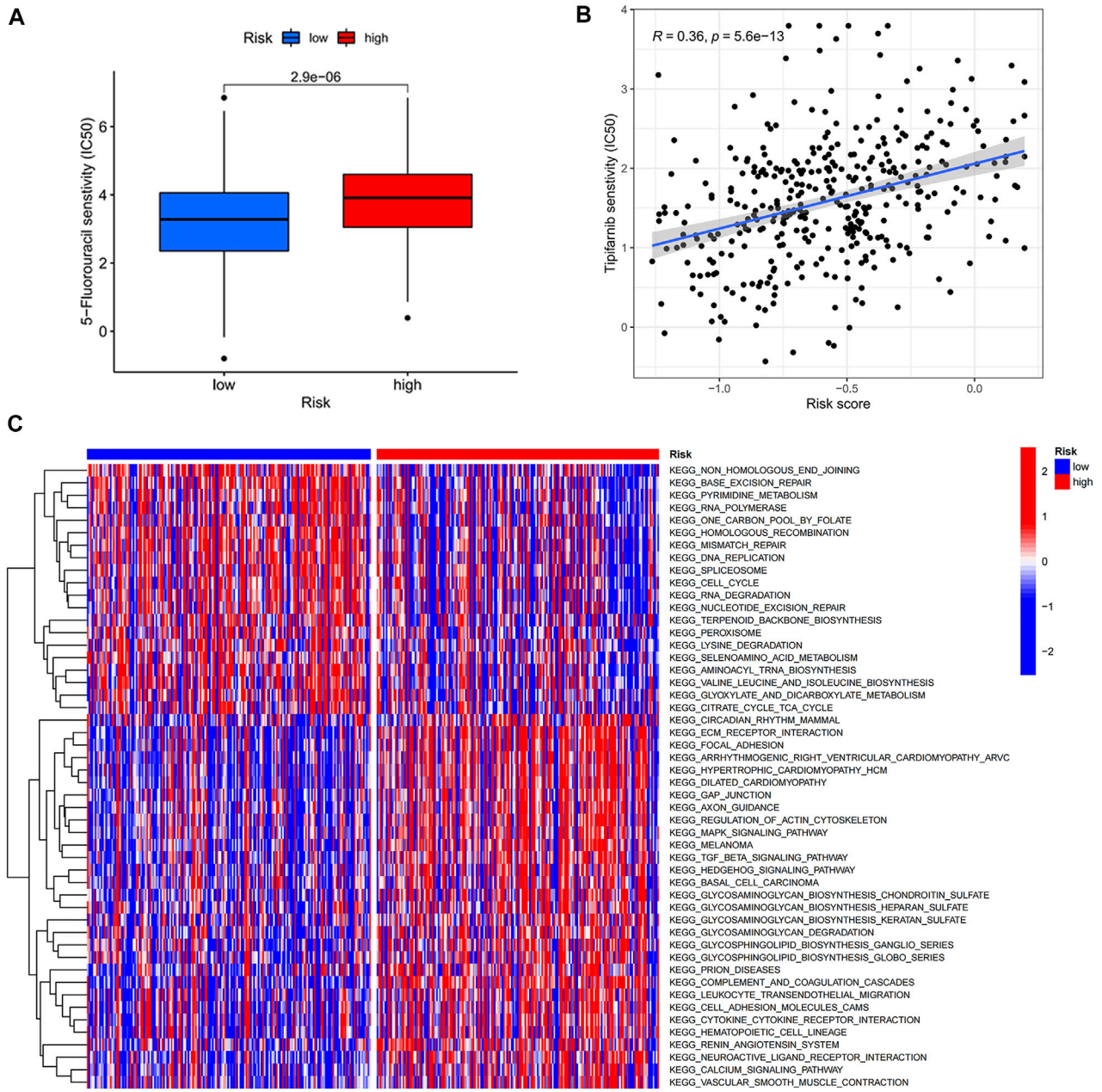


FIGURE 5 | Risk score-guided chemotherapy. **(A)** Box plot of IC_{50} differential analysis for high- and low-risk score groups. **(B)** Scatter plot of correlation between risk score and IC_{50}. **(C)** Heat map of the differential analysis of GSVA enrichment between high- and low-risk groups.

DAB2, SERPINE1, MATN3, and PLOD2), Liu's signature (NOX4, NOX5, GLS2, MYB, TGFBR1, NF2, AIFM2, ZFP36, SLC1A4 TXNIP, CXCL2, HAMP, and SP1), Meng's signature (CGB5, IGFBP1, OLFML2B, RAI14, SERPINE1, IQSEC2, and MPND), and Yin's signature (GPX3, ABCA1, NNMT, NOS3, SLCO4A1, ADH4, DHRS7, and TAP1) (Meng et al., 2020; Liu SJ et al., 2021; Dai et al., 2021; Yin et al., 2021). To highlight the advantages of the cellular senescence signature, we compared these five signatures, and the results are visualized in **Supplementary Figure S5**. We found that the cellular senescence signature was the best predictor of prognosis in gastric cancer patients, with a C-index of 0.642.

Risk Score Guide Clinical Treatment

Due to the increase in tumor resistance to chemotherapeutic drugs, most patients with gastric cancer currently have poor chemotherapy outcomes. We explored whether risk scores could play a role in chemotherapy. In our study, the risk score was significantly and positively correlated with half-maximal inhibitory concentration (IC_{50}), and the low-risk

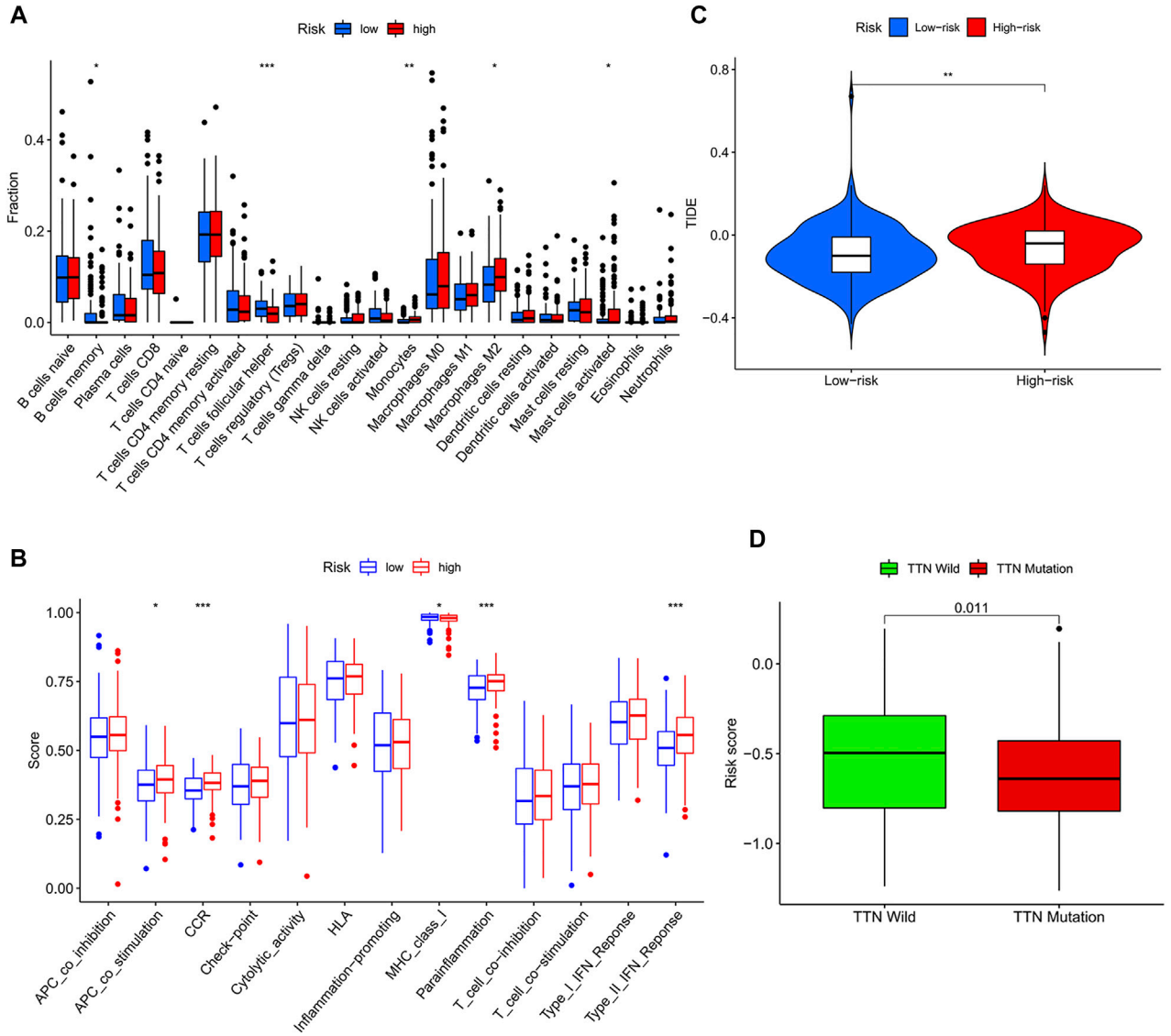


FIGURE 6 | Risk score-guided immunotherapy. **(A)** Box plot of the differential analysis of immune infiltration between the two risk score groups. **(B)** Box plot of the differential analysis of immune function between two risk score groups. **(C)** Violin plot of the response to immunotherapy between the two risk fraction groups calculated by the IDE algorithm. **(D)** Box plot of the differential analysis of the risk score between wild type and mutant type for the top 10 mutated genes in TCGA data.

score group had a lower IC_{50} value and was more sensitive to 5-FU (**Figures 5A,B**). By performing GSVA, most of the senescence pathways were found to be more active in the low-risk score group (**Figure 5C**). The high-risk score group had higher macrophage M2 infiltration, and the low-risk score group had higher B-cell memory and T-cell follicular helper infiltration (**Figure 6A**). In addition, immune function analysis showed that type_II_IFN_response and parainflammation were more active in the high-risk group, and MHC_class_I was more active in the low-risk group (**Figure 6B**). It suggested that the low-risk group might be more suitable for immunotherapy. It was confirmed by the TIDE algorithm that patients in the low-risk score group are more suitable for immunotherapy (**Figure 6C**). The risk score signature constructed using cellular senescence-related genes is a potential biomarker for assessing the clinical response to immunotherapy in gastric cancer patients. We also identified the top 10 mutated genes (TTN, TP53, MUC16, ARID1A, LRP1B, SYNE1, FLG, FAT4, CSMD3, and PCLO), SMARCA4, and ZFP36 in TCGA data (**Supplementary Table S4**). The samples were classified into mutation and wild types according to the mutation status of the genes. Among them, the mutation type of the six genes (TTN, ARID1A, LRP1B, FLG, FAT4, and PCLO) had lower risk scores (**Figure 6D**). We also found no difference in risk scores between mutation and wild type of SMARCA4 and ZFP36 (**Supplementary Figure S6**), so we speculated that co-mutation of SMARCA4 and ZFP36 does not affect the prognosis of gastric cancer.

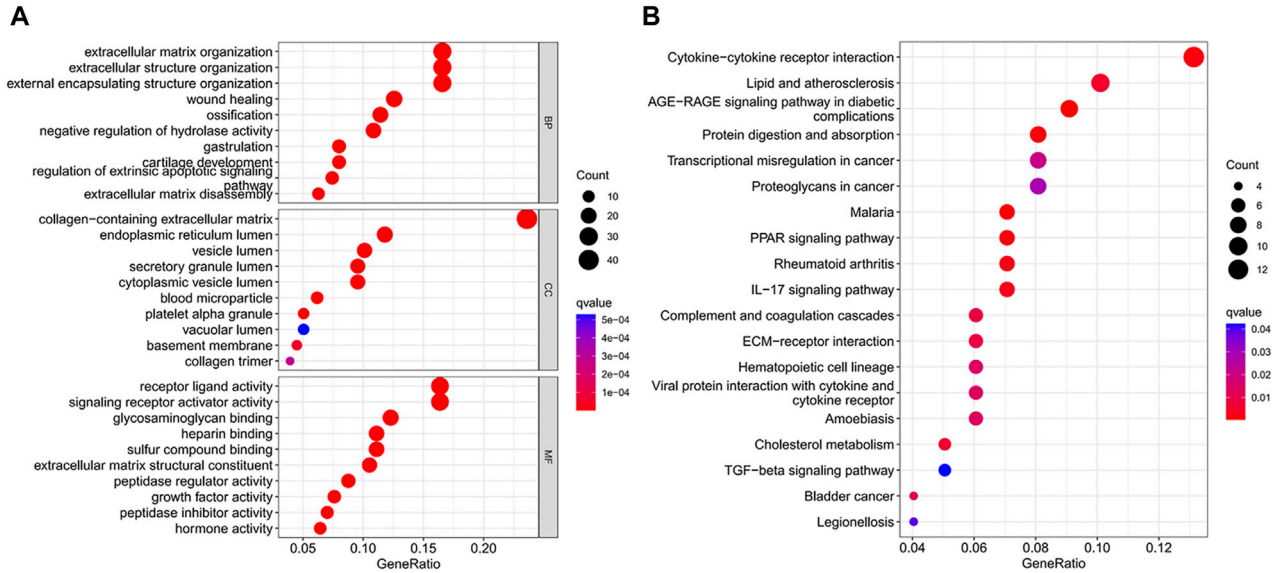


FIGURE 7 | Differential gene enrichment analysis. **(A)** GO enrichment analysis bubble map. **(B)** Bubble plot of KEGG enrichment analysis.

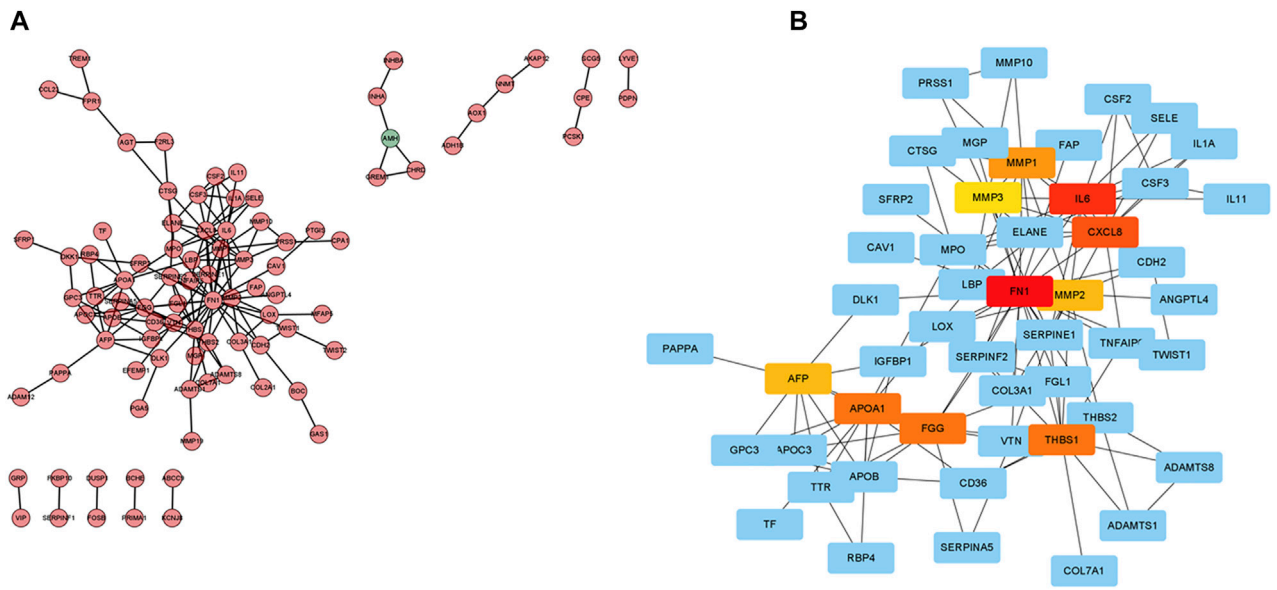


FIGURE 8 | Protein–protein interaction (PPI) network. **(A)** Cytoscape-treated PPI network. Red, highly expressed DEGs in the high-risk score group; green, highly expressed DEGs in the low-risk score group. **(B)** CytoHubba identified the top 10 hub genes.

Differential Gene Enrichment Analysis

We identified 186 differential genes in two risk groups. GO and KEGG enrichment analyses were performed on the differential genes, and the enrichment results were visualized in bubble plots. We found that extracellular matrix organization, extracellular structure organization, and external encapsulating structure organization were significantly enriched in the GO bubble map (**Figure 7A**), while cytokine–cytokine receptor interaction, protein digestion and absorption, transcriptional misregulation in cancer, and proteoglycans in cancer were significantly enriched in the KEGG bubble plots (**Figure 7B**).

Identification of 10 Hub Genes

The expression profiles of DEGs in two risk score groups were evaluated using the STRING database. PPI networks were constructed (**Supplementary Figure S7**). By using Cytoscape software, PPI network data were processed and displayed. The interactions of DEGs are shown in **Figure 8A**. A total of 10 hub

genes (FN1, IL6, CXCL8, THBS1, APOA1, FGG, MMP1, AFP, MMP2, and MMP3) of DEGs were identified using Cytoscape plugin cytoHubba and the extent method (**Figure 8B**). A total of six upregulated genes (FN1, APOA1, CXCL8, MMP1, MMP3, and THBS1) in tumor tissue were identified by the differential analysis of 10 hub genes (**Supplementary Figure S8**). By further analysis, five hub genes (FN1, APOA1, CXCL8, MMP1, and THBS1) with survival differences were identified (**Supplementary Figure S9**). Patients with low-expression levels had a better prognosis. We also analyzed the differences in the expression levels of genes in different clinical characteristics. FN1 was significantly more expressed in patients after stage T1 and unchanged in patients after stage T2. The expression levels of FN1 and THBS1 were higher in G3 patients than those in G2 patients (**Supplementary Figures S10, 14**). The expression levels of APOA1 were higher in G2 patients than those in G3 and in N2 than those in N0 (**Supplementary Figure S11**). The expression level of CXCL8 was significantly higher in patients over 65 years of age and after stage III (**Supplementary Figure S12**). The expression level of MMP1 was significantly higher in patients over 65 years of age and in stage IV than that in stage I (**Supplementary Figure S13**). Finally, we also performed the differential analysis of immune cell infiltration (**Supplementary Figure S15**). The FN1, APOA1, CXCL8, MMP1, and THBS1 low-expression groups all had higher immune cell infiltration and might be suitable for immunotherapy.

DISCUSSION

Cellular senescence is the result of irreversible cessation of cell division (Gorgoulis et al., 2019). Studies have shown that it can occur in the context of oncogene activation and is involved in tumor suppression (Di Micco et al., 2021). The latest studies have shown that senescent cancer cells have not only antitumor activity but also pro-tumor activity. Cellular senescence can play an essential role in immune surveillance to ensure that senescent cancer cells are eliminated (Prasanna et al., 2021). Nowadays, cellular senescence is emerging as a potentially novel anticancer strategy (Ramu et al., 2021). It could help guide effective anticancer therapy strategies by exploring the cellular senescence patterns of GC.

The main purpose of this study was to discuss the effect of cellular senescence on the prognosis and treatment of GC. We constructed a prognostic risk score signature for cellular senescence-related genes using TCGA data. Patients with low-risk scores had longer survival times, while the opposite was true for patients with high-risk scores. The same results were found in the GEO data. It indicated that the prognostic risk score signature could forecast the GC patients' prognosis. We also observed that the prognostic risk score signature could be an independent prognostic factor for GC by further Cox analysis. In addition, a nomogram was constructed for predicting gastric cancer patients' survival. The calibration curve confirmed the predictive accuracy of the nomogram. Encouragingly, the area under the ROC curve (AUC) of the nomogram was significantly higher than other clinical features, especially in traditional TNM stages. It showed that the nomogram has higher accuracy in predicting 1-year, 3-year, and 5-year survival rates of gastric cancer patients than clinical TNM stages. Moreover, the cellular senescence signature had the highest C-index and predicted the best prognosis among other prognosis-related signatures of gastric cancer.

Although chemotherapeutic agents are helpful in the therapy of GC, many GC patients appear resistant to chemotherapy, resulting in poorer chemotherapy outcomes (Wei et al., 2020). Therefore, it is increasingly essential to identify GC patients who are sensitive to chemotherapeutic drugs. According to these reasons, we investigated the differences in clinical response to chemotherapeutic drugs in two risk groups. In our research, GC patients with low-risk scores were more susceptible to 5-FU. It suggested that using the risk score could identify gastric cancer patients who are more suitable for chemotherapy. With the development of technology, more and more therapeutic approaches are available for GC (Hsu and Raufi, 2021). Immunotherapy is an emerging cancer treatment that activates the body's immune system to clear tumor cells (Kawazoe et al., 2021). The identification of patients with gastric cancer suitable for immunotherapy is particularly critical in the clinical environment. We observed higher immune infiltration levels in the low-risk score group, including B-cell memory and T-cell follicular helpers, and the high-risk score group had higher infiltration levels of macrophage M2 (tumor-promoting cells) (Xia et al., 2020; Overacre-Delgoffe et al., 2021). The results of immune function analysis also showed that high-risk score patients had active immune-related functions "type_II_IFN_response" and "parainflammation," whereas "MHC_class_I" was more active in the low-risk score group. Previous studies have shown that "type_II_IFN_response" is considered an anticancer immune-related function (Liu M et al., 2020). Interestingly, our findings showed the opposite that "type_II_IFN_response" might promote the development of gastric cancer. The immune-related function "parainflammation" is thought to promote tumor progression, which is consistent with our findings (Aran et al., 2016). The immune-related function "MHC_class_I," which mainly plays a role in the immunosurveillance of cancer, inhibits the immune escape of tumors and is considered a potential target for cancer immunotherapy (Cornel et al., 2020; Dersh et al., 2021). We speculated that patients with low-risk scores may be suitable for immunotherapy. Next, we demonstrated that low-risk score patients had a low immune escape potential and were more sensitive to immunotherapy using the TIDE algorithm. In conclusion, the prognostic risk score signature with cellular senescence genes not only predicts prognosis but also identifies patients with chemotherapy- and immunotherapy-sensitive gastric cancer. We also analyzed the top 10 mutated genes in the TCGA data. The mutation types of TTN, ARID1A, LRP1B, FLG, FAT4, and PCLO had lower risk scores than the wild type. This meant that patients with mutation types might have a better prognosis and be more suitable for chemotherapy and immunotherapy. We also identified a particular gene SMARCA4. It has the highest mutation frequency (6%), and it is linked to higher overall survival (OS) and progression-free survival (PFS). There was a co-mutation relationship between

SMARCA4 and ZFP36. But we found no difference between mutation and wild type in SMARCA4 and ZFP36, so we speculated that co-mutation of SMARCA4 and ZFP36 does not affect the prognosis of gastric cancer.

Because of the significant differences between the two risk groups, it is essential to study the differential genes in depth. We identified 10 hub genes (FN1, IL6, CXCL8, THBS1, APOA1, FGG, MMP1, AFP, MMP2, and MMP3) by constructing a PPI network. FN1, APOA1, CXCL8, MMP1, MMP3, and THBS1 were significantly upregulated in the tumor samples. This result was confirmed in the HPA database. We observed that FN1, APOA1, CXCL8, MMP1, and THBS1 were correlated with GC prognosis, with higher expression levels associated with a worse prognosis. This is consistent with previously published research studies (Li et al., 2019; Chen X et al., 2020; Chen YJ et al., 2020; Liu X et al., 2020; Zhang et al., 2021). We also found higher immune infiltration (plasma cells and T cells) in the low-expression group of FN1, APOA1, CXCL8, MMP1, and THBS1, while macrophages M2 and resting T cells showed higher infiltration in the high expression group. It suggested that patients in the low-expression group of FN1, APOA1, CXCL8, MMP1, and THBS1 might be more suitable for immunotherapy.

In summary, the cellular senescence risk score prognostic signature could be used to assess the prognosis of GC patients and guide clinical treatment. Our study not only provided a new predictive signature for the prognosis of GC but also offered guidance for the future therapy of gastric cancer.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

LD planned the research. LD, TB, JL, and BC downloaded data and performed the data analysis. XW, LD, TB, and JL drew the graphs. Finally, LD wrote the manuscript, and WY checked the manuscript.

ACKNOWLEDGMENTS

Authors thank the following databases for supporting this article: TCGA, GEO, CellAge, and STRING databases.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.909546/full#supplementary-material

REFERENCES

Aran, D., Lasry, A., Zinger, A., Biton, M., Pikarsky, E., Hellman, A., et al. (2016). Widespread Parainflammation in Human Cancer. *Genome Biol.* 17 (1), 145. doi:10.1186/s13059-016-0995-z

Birch, J., and Gil, J. (2020). Senescence and the SASP: Many Therapeutic Avenues. *Genes Dev.* 34 (23-24), 1565–1576. doi:10.1101/gad.343129.120

Bray, F., Laversanne, M., Weiderpass, E., and Soerjomataram, I. (2021). The Ever-increasing Importance of Cancer as a Leading Cause of Premature Death Worldwide. *Cancer* 127 (16), 3029–3030. doi:10.1002/cncr.33587

Chen X, X., Chen, R., Jin, R., and Huang, Z. (2020). The Role of CXCL Chemokine Family in the Development and Progression of Gastric Cancer. *Int. J. Clin. Exp. Pathol.* 13 (3), 484–492.

Chen YJ, Y. J., Liang, L., Li, J., Wu, H., Dong, L., Liu, T.-T., et al. (2020). IRF-2 Inhibits Gastric Cancer Invasion and Migration by Down-Regulating MMP-1. *Dig. Dis. Sci.* 65 (1), 168–177. doi:10.1007/s10620-019-05739-8

Coppé, J.-P., Patil, C. K., Rodier, F., Sun, Y., Muñoz, D. P., Goldstein, J., et al. (2008). Senescence-associated Secretory Phenotypes Reveal Cell-Nonautonomous Functions of Oncogenic RAS and the P53 Tumor Suppressor. *PLoS Biol.* 6 (12), 2853–2868. doi:10.1371/journal.pbio.0060301

Cornel, A. M., Mimpen, I. L., and Nierkens, S. (2020). MHC Class I Downregulation in Cancer: Underlying Mechanisms and Potential Targets for Cancer Immunotherapy. *Cancers* 12 (7), 1760. doi:10.3390/cancers12071760

Dai, W., Xiao, Y., Tang, W., Li, J., Hong, L., Zhang, J., et al. (2021). Identification of an EMT-Related Gene Signature for Predicting Overall Survival in Gastric Cancer. *Front. Genet.* 12, 661306. doi:10.3389/fgene.2021.661306

Demirci, D., Dayanc, B., Mazi, F. A., and Senturk, S. (2021). The Jekyll and Hyde of Cellular Senescence in Cancer. *Cells* 10 (2), 208. doi:10.3390/cells10020208

Dersh, D., Hollý, J., and Yewdell, J. W. (2021). A Few Good Peptides: MHC Class I-Based Cancer Immunosurveillance and Immunoevasion. *Nat. Rev. Immunol.* 21 (2), 116–128. doi:10.1038/s41577-020-0390-6

Di Micco, R., Krizhanovsky, V., Baker, D., and d'Adda di Fagagna, F. (2021). Cellular Senescence in Ageing: from Mechanisms to Therapeutic Opportunities. *Nat. Rev. Mol. Cell Biol.* 22 (2), 75–95. doi:10.1038/s41580-020-00314-w

Eusebi, L. H., Telese, A., Marasco, G., Bazzoli, F., and Zagari, R. M. (2020). Gastric Cancer Prevention Strategies: A Global Perspective. *J. Gastroenterol. Hepatol.* 35, 1495–1502. doi:10.1111/jgh.15037

Gorgoulis, V., Adams, P. D., Alimonti, A., Bennett, D. C., Bischof, O., Bishop, C., et al. (2019). Cellular Senescence: Defining a Path Forward. *Cell* 179 (4), 813–827. doi:10.1016/j.cell.2019.10.005

Hsu, A., and Raufi, A. G. (2021). Advances in Systemic Therapy for Gastric Cancer. *Gastrointest. Endosc. Clin. N. Am.* 31 (3), 607–623. doi:10.1016/j.giec.2021.03.009

Joshi, S. S., and Badgwell, B. D. (2021). Current Treatment and Recent Progress in Gastric Cancer. *CA Cancer J. Clin.* 71 (3), 264–279. doi:10.3322/caac.21657

Kawazoe, A., Shitara, K., Boku, N., Yoshikawa, T., and Terashima, M. (2021). Current Status of Immunotherapy for Advanced Gastric Cancer. *Jpn. J. Clin. Oncol.* 51 (1), 20–27. doi:10.1093/jjco/hyaa202

Li, L., Zhu, Z., Zhao, Y., Zhang, Q., Wu, X., Miao, B., et al. (2019). FN1, SPARC, and SERPINE1 are Highly Expressed and Significantly Related to a Poor Prognosis of Gastric Adenocarcinoma Revealed by Microarray and Bioinformatics. *Sci. Rep.* 9 (1), 7827. doi:10.1038/s41598-019-43924-x

Liu, Z., Liu, L., Weng, S., Guo, C., Dang, Q., Xu, H., et al. (2022). Machine Learning-Based Integration Develops an Immune-Derived lncRNA Signature for Improving Outcomes in Colorectal Cancer. *Nat. Commun.* 13 (1), 816. doi:10.1038/s41467-022-28421-6

Liu, Z., Guo, C., Dang, Q., Wang, L., Liu, L., Weng, S., et al. (2022). Integrative Analysis from Multi-Center Studies Identities a Consensus Machine Learning-Derived lncRNA Signature for Stage II/III Colorectal Cancer. *EBioMedicine* 75, 103750. doi:10.1016/j.ebiom.2021.103750

Liu M, M., Kuo, F., Capistrano, K. J., Kang, D., Nixon, B. G., Shi, W., et al. (2020). TGF-β Suppresses Type 2 Immunity to Cancer. *Nature* 587 (7832), 115–120. doi:10.1038/s41586-020-2836-1

Liu SJ, S. J., Yang, Y.-b., Zhou, J.-x., Lin, Y.-j., Pan, Y.-l., and Pan, J.-h. (2021). A Novel Ferroptosis-Related Gene Risk Signature for Predicting Prognosis and

Immunotherapy Response in Gastric Cancer. *Dis. Markers* 2021, 1–18. doi:10.1155/2021/2385406

Liu X, X., Gao, L., Ni, D., Ma, C., Lu, Y., and Huang, X. (2020). Candidate Genes for Predicting the Survival of Patients with Gastric Cancer: A Study Based on the Cancer Genome Atlas (TCGA) Database. *Transl. Cancer Res.* 9 (4), 2599–2608. doi:10.21037/tcr.2020.02.82

Liu Z, Z., Weng, S., Xu, H., Wang, L., Liu, L., Zhang, Y., et al. (2021). Computational Recognition and Clinical Verification of TGF-β-Derived miRNA Signature with Potential Implications in Prognosis and Immunotherapy of Intrahepatic Cholangiocarcinoma. *Front. Oncol.* 11, 757919. doi:10.3389/fonc.2021.757919

Meng, C., Xia, S., He, Y., Tang, X., Zhang, G., and Zhou, T. (2020). Discovery of Prognostic Signature Genes for Overall Survival Prediction in Gastric Cancer. *Comput. Math. Methods Med.* 2020, 1–9. doi:10.1155/2020/5479279

Overacre-Delgoffe, A. E., Bumgarner, H. J., Cillo, A. R., Burr, A. H. P., Tometich, J. T., Bhattacharjee, A., et al. (2021). Microbiota-specific T Follicular Helper Cells Drive Tertiary Lymphoid Structures and Anti-tumor Immunity against Colorectal Cancer. *Immunity* 54 (12), 2812–2824. doi:10.1016/j.immuni.2021.11.003

Prasanna, P. G., Citrin, D. E., Hildesheim, J., Ahmed, M. M., Venkatachalam, S., Riscuta, G., et al. (2021). Therapy-Induced Senescence: Opportunities to Improve Anticancer Therapy. *J. Natl. Cancer Inst.* 113 (10), 1285–1298. doi:10.1093/jnci/djab064

Ramu, D., Shan, T. W., Hirpara, J. L., and Pervaiz, S. (2021). Cellular Senescence: Silent Operator and Therapeutic Target in Cancer. *IUBMB Life* 73 (3), 530–542. doi:10.1002/iub.2460

Sexton, R. E., Al Hallak, M. N., Diab, M., and Azmi, A. S. (2020). Gastric Cancer: A Comprehensive Review of Current and Future Treatment Strategies. *Cancer Metastasis Rev.* 39 (4), 1179–1203. doi:10.1007/s10555-020-09925-3

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Wei, L., Sun, J., Zhang, N., Zheng, Y., Wang, X., Lv, L., et al. (2020). Noncoding RNAs in Gastric Cancer: Implications for Drug Resistance. *Mol. Cancer* 19 (1), 62. doi:10.1186/s12943-020-01185-7

Xia, Y., Rao, L., Yao, H., Wang, Z., Ning, P., and Chen, X. (2020). Engineering Macrophages for Cancer Immunotherapy and Drug Delivery. *Adv. Mat.* 32 (40), 2002054. doi:10.1002/adma.202002054

Yasuda, T., Koiwa, M., Yonemura, A., Miyake, K., Kariya, R., Kubota, S., et al. (2021). Inflammation-driven Senescence-Associated Secretory Phenotype in Cancer-Associated Fibroblasts Enhances Peritoneal Dissemination. *Cell Rep.* 34 (8), 108779. doi:10.1016/j.celrep.2021.108779

Yin, H. M., He, Q., Chen, J., Li, Z., Yang, W., and Hu, X. (2021). Drug Metabolism-related Eight-gene Signature Can Predict the Prognosis of Gastric Adenocarcinoma. *J. Clin. Lab. Anal.* 35 (12), e24085. doi:10.1002/jcla.24085

Zhang, S., Xiang, X., Liu, L., Yang, H., Cen, D., and Tang, G. (2021). Bioinformatics Analysis of Hub Genes and Potential Therapeutic Agents Associated with Gastric Cancer. *Cancer Manag. Res.* 13, 8929–8951. doi:10.2147/CMAR.S341485

Zhang, S.-X., Liu, W., Ai, B., Sun, L.-L., Chen, Z.-S., and Lin, L.-Z. (2022). Current Advances and Outlook in Gastric Cancer Chemoresistance: A Review. *Recent Pat. Anticancer Drug Discov.* 17 (1), 26–41. doi:10.2174/1574892816666210929165729

Zhou, L., Niu, Z., Wang, Y., Zheng, Y., Zhu, Y., Wang, C., et al. (2022). Senescence as a Dictator of Patient Outcomes and Therapeutic Efficacies in Human Gastric Cancer. *Cell Death Discov.* 8 (1), 13. doi:10.1038/s41420-021-00769-6

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 | Frontiers in Genetics

ORIGINAL RESEARCH
published: 15 June 2022
doi: 10.3389/fgene.2022.930248

Exploring the Relationship Between Senescence and Colorectal Cancer in Prognosis, Immunity, and Treatment

Kechen Dong[1†], *Jianping Liu*[2†], *Wei Zhou*[3*] *and Guanglin Zhang*[2*]

[1]*Department of Oncology of Head and Neck, Huangshi Central Hospital (Pu Ai Hospital), Affiliated Hospital of Hubei Polytechnic University, Edong Healthcare Group, Huangshi, China,* [2]*Department of Abdominal and Pelvic Medical Oncology II Ward, Huangshi Central Hospital (Pu Ai Hospital), Affiliated Hospital of Hubei Polytechnic University, Edong Healthcare Group, Huangshi, China,* [3]*Department of Urology, Huangshi Central Hospital (Pu Ai Hospital), Affiliated Hospital of Hubei Polytechnic University, Edong Healthcare Group, Huangshi, China*

Background: Senescence, as an effective barrier against tumorigenesis, plays a critical role in cancer therapy. However, the role of senescence in colorectal cancer (CRC) has not yet been reported. This study aimed to build a prognostic signature for the prognosis of patients with CRC based on senescence-related genes.

Methods: A prognostic signature was built from TCGA based on differentially expressed senescence-related genes by the least absolute shrinkage and selection operator (LASSO) and Cox regression analyses, which were further validated using two Gene Expression Omnibus (GEO) cohorts. The CIBERSORT and ssGSEA algorithms were utilized to analyze the infiltrating abundance of immune cells. The relationship of signature with the immune therapy and the sensitivity of different therapies was explored.

Results: We found 93 genes associated with senescence that were differentially expressed. Based on expression and clinical parameters, we developed a senescence-related prognostic signature and its effectiveness was verified using two external validation cohorts. Overall survival was predicted using a prognostic nomogram that incorporated the predictive values of the risk score and clinical traits. Additionally, the risk score was significantly correlated with immune cells infiltration, tumor immune microenvironment (TME) score, immune checkpoints, immunotherapeutic efficacy, and chemotherapy sensitivity.

Conclusion: The senescence-related prognostic model can well predict the prognosis, immunotherapeutic response, and identify potential drug targets, which can help guide individualized treatment.

Keywords: senescence, colorectal cancer, prognostic model, tumor immune microenvironment, immunothearpy

OPEN ACCESS

Edited by:
Rajkumar S. Kalra,
Okinawa Institute of Science and Technology Graduate University, Japan

Reviewed by:
Bin Fu,
The First Affiliated Hospital of Nanchang University, China
Aditya Sarode,
Columbia University, United States

***Correspondence:**
Wei Zhou
drweizhou@yeah.net
Guanglin Zhang
guanglinzhang@yeah.net

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

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

Received: *27 April 2022*
Accepted: *31 May 2022*
Published: *15 June 2022*

Citation:
Dong K, Liu J, Zhou W and Zhang G (2022) Exploring the Relationship Between Senescence and Colorectal Cancer in Prognosis, Immunity, and Treatment. Front. Genet. 13:930248. doi: 10.3389/fgene.2022.930248

INTRODUCTION

Colorectal cancer (CRC) is one of the most common malignant tumors. Its incidence rate ranks third in the world and the mortality rate is ranked second (Sung et al., 2021). There were nearly 1.9 million (10.0%) new cases of CRC worldwide, followed by breast and lung cancer in incidence (Sung et al., 2021). CRC is a malignant tumor of the digestive system and is the first tumor in the world in terms of morbidity and mortality and seriously threatens the life and health of individuals (Siegel et al., 2020). At present, the main treatment methods for CRC include a combination of endoscopic resection,

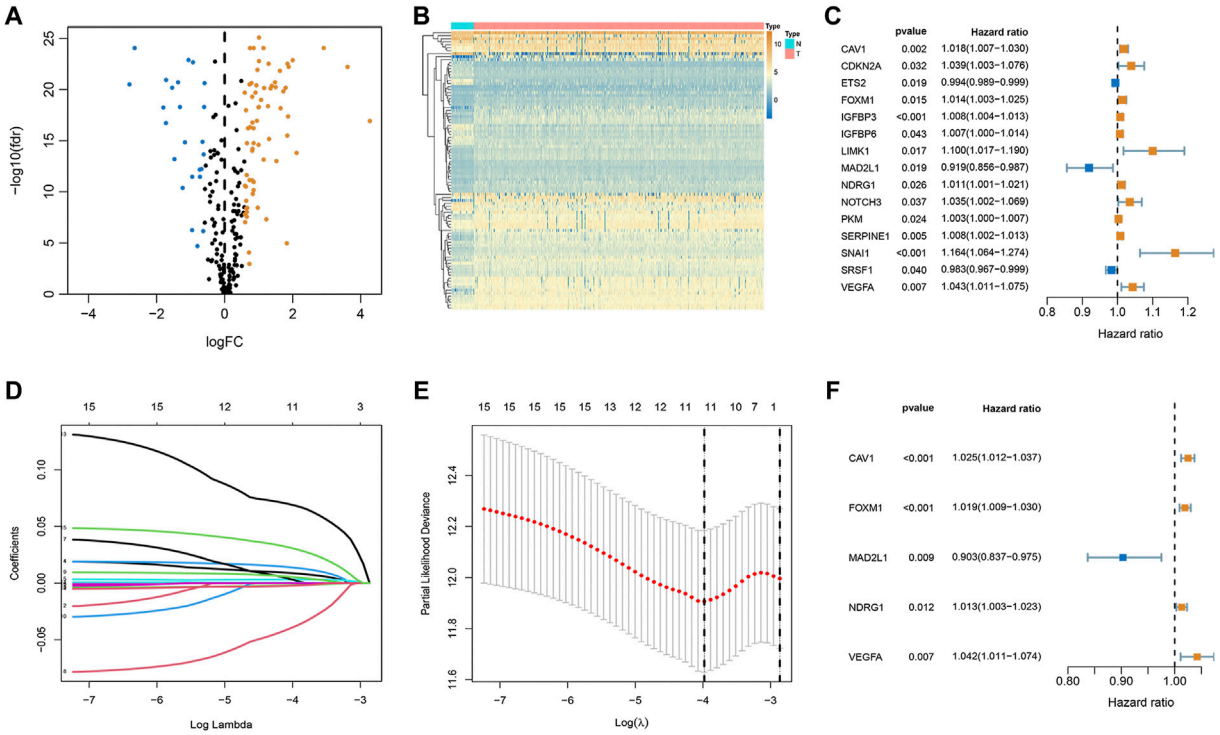


FIGURE 1 | Establishment of optimal senescence-related signature in the TCGA set. **(A,B)** Volcano and heatmap representations of lactate-related DEGs between normal and CRC groups. **(C)** The prognostic genes were selected by univariate Cox regression analysis. **(D,E)** Lasso regression analysis. **(F)** Multivariate Cox regression analyses of the association between genes and OS of patients.

surgical resection, chemotherapy, and radiotherapy (Modest et al., 2019; Dariya et al., 2020). As the surgical intervention is available for early CRC, a large majority of patients with advanced CRC suffer from a poor therapeutic outcome with higher rates of malignant recurrence and distant metastases, resulting in a 5-years survival rate of less than 10% (Chen et al., 2021a). Hence, it is particularly important to find a prognostic model that can accurately classify CRC patients, so that appropriate treatment method can be selected for patients with different precise prognoses.

Cellular senescence is defined as a permanent state of cell cycle termination. It is a response to endogenous and exogenous stresses, including DNA damage, telomere dysfunction, oncogene activation, and organelle stress, and is associated with processes such as tumor suppression, tissue repair, embryogenesis, and organ aging (López-Otín et al., 2013; Di Micco et al., 2021). The current state of aging research shares many similarities with cancer research over the past few decades. In the newly proposed third edition of cancer hallmarks in 2022, four new members have been added, and one of the hallmarks is senescent cells (Hanahan, 2022). Cancer is the result of abnormally enhanced cellular fitness, whereas senescence is characterized by loss of fitness. On the surface, cancer and aging appear to be opposite processes. However, at a deeper level, the two may have a common origin. Cellular senescence is caused by a time-dependent accumulation of cellular damage (Gems and Partridge, 2013). Meanwhile, cell damage occasionally confers abnormal benefits on certain cells, ultimately leading to cancer. Therefore, cancer and aging can be thought of as two distinct manifestations of the same underlying process, the accumulation of cellular damage. Numerous genes have been implicated in cellular senescence as biomarkers and causal drivers (Giovannini et al., 2012; Jia et al., 2018; Li et al., 2019; Shaosheng et al., 2021). Li et al. (Li et al., 2019) found that knockdown of BAZ1A-KD results in up-regulation of SMAD3 expression, which in turn activates transcription of the p21-encoding gene CDKN1A and causes senescence-related phenotypes in human cancer cells. However, it is unknown if these senescence-related genes have an impact on CRC prognosis.

In this study, for the first time, we established a prognostic signature based on differentially expressed senescence-related genes (DEGs) and verified its accuracy in two external databases. Following that, we developed a nomogram to predict the OS of patients with CRC. In addition, we investigated the prognostic value, and impact on tumor immune infiltration, immune checkpoint expression, immunotherapy, and chemotherapeutic drug sensitivity of senescence-related genes in HCC.

MATERIALS AND METHODS

Data Source

A total of 279 senescence-related genes were collected from CellAge database (https://genomics.senescence.info/cells/

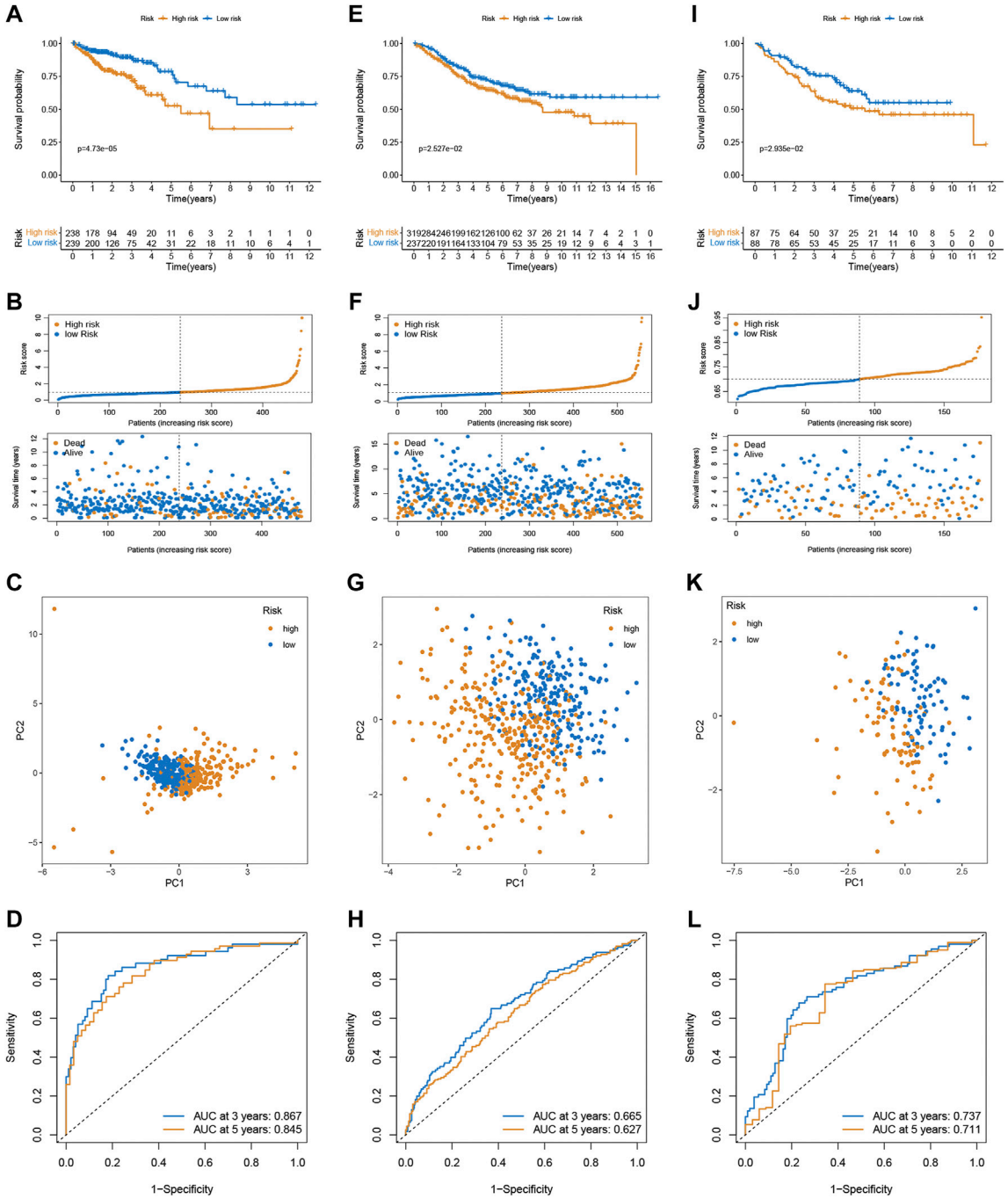


FIGURE 2 | Validation of the prognostic prediction performance of the signature. Kaplan-Meier survival analysis in the training **(A)**, GSE39582 **(E)**, and GSE17536 cohort **(I)**. Distribution of risk score and survival status in the training **(B)**, GSE39582 **(F)**, and GSE17536 cohort **(J)**. PCA analysis in the training **(C)**, GSE39582 **(G)**, and GSE17536 cohort **(K)**. Time-dependent ROC curves of risk scores in the training **(D)**, GSE39582 **(H)**, and GSE17536 cohort **(L)**.

signatures.php?) and were listed in **Supplementary Table S1**. The RNA-seq expression and clinical traits for CRC patients were obtained and extracted from three independent CRC cohorts (TCGA-COD, n = 477; GSE39582, n = 556; GSE17536, n = 175). IMvigor210 with immunotherapy data and clinical information were obtained from the IMvigor210CoreBiologies R package. TCGA cohort was used to build the signature, and two GEO cohorts were used to externally verify the signature.

Establishment and Identification of Prognostic Signature

The training cohort was employed to detect the senescence-related DEGs between normal and CRC tissues via the R package "limma" in RStudio, with the following cutoff for adjustment: p-value < 0.05 and | fold change (FC)| > 1.5. To screen senescence-related genes with prognostic significance, a univariate Cox regression analysis was conducted on DEGs. Following that, the Least absolute shrinkage and selection operator (LASSO) and multivariable Cox analysis was performed to build a predictive signature. The following formula was employed to calculate the risk scores of CRC samples:

$$Rish\,Score = \sum_{i=1}^{n} coef\,(Xi) \times \exp\ (Xi)$$

"Coef", "exp", and "n" represented the coefficient of the gene, the expression level, and the number of genes, respectively. The median risk score was used as the threshold. Patients with risk scores greater than the threshold (median risk score) were included in the high-risk group and the rest in the low-risk group. Receiver operating characteristics (ROC) and Kaplan-Meier survival curves were employed to assess the effectiveness of the signature. Principal component analysis (PCA) was conducted to verify whether the risk score could distinguish high- and low-risk score groups.

Two GEO validation cohorts were recruited to vertify the predictive accuracy of the model developed from the TCGA set. The above cut-off value was used to divide all CRC patients into high- and low-risk score groups, the same method was employed for the predictive power of the signature in OS prediction.

Nomogram Construction and Assessment

We explored the risk score with the corresponding CRC samples' clinical information, including age, gender, tumor site, and TNM stage. Additionally, we also explored whether the risk levels would affect the prognosis of patients in distinct clinical variable groups. Univariate and multivariate models were employed to ascertain whether the signature could be an independent predictive indicator for the prognosis of CRC patients. Then, a nomogram integrating risk score and clinical parameters was built using the "rms" R packages. ROC and calibration curves were employed to validate its accuracy is demonstrated.

Immune Activities Analysis

The ssGSEA algorithm was used to quantify the scores of 16 tumor immune infiltration cells (TIICs) and the function of 13 immune-related pathways. The proportion of 22 TIICs in two risk score groups was further quantified with CIBERSORT algorithm. The immune score, stromal score, ESTIMATE score were calculated through ESTIMATE algorithm to quantify the tumors microenvironment (Arbour et al., 2021). Two immune checkpoints (PD-1 and PD-L1) were chosen to assess the differences in their expression levels in two risk subgroups.

Targeted Drug and Immunotherapy

In this study, the capability of the signature in predicting sensitivity of chemotherapy and immunotherapy was investigated. In the aspect of chemotherapy, half maximal inhibitory concentration (IC50) was used to predict the sensitivity of chemotherapy drugs in the high- and low-risk groups. Meanwhile, potential immune checkpoint inhibitors (ICIs) response was predicted with TIDE algorithm (Jiang et al., 2018).

Gene Set Enrichment Analysis

To investigate the biological pathways of the subgroups, we further generated a gene set enrichment analysis (GSEA) for functional enrichment analysis. Gene sets with p-value and Q-value < 0.05 were the cutoff criterion for significant gene enrichment.

Statistical Analysis

Data were analyzed by R software version 4.1.0. Log-rank test was used for survival analysis. Wilcoxon rank-sum or Kruskal–Wallis tests were utilized to compare differences between two or three groups, respectively. The ROC curves were plotted to access the prognostic value of the model.

RESULTS

PPI Network and GO and KEGG Enrichment Analyses

Among 279 senescence-related genes, 93 DEGs in CRC patients of the TCGA cohort were identified with FDR <0.05 and FC > 1.5. Volcano and heatmap representations of senescence-related DEGs are provided in **Figures 1A,B**. Then, Protein-protein interaction (PPI) networks and functional enrichment analyses were constructed to comprehensively investigate these DEGs. As shown by PPI analysis, 84 of these 93 DEGs formed interaction modules (**Supplementary Figure S1A**). By using the cytoHubba plugin, we screened 10 hubgenes (**Supplementary Figure S1B**). GO functional annotation showed that the 93 DEGs are mainly related to regulation of cell cycle phase transition, transcription regulator complex, and DNA-binding transcription factor binding (**Supplementary Figure S1C**). KEGG signaling enrichment annotation showed that these DEGs are mainly enriched in the cell cycle, cellular senescence, p53 signaling pathway, and other tumor-related signal pathways (**Supplementary Figure S1D**).

Establishment and Validation of the Senescence-Related Signature

To explore whether these senescence-related genes are related to the prognosis of CRC, univariate COX regression analysis was applied. Based on the TCGA cohort, 15 genes were identified (**Figure 1C**). As shown in **Figures 1D,E**, 15 genes were subject to LASSO Cox regression analysis to avoid overfitting, and 11 out of 15 genes were chosen as the appropriate candidates for constructing a risk signature. Subsequently, multivariate Cox regression analysis obtained 5 genes (CAV1, FOXM1, MAD2L1, NDRG1, and VEGFA) to build a prognostic signature

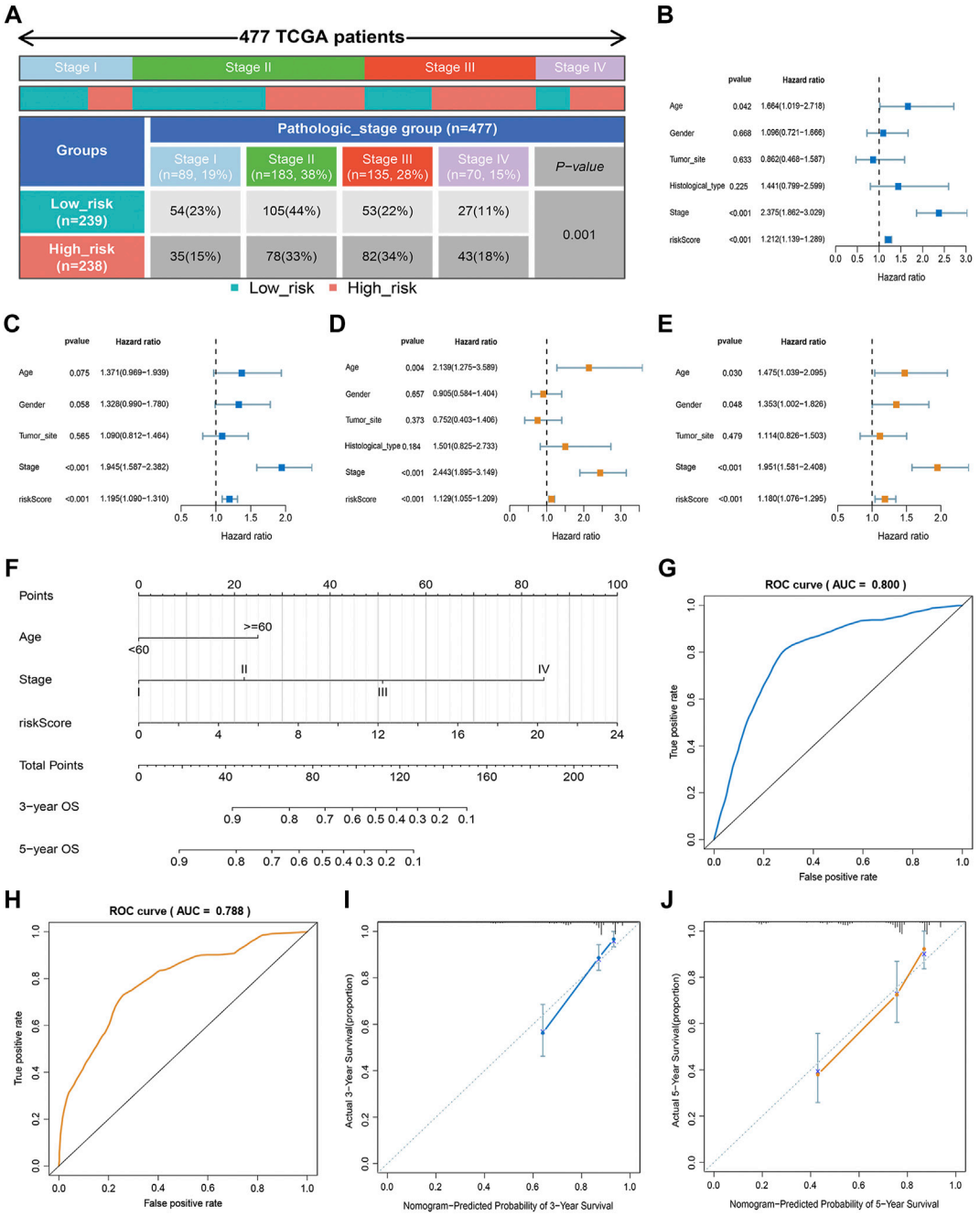


FIGURE 3 | Clinical value of the risk score. **(A)** Association of risk score with TNM stage. **(B,C)** Univariate analysis of risk scores and clinicopathological parameters in the training **(B)** and GSE39582 cohort **(C)**. **(D,E)** Multivariate Cox regression analysis of risk scores and clinicopathological parameters in the training **(D)** and GSE39582 cohort **(E)**. **(F)** Prediction of the nomogram based on clinical traits and risk score. **(G,H)** ROC curves of the nomogram for OS prediction at three **(G)** and 5 years **(H)**. **(I,J)** Calibration curve of the nomogram for predicting OS rates at three **(I)** and 5 years **(J)**.

(**Figure 1F**). Risk score = expression (CAV1) × 0.024 + expression (FOXM1) × (0.019) + expression (MAD2L1) × (-0.102) + expression (NDRG1) × (0.012) + expression (VEGFA) × (0.041). Median risk scores divided the cohort of CRC patients into the low- and high-risk subgroups. To analyze the translational levels of the signature genes, the Human Protein Atlas (HPA) database can be used, showing the expression and localization of the corresponding protein. The results showed that FOXM1, MAD2L1, NDRG1, and VEGFA was highly expressed in CRC tissue, wihle CAV1 was lowly expressed in CRC tissue (**Supplementary Figure S2**).

Internal and External Validation of the Signature

In the training cohort, the Kaplan-Meier analysis revealed that high-risk group had lower OS compared to low-risk group ($p < 0.001$, **Figure 2A**). Also, mortality was increased in CRC patients with increasing risk scores (**Figure 2B**). PCA analysis revealed that there was a clear division in two risk subgroups (**Figure 2C**). ROC plots were also used to assess diagnostic efficiency, with AUCs of 0.867 and 0.845 for 3 and 5-years survival, respectively (**Figure 2D**).

To confirm the robustness of the signature, the risk scores of CRC patients were calculated in two external validation sets (GSE39582 and GSE17536) using the same formula, and divided patients into the high- and low-risk subgroups according to the cutoff of the training cohort. Likewise, high-risk was associated with OS (**Figures 2E,I**), and the number of deaths increased with increasing risk scores (**Figures 2F,J**). PCA demonstrated overt separation of both subgroups (**Figures 2G,K**). The ROC further indicated the predicting accuracy of the signature (**Figures 2H, 2**L). Additionally, the Imvigor210 dataset of the treatment response data of patients who underwent anti-PD-L1 immunotherapy was retrieved to validate the predictive ability of the senescence-based signature in ICI therapy. Kaplan-Meier analysis showed that a high-risk score was associated with a poorer survival rate than a low risk score (**Supplementary Figure S3**).

Prognostic Value of the Signature

In our study, we analyzed the prognosis of patients in low- and high-risk groups among distinct clinical variable subgroups. As shown in **Supplementary Figure S4**, patients with high-risk scores had poorer survival probabilities than those with low-risk scores in all distinct clinical variable subgroups. In addition, we further investigated the association between risk scores and each clinical characteristic. The results showed that the risk score was linked to the TNM stage ($p < 0.01$; **Figure 3A**). Subsequently, we verified the independence and applicability of the risk score in the training and GSE39582 sets. Univariate and multivariate Cox regression analysis results showed that the signature could independently predict the prognosis of CRC patients, regardless of age, gender, tumor site, histological type, and TNM stage ($p < 0.001$, **Figures 3B–E**).

Development and Assessment of the Nomogram

An approach by which 3- and 5-years OS rates could be more accurately predicted was to construct a nomogram model based on Cox regression results (**Figure 3F**), which included risk score, age, and TNM stage. As shown in **Figure 3F**, this nomogram can predict the 3- and 5-survival for a patient based on the sum of the scores. The ROC curve revealed the high accuracy of the nomogram for 3-years (AUC = 0.80) and 5 -year (AUC = 0.788) survival rates (**Figures 3G,H**). The calibration curves comparing the predicted and actual survival rates of CRC patients indicated that the predicted survival rates were in good agreement with those actual rates (**Figures 3I,J**).

Correlations Between the Risk Scores and TME

To better investigate the relationship between risk score and immune characteristics, ssGSEA was used to calculate the enrichment scores of various immune cells. According to **Figure 4A**, the relative scale of fraction for $CD8^+$ T cells and NK cells was obviously lower in the high-risk group than that in the low-risk group. On the contrary, the fraction of macrophages and T helper cells were much lower in the low-risk group. We also found substantial variations in immune function in terms of T cell co-stimulation, type I IFN response, and type II IFN response (**Figure 4B**). Furthermore, CIBERSORT algorithms were employed to calculate the scores of various TIICs. Results suggested that the infiltration abundance of $CD8^+$ T cells, memory activated $CD4^+$ T cells, macrophages M1, naive B cells, and resting dendritic cells in the high-risk group was obviously lower than that in the low-risk group, and their infiltration abundance decreased with increasing risk score (**Figure 4C**). However, the infiltrative abundance of M2 macrophages, T cells regulatory (Tregs), and Tfh cells was distinctly higher in the high-risk group, and their abundance increased prominently with risk scores increased (**Figure 4C**). In addition, patients with a low-risk score presented a higher level of the immune score, stromal score, and ESTIMATE score than those with a high-risk score (**Figure 4D**).

Relationship Between the Signature and CRC Therapy

Given the significance of checkpoint treatment, we investigate more into the variations in immune checkpoint expression between different risk subgroups. The results indicated that the expression of PD-1 and PD-L1 in the low-risk group were higher than those in the low-risk group (**Figures 5A,B**). Furthermore, we applied the TIDE algorithms to evaluate the effectiveness of the signatures in forecasting ICIs responsiveness in CRC. TIDE scores were higher in the high-risk score group compared to the low-risk group (**Figure 5C**). Taken together, the signature can predict the benefit of CRC immunotherapy.

Chemotherapeutic drug sensitivity analysis will help guide the optimal selection of commonly used chemotherapeutic drugs for CRC. By comparing IC50 values in high- and low-risk groups, Wilcoxon signed-rank test was used to evaluate chemosensitivity. The result indicated that the patients with low-risk scores were more sensitive to cisplatin, docetaxel, gemcitabine, epothilone B, and Metformin, while patients with the high-risk score were more sensitive to nilotinib, saracatinib (AZD0530), dasatinib, and imatinib (**Figure 5D**-L).

GSEA Enrichment Analysis

To clarify the important pathway of signature enrichment related to pyroptosis, we conducted GSEA. As shown in **Supplementary Table S2**, 55 enrichment pathways with significant variations between low and high-risk groups were identified at the criteria of FDR <0.25, $p < 0.05$. The top five signaling pathways in the high-risk group were axon guidance, complement and coagulation cascades, ECM receptor interaction, focal adhesion, and hematopoietic cell lineage (**Figure 6A**). The top five signaling

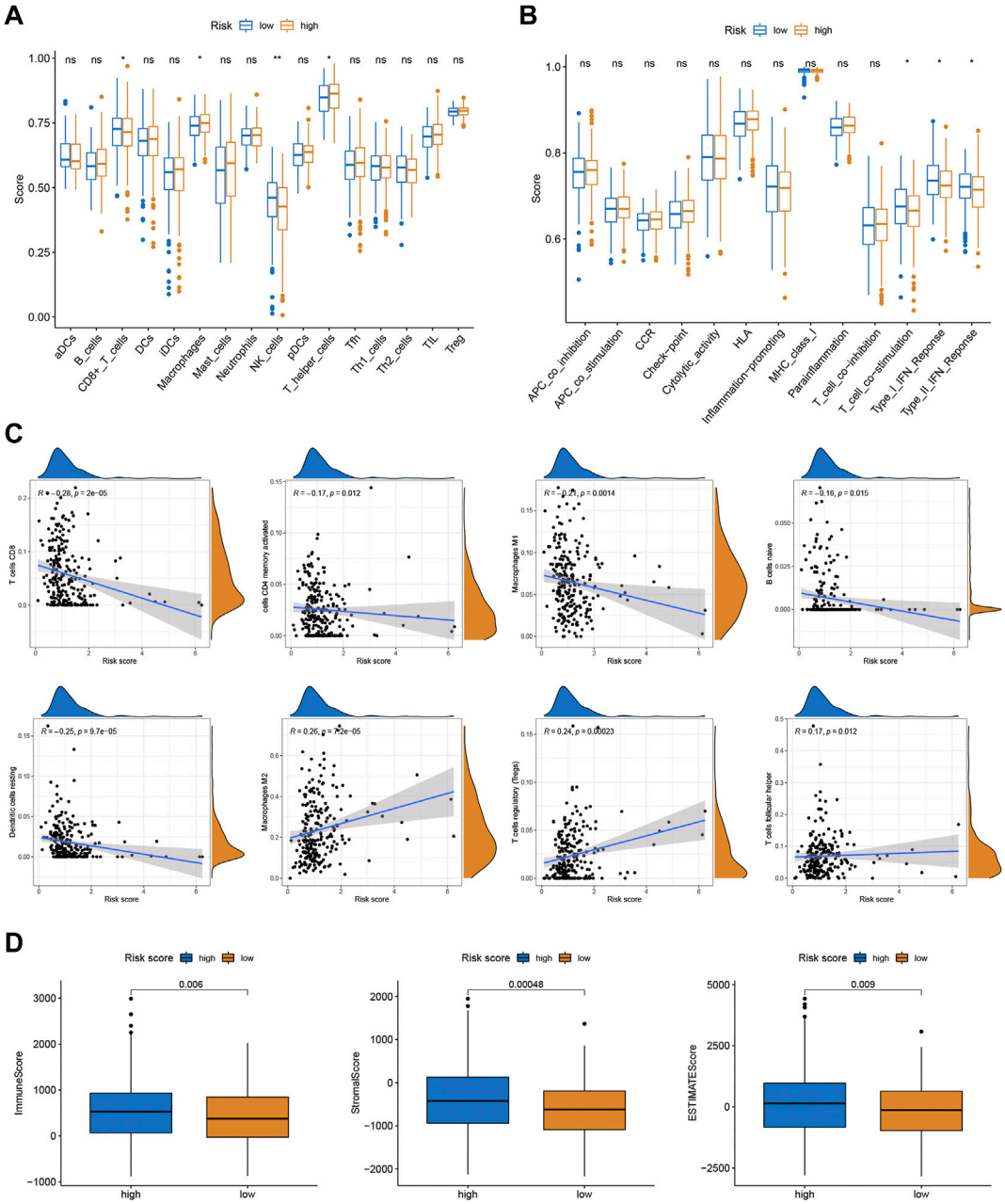


FIGURE 4 | Correlations between the risk scores and TME. **(A)** The ssGSEA scores of immune infiltrating cells. **(B)** The ssGSEA scores of immune functions. **(C)** The proportion of 22 immune infiltrating cells in two risk subgroups. **(D)** TME score in two risk subgroups.

pathways in the low-risk group were huntingtons disease, oxidative phosphorylation, parkinsons disease, proteasome, and ribosome (**Figure 6B**).

DISCUSSION

CRC is a highly heterogeneous disease, and survival time varies widely among patients with similar clinical stages. Cellular senescence is recognized as an effective barrier against tumorigenesis and can be promoted by immune surveillance (Ou et al., 2021). Most research on cellular senescence has focused on non-tumor cells, but tumor cells can also undergo senescence. The treatment of cancer consisting of pro-senescence and senolytic therapy has also been explored, which is expected to become new approaches for targeted therapy of cancer (Wang et al., 2022). Increasing evidence suggests that senescent cells can be eliminated by senescence-associated secretory phenotype (SASP)-elicited immune responses involving both innate and adaptive

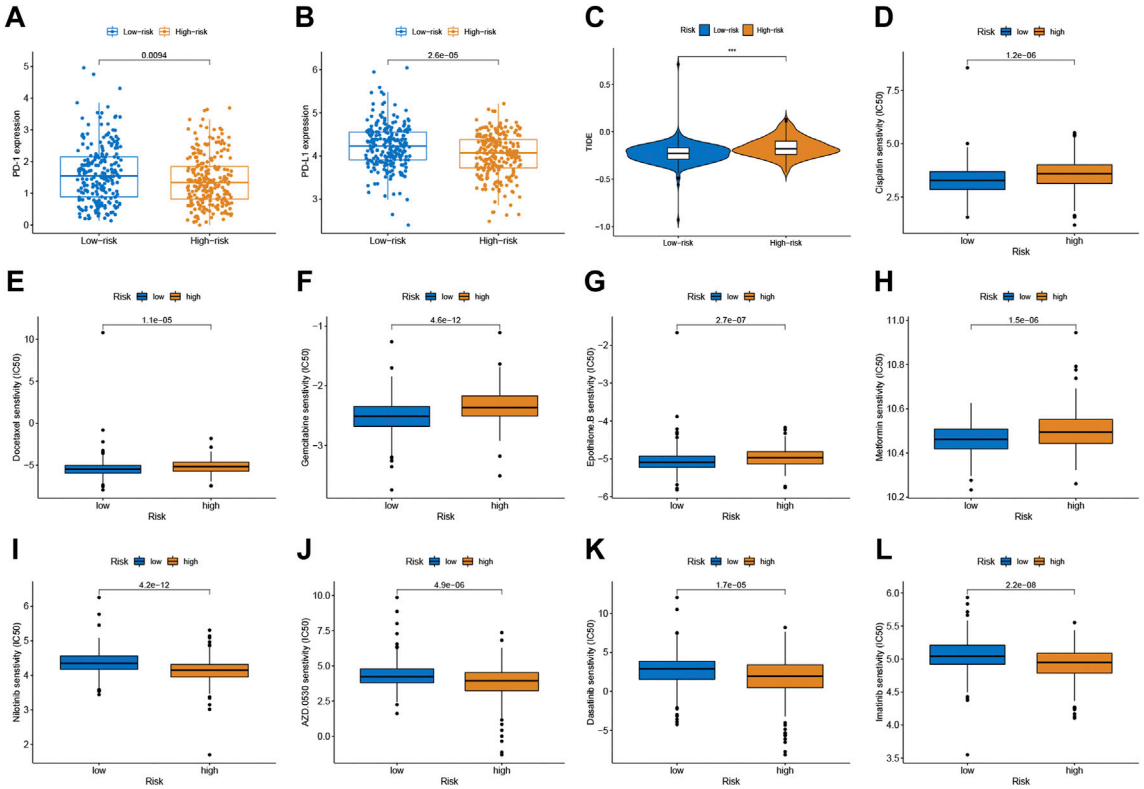


FIGURE 5 | Correlation between the predictive signature and CRC therapy. **(A,B)** The expression value of PD-1 and PD-L1 between two risk subgroups. **(C)** Comparison of TIDE score between low- and high-risk subgroups. **(D–L)** Estimated drug sensitivity in patients with high- and low-risk subgroups.

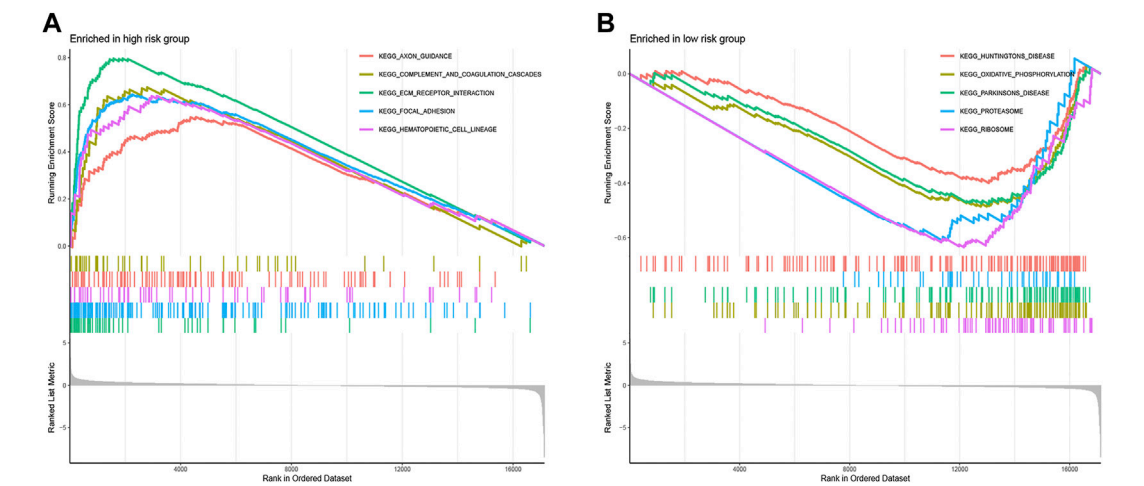


FIGURE 6 | Functional enrichment analysis between low- and high-risk groups. The top five signaling pathways in the high- **(A)** and low-risk subgroup **(B)**.

immunity, so activation of the host immune system is a particularly attractive approach to clearing senescent cancer cells (Schneider et al., 2021; Wang et al., 2022). However, the correlation between cellular senescence and TME remains unclear, and the value of cellular senescence-related genes in assessing immune infiltration and clinical outcome in CRC has not been reported. Therefore, this study aimed to establish a new prognostic signature based on senescence-related genes to help accurately predict the prognosis of CRC patients and guide individualized treatment.

In this work, we analyzed the role of senescence in CRC using the public databases. And 93 differentially expressed senescence-related genes were identified between CRC and normal samples. To comprehensively explore the mechanism of senescence in CRC, we performed univariate Cox regression analysis and LASSO Cox regression analysis on these DEGs to develop a senescence-related signature in the training cohort. The signature contained five senescence-associated genes: CAV1, FOXM1, MAD2L1, NDRG1, and VEGFA. CAV1 (caveolin-1) is a key structural component of caveolae and plays an important role in a variety of cellular processes including cholesterol homeostasis, vesicle transport, and tumor progression (Ha and Chi, 2012). CAV1 has been shown to play a dual role in tumorigenesis, inhibiting or promoting tumor growth depending on the cellular context (Ha and Chi, 2012; Kamposioras et al., 20221080). Several studies have reported the effect of CAV1 expression on CRC, but there were no consistent results (Alshenawy and Ali, 2013; Xue et al., 2015; Zhao et al., 2015). Typically, CAV1 expression is elevated in CRC tissue compared to adjacent normal tissue (Alshenawy and Ali, 2013; Xue et al., 2015). CAV1 expression was associated with clinicopathological traits and prognosis of CRC patients (Xue et al., 2015; Yang et al., 2018). CAV1 can affect the occurrence and development of CRC through different mechanisms, including via activation of SLC2A3/GLUT3 transcription (Ha and Chi, 2012), suppressing phosphorylation of epidermal growth factor receptor (Yang et al., 2018), and stimulating HMGA1-mediated GLUT3 transcription (Ha et al., 2012). FOXM1, a member of FOX superfamily, has been implicated in CRC progression and chemoresistance (Varghese et al., 2019; Yang et al., 2019; Yang et al., 2020). Yang et al. (Yang et al., 2019) revealed that FOXM1 expression significantly elavated in CRC tissues and was positively linked to tumor size, TNM stage, lymphatic and distant metastasis. Overexpression of FOXM1 promoted oncogenic effects on CRC by activating the β-catenin signaling pathway. Varghese et al. (Varghese et al., 2019) showed that FOXM1 regulates 5-FU resistance in CRC by regulating TYMS expression. Yang et al. (Yang et al., 2020) FOXM1 simultaneously promote migration, invasion, and drug resistance of CRC cells through upregulating Snail. MAD2L1, as a member of the spindle checkpoint functional complex, plays a crucial role in cell cycle regulation (Zhong et al., 2015). MAD2L1 has been reported as a novel oncogene that plays a role in regulating cancer cell growth and apoptosis (Ding et al., 2020; Ding et al., 2022). NDRG1 has been reported to act as a metastasis suppressor (Bae et al., 2013; Sahni et al., 2014). A recent study shows that NDRG1 regulates filopodia-induced CRC invasiveness by regulating CDC42 activity (Aikemu et al., 2021). VEGFA is an endothelial growth factor and regulator of vascular permeability (Claesson-Welsh and Welsh, 2013). Increasing evidence suggests that VEGFA-dependent signaling pathways play crucial roles in CRC progression (Terme et al., 2013; Dai et al., 2020; Liu et al., 2020).

Furthermore, all CRC patients were categorized into low- and high-risk subgroups depending upon the median value. Internal and external validation results showed that risk scores independently and effectively predicted 3- and 5-years survival in CRC patients. We also conducted univariate and multivariate Cox analyses to explore the effectiveness of the signature and clinical parameters as indicators of patient prognosis. It was concluded that the risk score served as an independent prognostic predictor for CRC patients. To better quantify the 3- and 5-years survival of CRC samples, a nomogram, combined with these independent indicators, was constructed. The predictive accuracy of the nomogram was verified by the ROC curve and calibration plot. Therefore, it may be used as a supplementary tool to better assist the prognosis evaluation and treatment of CRC.

We calculated the infiltration of immune cells and TME scores in the high-and low-risk groups. The ssGSEA and CIBERSORT results showed the risk score was closely related to the relative contents of TIICs, especially for T cells and macrophages. And with the increase of the risk score in the prognostic signature, relative contents of $CD8^+$ T cells tended to be downregulated, while the relative contents of macrophages tended to be upregulated. This discovery is in line with prior research that intratumoral T cell density has been shown to be an independent prognostic factor in CRC (Galon et al., 2006; Miller et al., 2021). $CD8^+$ T cells are considered major drivers of anti-tumor immunity (van der Leun et al., 2020). Accumulating evidence suggests that increased tumor-related macrophage infiltration results in a poor prognosis in CRC (Wei et al., 2019). Tumor-associated macrophage-induced immune responses were already considered critical determinants of tumor progression (Pan et al., 2020). Tumor-associated macrophages can also perform pre-tumor activities such as enhancing tumor cell proliferation, and invasion, angiogenesis, and inhibiting anti-tumor immune surveillance (Chen et al., 2021b; Boutilier and Elsawa, 2021). Also, patients with low risk score have a higher TME score than those with high risk score.

Emerging therapeutic strategies, including PD-1/PD-L1 inhibitors, are used for treating CRC (Yaghoubi et al., 2019). In our study, the expression levels of PD-1 and PD-L1 in the low-risk group were higher compared to those in the high-risk group, which implied that the signature would be able to predict their expression levels and provide guidance during immunotherapy with ICIs. Furthermore, we found that patients with high-risk scores had a higher TIDE score than those with the low-risk score. A lower TIDE score indicates a lower possibility of tumor immune evasion and may benefit from immunotherapy, which further explains the better prognosis of patients in the low-risk group in our study. These findings provide a basis for a more comprehensive understanding of anti-tumor immune responses in CRC patients, as well as guidance for personalized immunotherapy treatments. Chemotherapy and immunotherapy are the most important adjuvant therapies for CRC, which are of great significance for improving both the prognosis of patients and their quality of life. Patients with low risk score was more sensitive to cisplatin, docetaxel, gemcitabine, epothilone B, and Metformin, while patients with the high-risk score were more sensitive to nilotinib, saracatinib, dasatinib, and imatinib. The combination of chemotherapy and immunotherapy can provide precise and individualized therapy for patients with a different risk scores.

CONCLUSION

This study successfully constructed a 5-gene senescence-related signature that could be used to classify CRC patients. The prognostic model shows the convincing clinical value and may provide new ideas for improving the OS rate of CRC patients and facilitating personalized treatment.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The TCGA and GEO databases can be found here: https://portal.gdc.cancer.gov/ and https://www.ncbi.nlm.nih.gov/geo/, respectively.

AUTHOR CONTRIBUTIONS

KCD, WZ, and GLZ designed this study. KCD and JPL took part in the data collection. KCD, JPL, and WZ analyzed and visualized data. KCD, JPL, and WZ drafted the manuscript, and GLZ revised the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.930248/full#supplementary-material

REFERENCES

Aikemu, B., Shao, Y., Yang, G., Ma, J., Zhang, S., Yang, X., et al. (2021). NDRG1 Regulates Filopodia-Induced Colorectal Cancer Invasiveness via Modulating CDC42 Activity. *Int. J. Biol. Sci.* 17, 1716–1730. doi:10.7150/ijbs.56694

Alshenawy, H. A., and Ali, M. A. E.-H. A. E. -A. (2013). Differential Caveolin-1 Expression in Colon Carcinoma and its Relation to E-Cadherin-β-Catenin Complex. *Ann. Diagnostic Pathology* 17, 476–482. doi:10.1016/j.anndiagpath.2013.05.007

Arbour, K. C., Luu, A. T., Luo, J., Rizvi, H., Plodkowski, A. J., Sakhi, M., et al. (2021). Deep Learning to Estimate RECIST in Patients with NSCLC Treated with PD-1 Blockade. *Cancer Discov.* 11, 59–67. doi:10.1158/2159-8290.Cd-20-0419

Bae, D.-H., Jansson, P. J., Huang, M. L., Kovacevic, Z., Kalinowski, D., Lee, C. S., et al. (2013). The Role of NDRG1 in the Pathology and Potential Treatment of Human Cancers. *J. Clin. Pathol.* 66, 911–917. doi:10.1136/jclinpath-2013-201692

Boutilier, A. J., and Elsawa, S. F. (2021). Macrophage Polarization States in the Tumor Microenvironment. *Ijms* 22, 6995. doi:10.3390/ijms22136995

Chen, D., Zhang, X., Li, Z., and Zhu, B. (2021). Metabolic Regulatory Crosstalk between Tumor Microenvironment and Tumor-Associated Macrophages. *Theranostics* 11, 1016–1030. doi:10.7150/thno.51777

Chen, K., Collins, G., Wang, H., and Toh, J. W. T. (2021). Pathological Features and Prognostication in Colorectal Cancer. *Curr. Oncol.* 28, 5356–5383. doi:10.3390/curroncol28060447

Claesson-Welsh, L., and Welsh, M. (2013). VEGFA and Tumour Angiogenesis. *J. Intern Med.* 273, 114–127. doi:10.1111/joim.12019

Dai, J., Zhuang, Y., Tang, M., Qian, Q., and Chen, J. P. (2020). CircRNA UBAP2 Facilitates the Progression of Colorectal Cancer by Regulating miR-199a/VEGFA Pathway. *Eur. Rev. Med. Pharmacol. Sci.* 24, 7963–7971. doi:10.26355/eurrev_202008_22479

Dariya, B., Aliya, S., Merchant, N., Alam, A., and Nagaraju, G. P. (2020). Colorectal Cancer Biology, Diagnosis, and Therapeutic Approaches. *Crit. Rev. Oncog.* 25, 71–94. doi:10.1615/CritRevOncog.2020035067

Di Micco, R., Krizhanovsky, V., Baker, D., and d'Adda di Fagagna, F. (2021). Cellular Senescence in Ageing: from Mechanisms to Therapeutic Opportunities. *Nat. Rev. Mol. Cell Biol.* 22, 75–95. doi:10.1038/s41580-020-00314-w

Ding, X., Duan, H., and Luo, H. (2020). Identification of Core Gene Expression Signature and Key Pathways in Colorectal Cancer. *Front. Genet.* 11, 45. doi:10.3389/fgene.2020.00045

Ding, X., Fu, Q., Chen, W., Chen, L., Zeng, Q., Zhang, S., et al. (2022). Targeting of MAD2L1 by miR-515-5p Involves the Regulation of Cell Cycle Arrest and Apoptosis of Colorectal Cancer Cells. *Cell Biol. Int.* 46, 840–848. doi:10.1002/cbin.11774

Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., et al. (2006). Type, Density, and Location of Immune Cells within Human Colorectal Tumors Predict Clinical Outcome. *Science* 313, 1960–1964. doi:10.1126/science.1129139

Gems, D., and Partridge, L. (2013). Genetics of Longevity in Model Organisms: Debates and Paradigm Shifts. *Annu. Rev. Physiol.* 75, 621–644. doi:10.1146/annurev-physiol-030212-183712

Giovannini, C., Gramantieri, L., Minguzzi, M., Fornari, F., Chieco, P., Grazi, G. L., et al. (2012). CDKN1C/P57 Is Regulated by the Notch Target Gene Hes1 and Induces Senescence in Human Hepatocellular Carcinoma. *Am. J. Pathology* 181, 413–422. doi:10.1016/j.ajpath.2012.04.019

Ha, T.-K., and Chi, S.-G. (2012). CAV1/caveolin 1 Enhances Aerobic Glycolysis in Colon Cancer Cells via Activation of SLC2A3/GLUT3 Transcription. *Autophagy* 8, 1684–1685. doi:10.4161/auto.21487

Ha, T.-K., Her, N.-G., Lee, M.-G., Ryu, B.-K., Lee, J.-H., Han, J., et al. (2012). Caveolin-1 Increases Aerobic Glycolysis in Colorectal Cancers by Stimulating HMGA1-Mediated GLUT3 Transcription. *Cancer Res.* 72, 4097–4109. doi:10.1158/0008-5472.Can-12-0448

Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 12, 31–46. doi:10.1158/2159-8290.Cd-21-1059

Jia, Q., Nie, H., Wan, X., Fu, H., Yang, F., Li, Y., et al. (2018). Down-regulation of Cancer-Associated Gene CDC73 Contributes to Cellular Senescence. *Biochem. Biophysical Res. Commun.* 499, 809–814. doi:10.1016/j.bbrc.2018.03.228

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1

Kamposioras, K., Vassilakopoulou, M., Anthoney, A., Bariuoso, J., Mauri, D., Mansoor, W., et al. (202210802). Prognostic Significance and Therapeutic Implications of Caveolin-1 in Gastrointestinal Tract Malignancies. *Pharmacol. Ther.* 233, 108028. doi:10.1016/j.pharmthera.2021.108028

Li, X., Ding, D., Yao, J., Zhou, B., Shen, T., Qi, Y., et al. (2019). Chromatin Remodeling Factor BAZ1A Regulates Cellular Senescence in Both Cancer and Normal Cells. *Life Sci.* 229, 225–232. doi:10.1016/j.lfs.2019.05.023

Liu, H., Li, A., Sun, Z., Zhang, J., and Xu, H. (2020). Long Non-coding RNA NEAT1 Promotes Colorectal Cancer Progression by Regulating miR-205-5p/VEGFA axis. *Hum. Cell* 33, 386–396. doi:10.1007/s13577-019-00301-0

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The Hallmarks of Aging. *Cell* 153, 1194–1217. doi:10.1016/j.cell.2013.05.039

Miller, T. J., Anyaegbu, C. C., Lee-Pullen, T. F., Spalding, L. J., Platell, C. F., and McCoy, M. J. (2021). PD-L1+ Dendritic Cells in the Tumor Microenvironment Correlate with Good Prognosis and CD8+ T Cell Infiltration in Colon Cancer. *Cancer Sci.* 112, 1173–1183. doi:10.1111/cas.14781

Modest, D. P., Pant, S., and Sartore-Bianchi, A. (2019). Treatment Sequencing in Metastatic Colorectal Cancer. *Eur. J. Cancer* 109, 70–83. doi:10.1016/j.ejca.2018.12.019

Ou, H. L., Hoffmann, R., González-López, C., Doherty, G. J., Korkola, J. E., and Muñoz-Espín, D. (2021). Cellular Senescence in Cancer: from Mechanisms to Detection. *Mol. Oncol.* 15, 2634–2671. doi:10.1002/1878-0261.12807

Pan, Y., Yu, Y., Wang, X., and Zhang, T. (2020). Tumor-Associated Macrophages in Tumor Immunity. *Front. Immunol.* 11, 583084. doi:10.3389/fimmu.2020.583084

Sahni, S., Bae, D.-H., Lane, D. J. R., Kovacevic, Z., Kalinowski, D. S., Jansson, P. J., et al. (2014). The Metastasis Suppressor, N-Myc Downstream-Regulated Gene

1 (NDRG1), Inhibits Stress-Induced Autophagy in Cancer Cells. *J. Biol. Chem.* 289, 9692–9709. doi:10.1074/jbc.M113.529511

Schneider, J. L., Rowe, J. H., Garcia-de-Alba, C., Kim, C. F., Sharpe, A. H., and Haigis, M. C. (2021). The Aging Lung: Physiology, Disease, and Immunity. *Cell* 184, 1990–2019. doi:10.1016/j.cell.2021.03.005

Shaosheng, W., Shaochuang, W., Lichun, F., Na, X., and Xiaohong, Z. (2021). ITPKA Induces Cell Senescence, Inhibits Ovarian Cancer Tumorigenesis and Can Be Downregulated by miR-203. *Aging* 13, 11822–11832. doi:10.18632/aging.202880

Siegel, R. L., Miller, K. D., Goding Sauer, A., Fedewa, S. A., Butterly, L. F., Anderson, J. C., et al. (2020). Colorectal Cancer Statistics, 2020. *CA A Cancer J. Clin.* 70, 145–164. doi:10.3322/caac.21601

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Terme, M., Pernot, S., Marcheteau, E., Sandoval, F., Benhamouda, N., Colussi, O., et al. (2013). VEGFA-VEGFR Pathway Blockade Inhibits Tumor-Induced Regulatory T-Cell Proliferation in Colorectal Cancer. *Cancer Res.* 73, 539–549. doi:10.1158/0008-5472.Can-12-2325

van der Leun, A. M., Thommen, D. S., and Schumacher, T. N. (2020). CD8+ T Cell States in Human Cancer: Insights from Single-Cell Analysis. *Nat. Rev. Cancer* 20, 218–232. doi:10.1038/s41568-019-0235-4

Varghese, V., Magnani, L., Harada-Shoji, N., Mauri, F., Szydlo, R. M., Yao, S., et al. (2019). FOXM1 Modulates 5-FU Resistance in Colorectal Cancer through Regulating TYMS Expression. *Sci. Rep.* 9, 1505. doi:10.1038/s41598-018-38017-0

Wang, L., Lankhorst, L., and Bernards, R. (2022). Exploiting Senescence for the Treatment of Cancer. *Nat. Rev. Cancer* 22, 340–355. doi:10.1038/s41568-022-00450-9

Wei, C., Yang, C., Wang, S., Shi, D., Zhang, C., Lin, X., et al. (2019). Crosstalk between Cancer Cells and Tumor Associated Macrophages Is Required for Mesenchymal Circulating Tumor Cell-Mediated Colorectal Cancer Metastasis. *Mol. Cancer* 18, 64. doi:10.1186/s12943-019-0976-4

Xue, J., Wu, X.-L., Huang, X.-T., Guo, F., Yu, H.-F., Zhang, P.-C., et al. (2015). Correlation of Caveolin-1 Expression with Microlymphatic Vessel Density in Colorectal Adenocarcinoma Tissues and its Correlation with Prognosis. *Asian Pac. J. Trop. Med.* 8, 655–657. doi:10.1016/j.apjtm.2015.07.004

Yaghoubi, N., Soltani, A., Ghazvini, K., Hassanian, S. M., and Hashemy, S. I. (2019). PD-1/PD-L1 Blockade as a Novel Treatment for Colorectal Cancer. *Biomed. Pharmacother.* 110, 312–318. doi:10.1016/j.biopha.2018.11.105

Yang, J., Zhu, T., Zhao, R., Gao, D., Cui, Y., Wang, K., et al. (2018). Caveolin-1 Inhibits Proliferation, Migration, and Invasion of Human Colorectal Cancer Cells by Suppressing Phosphorylation of Epidermal Growth Factor Receptor. *Med. Sci. Monit.* 24, 332–341. doi:10.12659/msm.907782

Yang, K., Jiang, B., Lu, Y., Shu, Q., Zhai, P., Zhi, Q., et al. (2019). FOXM1 Promotes the Growth and Metastasis of Colorectal Cancer via Activation of β-Catenin Signaling Pathway. *Cmar* Vol. 11, 3779–3790. doi:10.2147/cmar.S185438

Yang, Y., Jiang, H., Li, W., Chen, L., Zhu, W., Xian, Y., et al. (2020). FOXM1/DVL2/Snail axis Drives Metastasis and Chemoresistance of Colorectal Cancer. *Aging* 12, 24424–24440. doi:10.18632/aging.202300

Zhao, Z., Han, F. H., Yang, S. B., Hua, L. X., Wu, J. H., and Zhan, W. H. (2015). Loss of Stromal Caveolin-1 Expression in Colorectal Cancer Predicts Poor Survival. *Wjg* 21, 1140–1147. doi:10.3748/wjg.v21.i4.1140

Zhong, R., Chen, X., Chen, X., Zhu, B., Lou, J., Li, J., et al. (2015). MAD1L1 Arg558His and MAD2L1 Leu84Met Interaction with Smoking Increase the Risk of Colorectal Cancer. *Sci. Rep.* 5, 12202. doi:10.1038/srep12202

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 | Frontiers in Genetics

SYSTEMATIC REVIEW
published: 12 July 2022
doi: 10.3389/fgene.2022.931785

Association Between Telomere Length and Skin Cancer and Aging: A Mendelian Randomization Analysis

Nannan Son, Yankun Cui and Wang Xi**

Jiangxi University of Chinese Medicine, Nanchang, China

Background: Telomere shortening is a hallmark of cellular senescence. However, telomere length (TL)-related cellular senescence has varying effects in different cancers, resulting in a paradoxical relationship between senescence and cancer. Therefore, we used observational epidemiological studies to investigate the association between TL and skin cancer and aging, and to explore whether such a paradoxical relationship exists in skin tissue.

Methods: This study employed two-sample Mendelian randomization (MR) to analyze the causal relationship between TL and skin cancer [melanoma and non-melanoma skin cancers (NMSCs)] and aging. We studied single nucleotide polymorphisms (SNPs) obtained from pooled data belonging to genome-wide association studies (GWAS) in the literature and biobanks. Quality control was performed using pleiotropy, heterogeneity, and sensitivity analyses.

Results: We used five algorithms to analyze the causal relationship between TL and skin aging, melanoma, and NMSCs, and obtained consistent results. TL shortening reduced NMSC and melanoma susceptibility risk with specific odds ratios (ORs) of 1.0344 [95% confidence interval (CI): 1.0168–1.0524, p = 0.01] and 1.0127 (95% CI: 1.0046–1.0209, p = 6.36E-07), respectively. Conversely, TL shortening was validated to increase the odds of skin aging (OR = 0.96, 95% CI: 0.9332–0.9956, p = 0.03). Moreover, the MR-Egger, maximum likelihood, and inverse variance weighted (IVW) methods found significant heterogeneity among instrumental variable (IV) estimates (identified as MR-Egger skin aging Q = 76.72, p = 1.36E-04; melanoma Q = 97.10, p = 1.62E-07; NMSCsQ = 82.02, p = 1.90E-05). The leave-one-out analysis also showed that the SNP sensitivity was robust to each result.

Conclusion: This study found that TL shortening may promote skin aging development and reduce the risk of cutaneous melanoma and NMSCs. The results provide a reference for future research on the causal relationship between skin aging and cancer in clinical practice.

Keywords: telomere length, skin cancer, skin aging, mendelian randomization, age

Edited by:
Rajkumar S. Kalra, Okinawa Institute of Science and Technology Graduate University, Japan

Reviewed by:
Mangala Hegde, Indian Institute of Technology Guwahati, India
Jiang Li, Geisinger Medical Center, United States

****Correspondence:***
Yankun Cui
20201020@jxutcm.edu.cn
Wang Xi
20201019@jxutcm.edu.cn

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

Received: *29 April 2022*
Accepted: *20 June 2022*
Published: *12 July 2022*

Citation:
Son N, Cui Y and Xi W (2022) Association Between Telomere Length and Skin Cancer and Aging: A Mendelian Randomization Analysis. Front. Genet. 13:931785. doi: 10.3389/fgene.2022.931785

1 INTRODUCTION

Telomeres are DNA–protein complexes, located at the chromosome ends of eukaryotic cells, that protect the chromosomes from degradation and fusion (Shay, 2018). Defective telomere function has been shown to lead to genetic instabilities in cancer, with the telomeres shortening as cells age (Blasco, 2005). Telomere length (TL) in cells has been extensively studied as an aging biomarker and risk factor for age-related diseases. However, the extent to which TL can reflect cancer relevance remains unclear (Arsenis et al., 2017). Telomere shortening accelerates skin aging while acting as a mitotic clock, preventing abnormal proliferation in cancer (Buckingham and Klingelhutz, 2011). Skin is a highly self-renewing tissue that must undergo extensive proliferation throughout an organism's life cycle. It is generally believed that aging caused by telomere shortening can increase cancer risk. However, many studies have found that cancers such as melanoma may occur due to excessive telomere lengthening (Ismail et al., 2021). Clarifying this contradiction requires further clinical and epidemiological research.

The skin is the largest organ in the human body, accounting for approximately 15% of an adult's body weight. Skin aging is a major problem and involves multiple complex factors, such as cellular DNA damage and changes in mitochondrial function (Lowry, 2020). As one of the most common cancers, the prevalence of skin cancer has been increasing over the past 3 decades. Skin cancers are mainly divided into melanoma and non-melanoma skin cancers (NMSCs); the latter includes basal cell carcinoma and squamous cell carcinoma cancer (Linares et al., 2015). According to the World Health Organization, as many as 60,000 people worldwide die of skin cancer each year, with melanoma being responsible for most cancer-related deaths. Therefore, exploring the correlation between multiple risk factors, skin cancer, and aging is necessary.

We conducted Mendelian randomization (MR) analysis to explore the causal relationship between TL and skin aging and the risk of skin cancer (NMSCs and melanoma). MR is also known as "Mendelian deconfounding" because it aims to give an estimate of causality without bias due to confounders. The instrumental variables (IVs) in MR studies must satisfy three core assumptions: 1) genetic IVs are related to exposure factors, 2) genetic IV formation can be regarded as a random assignment and has nothing to do with confounding factors, and 3) IVs can only be associated with exposure factors that affect the outcomes (Bowden and Holmes, 2019) (**Figure 1**). To avoid violating the assumptions of MR, first, we need to perform horizontal pleiotropy analysis to prevent IVs from directly affecting the results without exposure factor (Morrison et al., 2020). Second, removing SNPs in linkage disequilibrium (LD) (Cheng et al., 2020), avoiding IVs associated with causal variants may contribute to confounders. Third, intrinsic differences between populations (confounding factors) can be mitigated by restricting study populations to the same ethnic background.

2 MATERIALS AND METHODS

2.1 Study Design Overview

An overview of the study design is drawn in **Figure 2**. we used a two-sample MR study to explore possible causal relationships between our study's exposure and outcome. We used genome-wide association study (GWAS) datasets to estimate the effect of the exposure (TL) on the outcomes (skin cancer and skin aging).

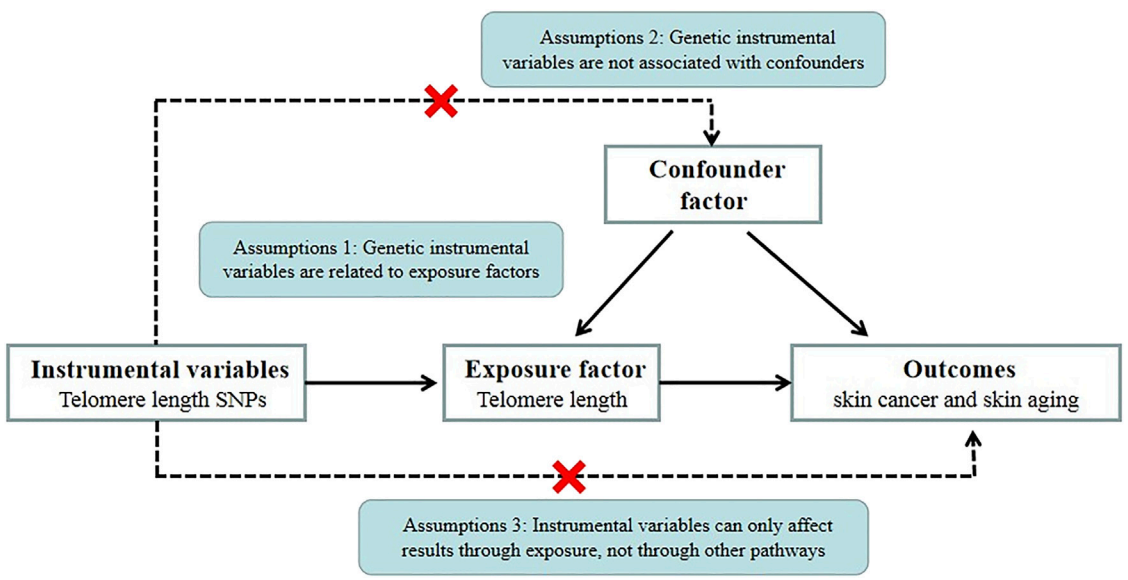


FIGURE 1 | MR basic requirements framework. Two-sample MR studies need to satisfy three assumptions. Assumption 1: Instrumental variables should be associated with telomere length. Assumption 2: The selected instrumental variables should be independent of confounders. Assumption 3: Instrumental variables should affect outcomes only through exposure and not through direct correlation.

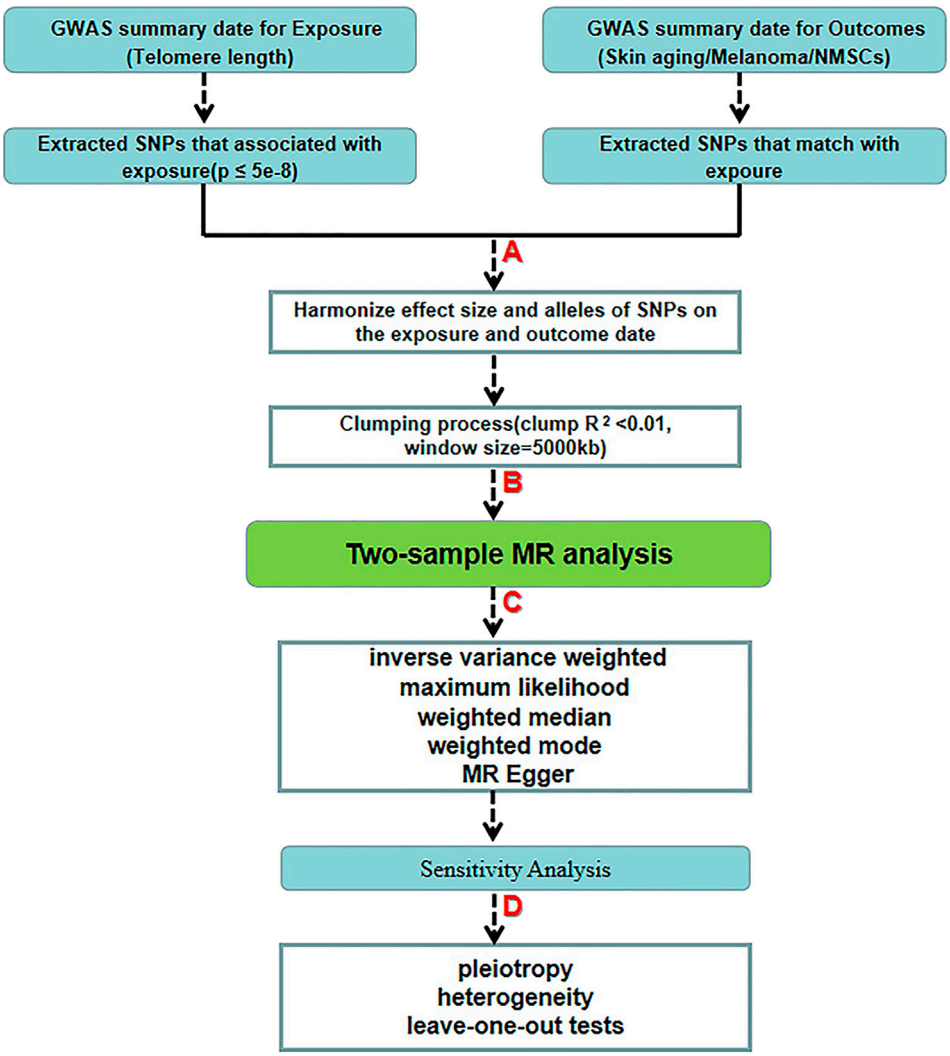


FIGURE 2 | Schematic diagram of the two-sample-MR analysis process. **(A)** Screening from GWAS data for selected SNPs significantly associated with exposure (p < 5.00 E-08 at genome-wide threshold). **(B)** Obtaining independent SNPs and examining effect size outliers with linkage disequilibrium (LD) with r < 0.01 or physical distance greater than 5000 kb. **(C)** Two-sample MR analysis by five algorithms. **(D)** Quality control by sensitivity analysis.

We selected single nucleotide polymorphisms (SNPs) closely associated with TL as IVs based on previously published GWAS databases and literature reports. The effects of the IVs on the exposure and outcomes were obtained from two independent samples. Ethical approval was provided in the original article for the GWAS-pooled dataset used in this study. Therefore, informed consent was no longer required.

2.2 Genetic Instrument Selection

Based on literature reports, SNP sites related to TL were screened. The SNPs associated with TL used in the present study came from a large GWAS of 7859 individuals from Europe. TL was measured in a mixed population of leukocytes using the established quantitative polymerase chain reaction technique, which expresses TL as the ratio of the number of telomere repeats to single-copy genes. Normalizing leukocyte TL measurements required the use of calibration samples or the quantification of a standard curve (Li et al., 2020). Statistics using cohort data identified 17 genome-wide significant loci, including several novel genes (*SENP7*, *MOB1B*, *CARMIL1*, *PRRC2A*, *TERF2*, and *RFWD3*), and confirmed the presence of other relevant genes (**Table 1**).

2.3 Skin Cancer and Skin Aging Genome-Wide Association Studies Selection

We searched the United Kingdom Biobank for aggregated GWAS data on common skin aging and cancers (NMSCs and melanoma) (Sudlow et al., 2015). Data on skin aging were obtained from a facial skin aging survey in 423,999 European participants; the skin

TABLE 1 | Association of 42 TL SNPs.

SNP ID	Closest gene	EA	NEA	EAF	Beta	SE	*p* value
rs10936600	LRRC34	T	A	0.2430	−0.0858	0.0057	6.42E-51
rs7705526	TERT	A	C	0.3283	0.0820	0.0058	4.82E-45
rs4691895	TERT	C	G	0.7829	0.0577	0.0061	1.47E-21
rs9419958	STN1	C	T	0.8616	−0.0636	0.0071	4.77E-19
rs75691080	STMN3	T	C	0.0912	−0.0671	0.0089	5.75E-14
rs59294613	POT1	A	C	0.2928	−0.0407	0.0055	1.12E-13
rs8105767	ZNF208	G	A	0.2887	0.0392	0.0054	5.21E-13
rs3219104	PARP1	C	A	0.8302	0.0417	0.0064	9.31E-11
rs2736176	PRRC2A	C	G	0.3134	0.0345	0.0055	3.41E-10
rs3785074	TERF2	G	A	0.2628	0.0351	0.0056	4.50E-10
rs7194734	MPHOSPH6	T	C	0.7816	−0.0369	0.0060	6.72E-10
rs34978822	RTEL1	G	C	0.0148	−0.1397	0.0227	7.04E-10
rs34991172	CARMIL1	G	T	0.0684	−0.0608	0.0105	6.03E-09
rs228595	ATM	A	G	0.4169	−0.0285	0.0050	1.39E-08
rs2302588	DCAF4	C	G	0.1003	0.0476	0.0084	1.64E-08
rs13137667	MOB1B	C	T	0.9591	0.0765	0.0137	2.37E-08
rs55749605	SENP7	A	C	0.5790	−0.0373	0.0067	2.38E-08
rs62053580	RFWD3	G	A	0.1694	−0.0389	0.0071	3.96E-08
rs12909131	ATP8B4	T	C	0.2309	−0.0308	0.0058	1.15E-07
rs1744757	MROH8	T	C	0.8507	0.0359	0.0068	1.38E-07
rs2124616	TYMS	A	G	0.1400	−0.0374	0.0072	1.72E-07
rs2613954	RP11	T	C	0.8858	−0.0381	0.0078	1.10E-06
rs12065882	MAGI3	G	A	0.2084	0.0298	0.0062	1.36E-06
rs2386642	ASB13	A	G	0.6732	−0.0256	0.0053	1.44E-06
rs56810761	UNC80	T	C	0.2698	0.0275	0.0057	1.45E-06
rs62365174	TENT2	G	A	0.0882	−0.0544	0.0113	1.50E-06
rs112655343	ATF7IP	T	C	0.1017	0.0425	0.0090	2.22E-06
rs60160057	DCLK2	A	G	0.2115	−0.0287	0.0062	3.15E-06
rs117536281	CDCA4	G	A	0.0342	0.0850	0.0183	3.31E-06
rs59192843	BBOF1	G	T	0.0592	0.0655	0.0141	3.52E-06
rs57415150	CSMD1	A	G	0.0417	−0.0584	0.0126	3.68E-06
rs6038821	LINC01706	T	A	0.0383	0.0596	0.0129	3.98E-06
rs144204502	TK1	T	C	0.0142	−0.0896	0.0196	4.92E-06
rs6107615	PROKR2	C	T	0.4217	−0.0228	0.0050	5.30E-06
rs117037102	CEP295	T	C	0.0179	0.0979	0.0218	6.81E-06
rs7276273	KRTAP10-4	C	A	0.0074	−0.1502	0.0334	6.90E-06
rs11665818	IFNL2	A	G	0.1946	0.0278	0.0062	7.04E-06
rs3213718	CALM1	T	C	0.5828	0.0224	0.0050	7.22E-06
rs143276018	NMRK2	C	T	0.0182	−0.1015	0.0229	9.02E-06
rs7311314	SMUG1	A	G	0.3174	0.0240	0.0054	9.50E-06
rs35675808	CD247	G	C	0.0281	0.0736	0.0166	9.54E-06
rs117610974	UNC13C	G	C	0.0094	−0.1540	0.0350	1.05E-05

EA, indicates effect allele; NEA, non-effect allele; SE, standard error; SNP, single nucleotide polymorphism; EAF, effect allele frequency.

cancer GWAS included 3,751 melanoma cases and 23,694 NMSCs, while 372,016 European participants were collected as controls. Analyses were adjusted for age, sex, and principal components when necessary. In addition, all SNPs in the MR analysis were derived from a GWAS of European ancestry to minimize potential bias due to population heterogeneity.

2.4 Single Nucleotide Polymorphisms Inclusion and Exclusion Criteria

To verify the validity of the IVs included in the MR analysis, we set the following screening criteria for eligible SNPs in the previously identified GWAS set. We selected SNPs that were significantly associated with our exposure ($p \leq 5E\text{-}8$) and that had a certain probability of mutation (minor allele frequency, MAF ≥ 5%) with no reported locus coincidence. To estimate LD between SNPs, 1000 Genome Project samples were used ($R^2 < 0.01$) (Siva, 2008). When there was an LD effect between SNPs, we selected the genetic variant with the lowest *p* value. We excluded all palindromic SNPs that could introduce ambiguity to the identity of the effector allele in the exposure GWAS. To limit bias from weak IVs, the F statistic should have been greater than 10. The formula for calculating F was as follows: $R^2 \times (n-k-1)/[(1-R^2) \times k]$, where n is the sample size of the GWAS, k is the number of SNPs, and R^2 is the proportion of telomere variability explained by each SNP. Specifically, the R formula calculates: $2 \times beta^2 \times (1\text{-}EAF) \times EAF$, where EAF is the effect allele frequency, and beta is an estimate of the genetic effect of each SNP on TL. To satisfy the third core hypothesis, SNPs associated with skin aging, and skin cancer were excluded, as were pathways that did not include TL.

2.5 Method Selection

We estimated the risk relationship between TL and skin aging and cancer using the MR Egger, inverse variance weighted (IVW), weighted median, maximum likelihood, and weighted mode MR methods. Considering the potential pleiotropic genetic variation effects, to avoid bias, we focused on the MR-Egger regression results, the slope of which can estimate the directed pleiotropy magnitude. IVW is used to take a weighted average of random variable measurements. Each random variable is weighted using the inverse of its variance. This method minimizes the mean variance and is often used in meta-analyses to integrate independent measurement results. Maximum likelihood uses known sample results to infer the parameter values that are most likely (maximum probability) to lead to such an outcome, which outperforms naive regression methods and reduces bias in misspecification. In addition (Luque-Fernandez et al., 2018), we also performed weighted median and weighted mode analysis using IVs to accurately estimate causal effects for more than 50% of the weights. Results were presented as odds ratios (ORs) and 95% confidence intervals (CIs).

2.6 Sensitivity Analysis

A sensitivity analysis was performed using the heterogeneity, pleiotropy, and leave-one-out tests. First, the MR-Egger method was used to analyze pleiotropy and to verify whether a single locus affected multiple phenotypes. The pleiotropy refers to the phenomenon of a single locus affecting multiple phenotypes. Horizontal pleiotropy occurs when genetic variants are associated with multiple phenotypes along multiple pathways, which can invalidate results derived from MR analysis (Bowden et al., 2017). MR-Egger regression analysis can be used to evaluate the bias generated by horizontal pleiotropy, and its regression intercept can evaluate the size of pleiotropy. The closer the intercept is to 0, the smaller the possibility of gene pleiotropy. If $p > 0.05$, it is considered that the possibility of gene pleiotropy in the causal analysis is weak, and its effect can be ignored.Second, the SNPs were individually removed through a leave-one-out sensitivity test, and the drawn forest map was viewed after analysis. If a certain SNP is eliminated, the result changes greatly, indicating that this SNP is an outlier and needs to be eliminated. If the overall solid line do not change much after removing a certain SNP (all solid line are on the same side of 0), the results are reliable. Third, there may be heterogeneity in IVs from different platforms or populations affecting resultsthe. Combined MR-Egger, maximum likelihood, and IVW methods were used for heterogeneity analysis, and the Cochran Q statistic was used to standardize the heterogeneity analysis. We used a two-sided p-value, with statistical significance set at $p < 0.05$. Statistical analyses were performed using the "TwoSampleMR" package in R 3.4.2.

3 RESULT

By screening and selecting SNPs based on the above criteria, we identified 42 SNPs that met the TL criteria by means of multiple tests such as LD (**Table 1**). The analysis found that the causal relationship between TL and skin aging was consistent across the MR-Egger, weighted median, maximum likelihood, and weighted mode calculation methods.These scatterplots representing SNPs reflect the effects of TL on skin aging and skin cancer, as shown in **Figure 3A**. The results indicated that the risk of skin aging increased with TL shortening. The MR-Egger test showed that TL was significantly associated with skin aging (OR = 0.96, 95% CI: 0.9332–0.9956, $p = 0.03$). In addition, the weighted median, maximum likelihood, and weighted mode methods showed that TL shortening increased the risk of skin aging (**Table 2**). However, IVW did not reveal any causal link between TL and skin aging (OR = 0.9880, 95% CI: 0.9735–1.0027, $p = 0.11$). The IVW method produces consistent causal estimates by combining the Wald ratios of the causal effects of each SNP, but this may also introduce an invalid IV (Bowden et al., 2016). IVW methods are susceptible to being hampered by extreme propensity scores, leading to biased estimates and excessive variance (Li et al., 2019). Based on the above five analyses, we concluded that the causal relationship between TL and skin aging was significant. According to the forest map drawn by SNP (**Figure 4**), the analysis for skin cancer was adjusted to exclude four palindromic SNPs (rs55749605, rs59294613, rs2386642, and rs59192843). The risk relationship between TL and melanoma and NMSCs showed consistent results across the five methods. As shown by the MR-Egger method, TL was significantly associated with NMSCs (OR = 1.0344, 95% CI: 1.0168–1.0524, $p = 4.60E-04$) and melanoma (OR = 1.0127, 95% CI: 1.0046–1.0209, $p = 6.36E-07$). Three methods, including IVW, also demonstrated that the risk of both melanoma and NMSCs decreased with TL shortening. As shown in **Figure 3**, the five methods could intuitively determine the direction of agreement. Therefore, we concluded that TL was significantly associated with both melanoma and NMSCs.

The quality controls for this study included pleiotropy, heterogeneity, and sensitivity tests. MR-Egger regression was used to test the pleiotropic effects of TL on skin aging, NMSCs, and melanoma. The results of each group showed that the effect of TL on skin aging and skin cancer had no significant horizontal pleiotropic bias (skin aging, $p = 0.1$; melanoma, $p = 0.72$; NMSCs, $p = 0.72$) (**Table 3**). The funnel plot according to IVW and MR-Egger also suggests that there was no horizontal pleiotropy (**Figure 5**). We utilized three algorithms (MR-Egger, maximum likelihood, and IVW) to analyze whether there was statistical heterogeneity among the IV estimates. We found that there was substantial heterogeneity among these IVs across the three outcomes (i.e., for skin aging, MR-Egger Q = 76.72, $p = 1.36E-04$; for melanoma, MR-Egger Q = 97.10, $p = 1.62E-07$; for NMSCs, MR-Egger Q = 538.50, $p = 1.90E-05$). We analyzed the sensitivity by performing the leave-one-out sensitivity test and found that regardless of which SNP was removed, it would not fundamentally impact the results (all lines were on the same side of 0) (**Figure 6**), This indicated that the MR

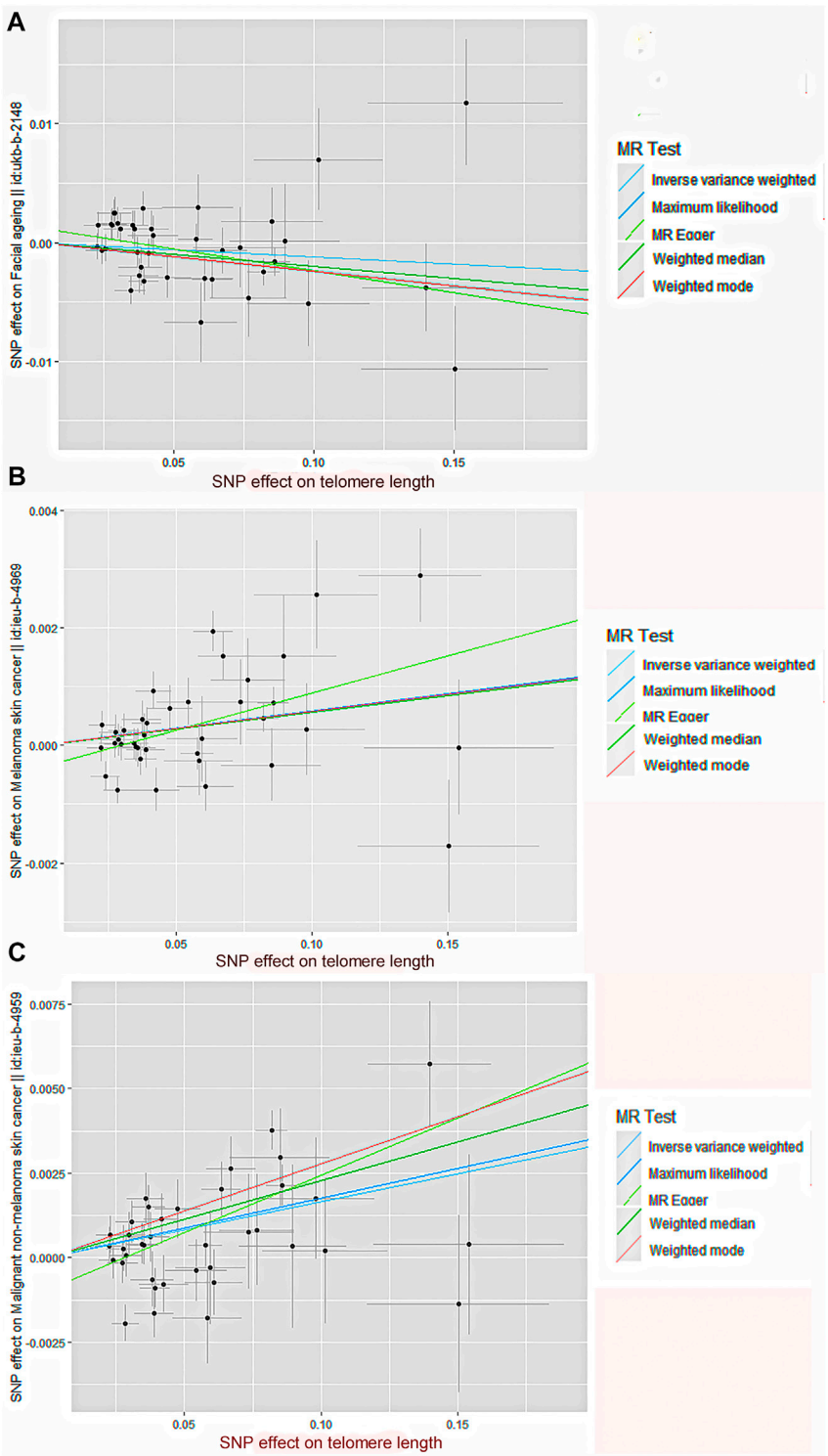


FIGURE 3 | Scatter plot of genetic causality between TL and skin aging and cancer using different MR methods.**(A)** Skin aging; **(B)** Melanoma; **(C)** NMSCs. The slope of the line represents the causality of the different methods. The dark blue line represents Maximum likelihood, the light green line represents MR Egger, the dark green line represents Weighted median, the light blue line represents IVW, and the red line represents Weighted mode.

TABLE 2 | MR estimates for each method to assess the effect of TL on skin aging and skin cancer.

Outcome	MR Methods	Number of SNPs	OR (95%CI)	SE	*P*
Neuroblastoma	MR-Egger	39	0.9639 (0.9332~0.9956)	0.0164	0.03
	Weighted median	39	0.9799 (0.9632~0.9969)	0.0087	0.01
Simple mode	Inverse variance weighted	39	0.9880 (0.9735~1.003)	0.0075	0.11
	Maximum likelihood	39	0.9878 (0.9777~0.9982)	0.0053	0.02
	Weighted mode	39	0.9758 (0.9582~0.9900)	0.0200	0.01
	MR-Egger	38	1.0344 (1.0168~1.0524)	0.0088	4.60E-04
	Weighted median	38	1.0231 (1.0138~1.0324)	0.0045	1.19E-06
	Inverse variance weighted	38	1.0166 (1.0084~1.0249)	0.0041	6.33E-05
	Maximum likelihood	38	1.0178 (1.0126~1.0231)	0.0027	1.09E-10
	Weighted mode	38	1.0282 (1.0120~1.0447)	0.0079	1.79E-03
melanoma	MR-Egger	38	1.0127 (1.0046~1.0209)	0.0041	6.36E-07
	Weighted median	38	1.0057 (1.0019~1.0095)	0.0019	4.99E-03
	Inverse variance weighted	38	1.0057 (1.0020~1.0095)	0.0019	2.68E-03
	Maximum likelihood	38	1.0059 (1.0036~1.0082)	0.0011	6.36E-07
	Weighted mode	38	1.0058 (1.0017~1.0103)	0.0023	1.96E-02

results were robust. The solid line is completely to the left of 0, indicating that the estimated result from this SNP is that TL shortening can increase skin aging. The solid line is completely to the right of 0, indicating that the estimated result from this SNP is that TL shortening can reduce the risk of skin cancer. All SNPs are on the side of 0, representing the stabilityof the results.In **Figure 6A**, the solid line is completely to the left of 0, indicating that the estimated result for this SNP is that TL shortening increases skin aging. In **Figure 6B,C**, the solid line is completely to the right of 0, indicating that the estimated result for this SNP is that TL shortening reduces the risk of skin cance.

4 DISCUSSION

Telomere shortening is observed in most human cancers, but it is worth noting that this phenomenon is controversial.New research suggests that longer-than-expected telomeres (made up of repetitive DNA sequences) are associated with an increased risk of several cancers, including melanoma (De Vitis et al., 2018). A novel point of this study is the paradoxical issue of factor analysis from genetics.Exploring epigenetic drivers of skin cancer and skin aging by means of TL-related SNPs.There are 42 SNPs that met the inclusion criteria for the core assumption of MR. Utilizing these IVs not only enables inferring a causal relationship between outcomes and exposure, but also effectively avoids confounding bias in traditional epidemiological studies (Tin and Köttgen, 2021). For example, the two SNPs (rs7705526, rs4691895) that we screened for maintaining TL are located near the TERT gene (**Table 1**). TERT is a cancer-related gene, and mutations in the noncoding region of the TERT gene are considered to be the cause of most melanomas (Toussi et al., 2020). The second innovation of this study, the MR study used an epidemiological approach to explore the association between skin aging and skin cancer. This study uses TL genetic variants as probes, reduces confounding factors and reverse causality and may be more convincing than classical observational experiments (Davey Smith and Hemani, 2014).

Telomeres are nuclear protein complexes at chromosome ends that maintain chromosomal stability, upon which normal cell division, differentiation, and regeneration depend (Turner et al., 2019). Normal telomeres and telomerase can regulate skin cell physiological function and abnormal proliferation; therefore, TLs play an essential role in skin aging and cancer development (Ventura et al., 2019). We used a two-sample MR to assess the association between TL and skin aging and cancer. We found an increased likelihood of skin aging with TL shortening. In contrast, the risk of NMSCs and melanoma was significantly reduced with TL shortening. Contrary to the general belief that cell senescence can trigger cancer, TL shortening as a risk factor for skin aging did not increase the risk of skin cancer. This is consistent with recent clinical reports that long telomeres are associated with increased mortality in more than 2000 melanoma patients from hospital clinics and the general population (Ismail et al., 2021).

Telomeres are closely linked to cellular aging, especially in dermal cells. Telomere shortening in skin fibroblasts may lead to epidermal aging and barrier function defects (Quan et al., 2015). Long-term exposure to ultraviolet (UVA) radiation has long been recognized as the most important factor in skin aging. Studies have found that in fibroblasts exposed to UVA at doses of 1000 or 10,000 mJ/cm^2, TL was significantly shorter than in unirradiated controls, negatively correlating with the UVA dose (Ma et al., 2012). In addition, reactive oxygen species (ROS) are considered another cause of skin aging (Kammeyer and Luiten, 2015). Studies have shown that ROS cause cellular senescence owing to accelerated telomere shortening (Anderson et al., 2014). In the present study, all four MR methods showed a significant relationship between TL shortening and skin aging. Each standard deviation of TL shortening was genetically predicted to increase the risk of skin aging by 4.61%. The sensitivity analysis verified the reliability

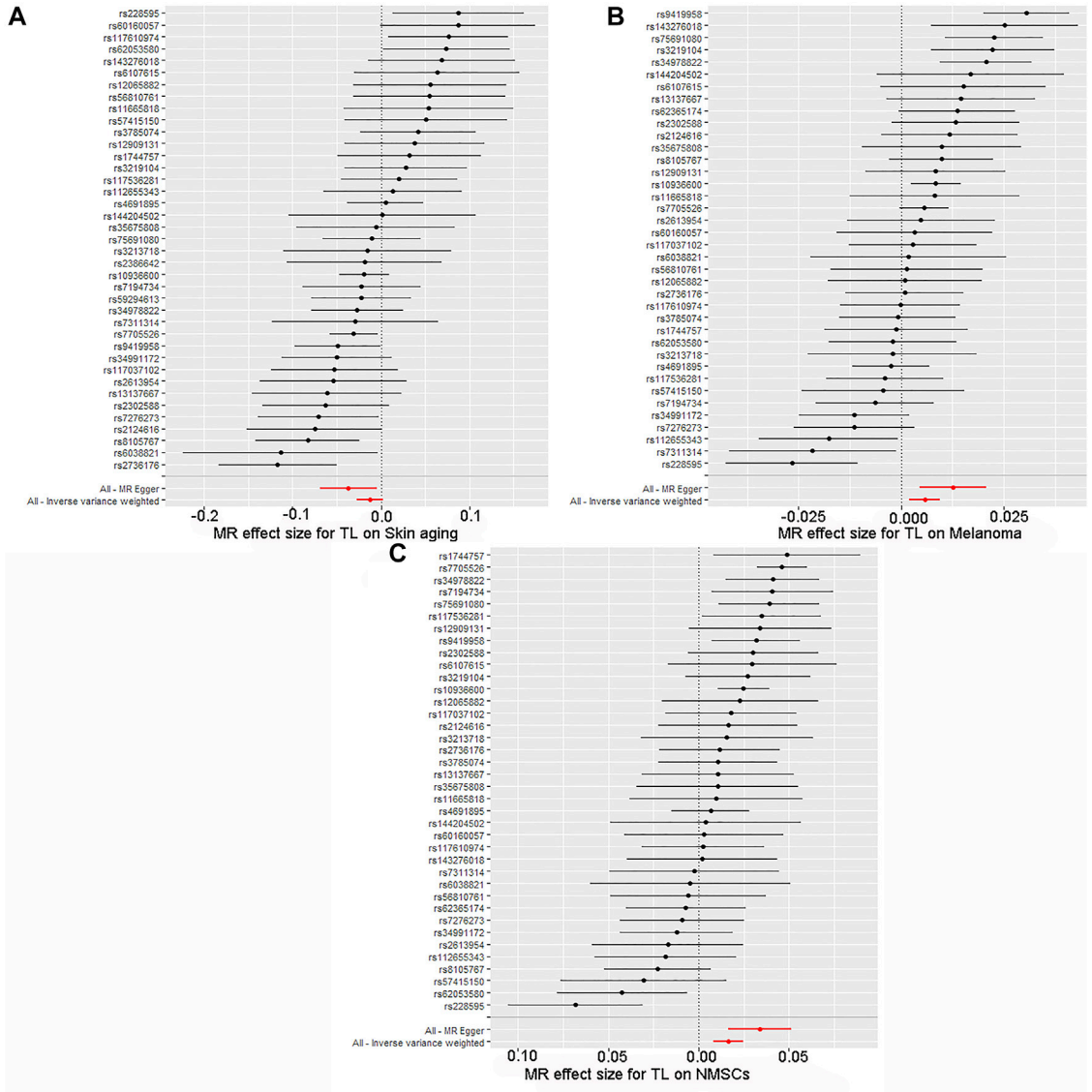


FIGURE 4 | Forest map of skin cancer and aging risk based on TL genetic variants. **(A)** Skin aging; **(B)** Melanoma; **(C)** NMSCs.SNPs of TL were analyzed using IVW and MR-Egger. Black dots represent estimates of causal effects of TL on skin cancer and aging (beta coefficients). The black line represents the estimated 95% confidence interval.

TABLE 3 | Heterogeneity and pleiotropy analysis by MR Egger, IVW, Maximum likelihood.

Outcome	MR Methods	Cochran Q statistic	Heterogeneity *p*-value	Pleiotropy *p*-value
skin aging	MR-Egger	76.72	1.36E-04	0.10
	Inverse variance weighted	82.55	3.86E-05	
	Maximum likelihood	82.30	4.15E-05	
NMSCs	MR-Egger	82.02	1.90E-05	0.10
	Inverse variance weighted	93.12	9.78E-07	
	Maximum likelihood	90.62	2.15E-06	
melanoma	MR-Egger	97.10	1.62E-07	0.72
	Inverse variance weighted	106.60	1.15E-08	
	Maximum likelihood	105.44	1.71E-08	

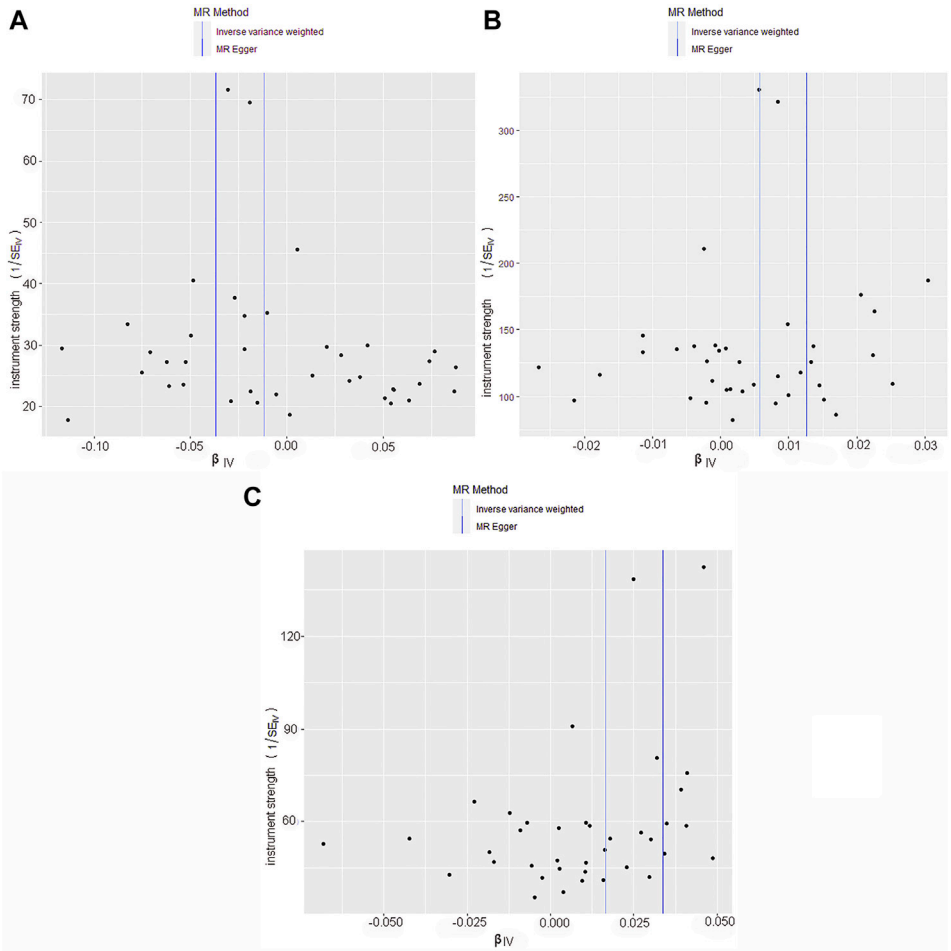


FIGURE 5 | Funnel plot of skin cancer and aging risk based on TL genetic variants. **(A)** Skin aging; **(B)** Melanoma; **(C)** NMSCs.Causal effects were expressed as log odds ratios for skin aging and skin cancer per unit shortening of telomere length. Overall causal estimates (beta coefficients) of telomere length and skin aging and skin cancer estimated by the IVW (light blue line) and MR-Egger (dark blue line) methods are shown.

of all SNP results using the leave-one-out method. Therefore, maintaining TL can protect chromosome stability and prevent skin aging caused by DNA breakage damage.

TL is a key factor in cell proliferative potential, and much evidence supports the vital role of altered TL in cancer pathogenesis (Xu et al., 2013). However, studies analyzing the association between TL and cancer risk have yielded conflicting conclusions (Aviv et al., 2017). Studies have found that most solid cancers originating from proliferating tissues display short telomere characteristics, and most cancer incidence increases with age. In contrast, in the general population, individuals with constitutively long telomeres also have an increased risk of some serious cancers (Savage et al., 2013). From **Table 1**, the genes of SNPs that maintain TL are all related to cancer. For example, the protection of telomere 1 (POT1) protein is an important subunit of the Shelterin telomere-binding complex, which can promote the development of various cancers by leading to immortalization (Wu et al., 2020). Therefore, combined with the multiple related genes we screened (including STN1, STMN3, PRRC2A), we speculate that the telomere shortening inhibitory pathway in tumors has been determined at birth.when telomeres are too long, the telomere reserve will not be depleted in time, which will provide additional divisions for cancer cells, especially for familial-prone tumors (melanoma glioma, non-Hodgkin lymphoma, etc.) (Nieters et al., 2012; Feng et al., 2018; Ali et al., 2021).

4.1 Limitations

Because of the small study population, the effect of TL on skin aging and cancer was not found to be very significant. To bring our results closer to reality, we used five algorithms with different characteristics, such as the MR-Egger method, as a reference. This method assessed whether genetic variation had pleiotropic effects on the results that differed from zero on average (directional pleiotropy) and provided consistent causal effect estimates under a weaker assumption–the

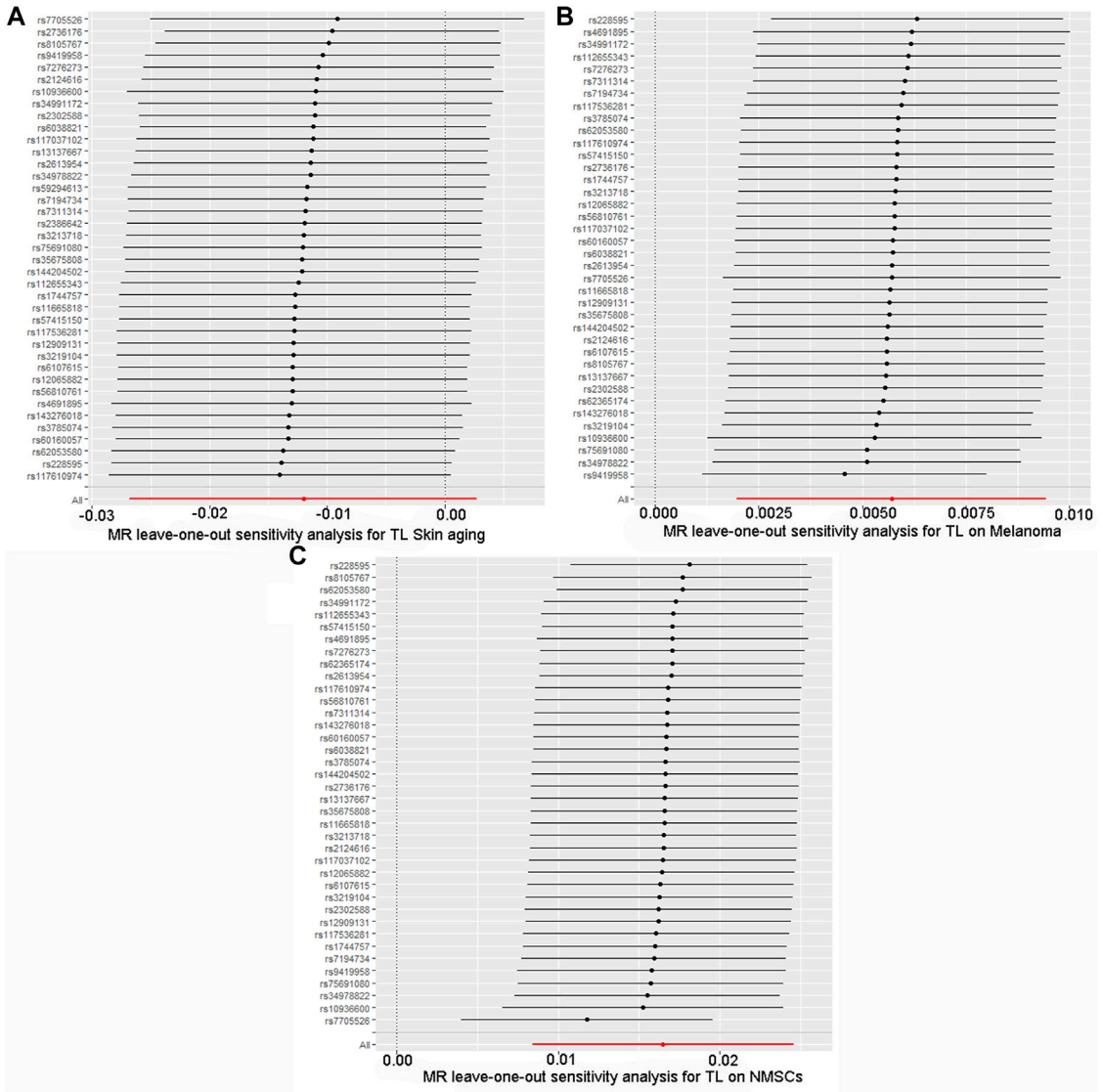


FIGURE 6 | Forest plot for leave-one-out sensitivity analysis. **(A)** Skin aging; **(B)** Melanoma; **(C)** NMSCs.Each horizontal solid line reflects the result estimated by a single SNP using the Wald ratio method.The solid line is completely to the left of 0, indicating that the estimated result from this SNP is that TL shortening can increase skin aging. The solid line is completely to the right of 0, indicating that the estimated result from this SNP is that TL shortening can reduce the risk of skin cancer. All SNPs are on the side of 0, representing the stabilityof the results.

InSIDE (instrumental strength independent of direct effects) assumption (Burgess and Thompson, 2017). Furthermore, we selected the largest and most reliable GWAS available to explore the causal relationship between TL and skin aging and cancer. In our MR research framework, the interference of confounders and reverse causality were minimized. In addition, the original United Kingdom Biobank study lacked a breakdown of the population (including gender and age) and was unable to conduct further subgroup analyses. Clinically, the complex physiological mechanisms of TL and skin aging and cancer go well beyond these simple models. Further studies are needed to identify the underlying mechanisms that provide insights into skin aging and cancer and to facilitate prevention.

5 CONCLUSION

This study supported the causal relationship that TL shortening may promote the development of skin aging and reduce the risk of cutaneous melanoma and NMSCs. The results provide a reference for future research on the relationship between skin aging and cancer in clinical practice. Also, our study provides evidence for skin aging and cancer (melanoma and NMSCs) treatment and diagnosis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

Both WX and YC are the corresponding authors of this article and were responsible for data calculations. NS completed the final writing of the article.

FUNDING

This work was financially supported by The National Natural Science Foundation of China (Youth Science Program) No. 82104725.

ACKNOWLEDGMENTS

We thank the United Kingdom Biobank for providing GWAS aggregated data on skin cancer and skin aging.

REFERENCES

Ali, M. W., Patro, C. P. K., Zhu, J. J., Dampier, C. H., Plummer, S. J., Kuscu, C., et al. (2021). A Functional Variant on 20q13.33 Related to Glioma Risk Alters Enhancer Activity and Modulates Expression of Multiple Genes. *Hum. Mutat.* 42 (1), 77–88. doi:10.1002/humu.24134

Anderson, A., Bowman, A., Boulton, S. J., Manning, P., and Birch-Machin, M. A. (2014). A Role for Human Mitochondrial Complex II in the Production of Reactive Oxygen Species in Human Skin. *Redox Biol.* 2, 1016–1022. doi:10.1016/j.redox.2014.08.005

Arsenis, N. C., You, T., Ogawa, E. F., Tinsley, G. M., and Zuo, L. (2017). Physical Activity and Telomere Length: Impact of Aging and Potential Mechanisms of Action. *Oncotarget* 8 (27), 45008–45019. doi:10.18632/oncotarget.16726

Aviv, A., Anderson, J. J., and Shay, J. W. (2017). Mutations, Cancer and the Telomere Length Paradox. *Trends Cancer* 3 (4), 253–258. doi:10.1016/j.trecan.2017.02.005

Blasco, M. A. (2005). Telomeres and Human Disease: Ageing, Cancer and beyond. *Nat. Rev. Genet.* 6 (8), 611–622. doi:10.1038/nrg1656

Bowden, J., and Holmes, M. V. (2019). Meta-analysis and Mendelian Randomization: A Review. *Res. Syn. Meth* 10 (4), 486–496. doi:10.1002/jrsm.1346

Bowden, J., Davey Smith, G., Haycock, P. C., and Burgess, S. (2016). Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* 40 (4), 304–314. doi:10.1002/gepi.21965

Bowden, J., Del Greco M, F., Minelli, C., Davey Smith, G., Sheehan, N., and Thompson, J. (2017). A Framework for the Investigation of Pleiotropy in Two-Sample Summary Data Mendelian Randomization. *Stat. Med.* 36 (11), 1783–1802. doi:10.1002/sim.7221

Buckingham, E. M., and Klingelhutz, A. J. (2011). The Role of Telomeres in the Ageing of Human Skin. *Exp. Dermatol.* 20 (4), 297–302. doi:10.1111/j.1600-0625.2010.01242.x

Burgess, S., and Thompson, S. G. (2017). Interpreting Findings from Mendelian Randomization Using the MR-Egger Method. *Eur. J. Epidemiol.* 32 (5), 377–389. doi:10.1007/s10654-017-0255-x

Cheng, Q., Yang, Y., Shi, X., Yeung, K.-F., Yang, C., Peng, H., et al. (2020). MR-LDP: a Two-Sample Mendelian Randomization for GWAS Summary Statistics Accounting for Linkage Disequilibrium and Horizontal Pleiotropy. *Nar. Genom Bioinform.* 2 (2), lqaa028. doi:10.1093/nargab/lqaa028

Davey Smith, G., and Hemani, G. (2014). Mendelian Randomization: Genetic Anchors for Causal Inference in Epidemiological Studies. *Hum. Mol. Genet.* 23 (R1), R89–R98. doi:10.1093/hmg/ddu328

De Vitis, M., Berardinelli, F., and Sgura, A. (2018). Telomere Length Maintenance in Cancer: At the Crossroad between Telomerase and Alternative Lengthening of Telomeres (ALT). *Int. J. Mol. Sci.* 19 (2), 606. doi:10.3390/ijms19020606

Feng, X., Hsu, S.-J., Bhattacharjee, A., Wang, Y., Diao, J., and Price, C. M. (2018). CTC1-STN1 Terminates Telomerase while STN1-TEN1 Enables C-Strand Synthesis during Telomere Replication in Colon Cancer Cells. *Nat. Commun.* 9 (1), 2827. doi:10.1038/s41467-018-05154-z

Ismail, H., Helby, J., Hölmich, L. R., Chakera, A. H., Bastholt, L., Klyver, H., et al. (2021). Genetic Predisposition to Long Telomeres Is Associated with Increased Mortality after Melanoma: A Study of 2101 Melanoma Patients from Hospital Clinics and the General Population. *Pigment. Cell Melanoma Res.* 34 (5), 946–954. doi:10.1111/pcmr.12971

Kammeyer, A., and Luiten, R. M. (2015). Oxidation Events and Skin Aging. *Ageing Res. Rev.* 21, 16–29. doi:10.1016/j.arr.2015.01.001

Li, F., Thomas, L. E., and Li, F. (2019). Addressing Extreme Propensity Scores via the Overlap Weights. *Am. J. Epidemiol.* 188 (1), 250–257. doi:10.1093/aje/kwy201

Li, C., Stoma, S., Lotta, L. A., Warner, S., Albrecht, E., Allione, A., et al. (2020). Genome-wide Association Analysis in Humans Links Nucleotide Metabolism to Leukocyte Telomere Length. *Am. J. Hum. Genet.* 106 (3), 389–404. doi:10.1016/j.ajhg.2020.02.006

Linares, M. A., Zakaria, A., and Nizran, P. (2015). Skin Cancer. *Prim. Care Clin. Office Pract.* 42 (4), 645–659. doi:10.1016/j.pop.2015.07.006

Lowry, W. (2020). Its Written All over Your Face: The Molecular and Physiological Consequences of Aging Skin. *Mech. Ageing Dev.* 190, 111315. doi:10.1016/j.mad.2020.111315

Luque-Fernandez, M. A., Schomaker, M., Rachet, B., and Schnitzer, M. E. (2018). Targeted Maximum Likelihood Estimation for a Binary Treatment: A Tutorial. *Stat. Med.* 37 (16), 2530–2546. doi:10.1002/sim.7628

Ma, H.-M., Liu, W., Zhang, P., and Yuan, X.-Y. (2012). Human Skin Fibroblast Telomeres Are Shortened after Ultraviolet Irradiation. *J. Int. Med. Res.* 40 (5), 1871–1877. doi:10.1177/030006051204000526

Morrison, J., Knoblauch, N., Marcus, J. H., Stephens, M., and He, X. (2020). Mendelian Randomization Accounting for Correlated and Uncorrelated Pleiotropic Effects Using Genome-wide Summary Statistics. *Nat. Genet.* 52 (7), 740–747. doi:10.1038/s41588-020-0631-4

Nieters, A., Conde, L., Slager, S. L., Brooks-Wilson, A., Morton, L., Skibola, D. R., et al. (2012). PRRC2A and BCL2L11 Gene Variants Influence Risk of Non-Hodgkin Lymphoma: Results from the InterLymph Consortium. *Blood* 120 (23), 4645–4648. doi:10.1182/blood-2012-05-427989

Quan, C., Cho, M. K., Perry, D., and Quan, T. (2015). Age-associated Reduction of Cell Spreading Induces Mitochondrial DNA Common Deletion by Oxidative Stress in Human Skin Dermal Fibroblasts: Implication for Human Skin Connective Tissue Aging. *J. Biomed. Sci.* 22 (1), 62. doi:10.1186/s12929-015-0167-6

Savage, S. A., Gadalla, S. M., and Chanock, S. J. (2013). The Long and Short of Telomeres and Cancer Association Studies. *JNCI J. Natl. Cancer Inst.* 105 (7), 448–449. doi:10.1093/jnci/djt041

Shay, J. W. (2018). Telomeres and Aging. *Curr. Opin. Cell Biol.* 52, 1–7. doi:10.1016/j.ceb.2017.12.001

Siva, N. (2008). 1000 Genomes Project. *Nat. Biotechnol.* 26 (3), 256. doi:10.1038/nbt0308-256b

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., et al. (2015). UK Biobank: an Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med.* 12 (3), e1001779. doi:10.1371/journal.pmed.1001779

Tin, A., and Köttgen, A. (2021). Mendelian Randomization Analysis as a Tool to Gain Insights into Causes of Diseases: A Primer. *J. Am. Soc. Nephrol.* 32 (10), 2400–2407. doi:10.1681/ASN.2020121760

Toussi, A., Mans, N., Welborn, J., and Kiuru, M. (2020). Germline Mutations Predisposing to Melanoma. *J. Cutan. Pathol.* 47 (7), 606–616. doi:10.1111/cup.13689

Turner, K., Vasu, V., and Griffin, D. (2019). Telomere Biology and Human Phenotype. *Cells* 8 (1), 73. doi:10.3390/cells8010073

Ventura, A., Pellegrini, C., Cardelli, L., Rocco, T., Ciciarelli, V., Peris, K., et al. (2019). Telomeres and Telomerase in Cutaneous Squamous Cell Carcinoma. *Int. J. Mol. Sci.* 20 (6), 1333. doi:10.3390/ijms20061333

Wu, Y., Poulos, R. C., and Reddel, R. R. (2020). Role of POT1 in Human Cancer. *Cancers* 12 (10), 2739. doi:10.3390/cancers12102739

Xu, L., Li, S., and Stohr, B. A. (2013). The Role of Telomere Biology in Cancer. *Annu. Rev. Pathol. Mech. Dis.* 8, 49–78. doi:10.1146/annurev-pathol-020712-164030

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 | Frontiers in Genetics

TYPE Original Research
PUBLISHED 04 August 2022
DOI 10.3389/fgene.2022.961554

OPEN ACCESS

EDITED BY
Dhanendra Tomar,
Wake Forest School of Medicine,
United States

REVIEWED BY
Shivani Ror,
The University of Iowa, United States
Milton Roy,
Johns Hopkins Medicine, United States

*CORRESPONDENCE
Wenqi Yang,
doctorwenqiyang@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Genetics of Aging, a section of the journal Frontiers in Genetics

RECEIVED 04 June 2022
ACCEPTED 11 July 2022
PUBLISHED 04 August 2022

CITATION
Dai L, Wang X, Bai T, Liu J, Chen B, Li T and Yang W (2022), Identification of a novel cellular senescence-related signature for the prediction of prognosis and immunotherapy response in colon cancer.
Front. Genet. 13:961554.
doi: 10.3389/fgene.2022.961554

Identification of a novel cellular senescence-related signature for the prediction of prognosis and immunotherapy response in colon cancer

Longfei Dai†, Xu Wang†, Tao Bai, Jianjun Liu, Bo Chen, Ting Li and Wenqi Yang*

Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei, China

The study was conducted to construct a cellular senescence-related risk score signature to predict prognosis and immunotherapy response in colon cancer. Colon cancer data were acquired from the Gene Expression Omnibus and The Cancer Genome Atlas databases. And cellular senescence-related genes were obtained from the CellAge database. The colon cancer data were classified into different clusters based on cellular senescence-related gene expression. Next, prognostic differential genes among clusters were identified with survival analysis. A cellular senescence-related risk score signature was developed by performing the LASSO regression analysis. Finally, PCA analysis, t-SNE analysis, Kaplan-Meier survival analysis, ROC analysis, univariate Cox regression analysis, multivariate Cox regression analysis, C-index analysis, meta-analysis, immune infiltration analysis, and IPS score analysis were used to evaluate the significance of the risk signature for predicting prognosis and immunotherapy response in colon cancer. The colon cancer data were classified into three clusters. The patients in cluster A and cluster B had longer survival. A cellular senescence-related risk score signature was developed. Patients in the low-risk score group showed a better prognosis. The risk score signature could predict colon cancer patients' prognosis independently of other clinical characteristics. The risk score signature predicted the prognosis of colon cancer patients more accurately than other signatures. Patients in the low-risk score group showed a better response to immunotherapy. The opposite was true for the high-risk score group. In conclusion, the cellular senescence-related risk score signature could be used for the prediction of prognosis and immunotherapy response in colon cancer.

KEYWORDS

colon cancer, cellular senescence, prognosis, immunotherapy, signature

Introduction

Nowadays, colon cancer has high morbidity and mortality all over the world, which seriously threatens human life (Sung et al., 2021). Surgical resection is preferred for early-stage colon cancer, while systemic chemotherapy is the main treatment for advanced colon cancer (Chakrabarti et al., 2020; Body et al., 2021). However, the effectiveness of chemotherapy for partial colon cancer patients is often unsatisfactory due to the emergence of drug resistance (Azwar et al., 2021). In recent years, immunotherapy has brought new hope for the treatment of cancer (Lichtenstern et al., 2020). Immunotherapy drugs (PD1/PD-L1 blocker and CTLA-4 blocker) have been shown to improve the prognosis of colon cancer patients (Yaghoubi et al., 2019; Ben et al., 2021). Unfortunately, the prognosis and immunotherapy responses of different colon cancer patients are significantly differentiated due to the existence of tumor heterogeneity (Marisa et al., 2021; Guo et al., 2022). Therefore, it is particularly important to distinguish between patients with colon cancer who show a better prognosis and immunotherapy response.

Cellular senescence, the permanent cessation of cell proliferation, is thought to be able to prevent the development and metastasis of tumor cells (Calcinotto et al., 2019; Di Micco et al., 2021). However, recent studies have shown that senescent cancer cells promote tumorigenesis in neighboring cells through the release of SASP (Prieto and Baker, 2019). Demirci et al. revealed that it was the Jekyll and Hyde nature of cancer cell senescence (Demirci et al., 2021). Furthermore, cellular senescence had been demonstrated to be a potential target for cancer in clinical therapy (Prasanna et al., 2021; Wang et al., 2022). Lin et al. found that cellular senescence was important in the prognosis and immunotherapy of lung cancer (Lin et al., 2021). And Zhou et al. demonstrated that cellular senescence was a potential marker of prognosis and therapeutic outcome in gastric cancer (Zhou et al., 2022). However, the role of cellular senescence in the prognosis and immunotherapy of colon cancer is not well understood.

In this study, we aimed to investigate the significance of cellular senescence in the prognosis and immunotherapy of colon cancer. Meanwhile, a cellular senescence-related risk score signature was constructed to distinguish patients with a better prognosis and immunotherapy response.

Methods

Acquisition of colon cancer information and cellular senescence-related genes

Transcriptome information, clinical information, and mutation information were acquired from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). Next, the gene ID from the transcriptome information was converted into gene names to obtain TCGA expression data. Then the FPKM value of TCGA expression data was converted into the TPM value. The survival time, survival status, age, gender, pathological TNM stage, pathological T-stage, pathological N-stage, and pathological M-stage were extracted from the clinical information. The platform file GPL570 and probe matrix GSE39582 were also downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The probe matrix was transformed into a gene matrix by finding the correspondence between the probe matrix and gene names based on the platform file information. The copy number data of colon cancer were obtained from UCSC Xena (http://xena.ucsc.edu/). Finally, the TCGA expression data and GEO expression data were merged to obtain the expression of genes in the merged data. Cellular senescence-related genes were acquired from the CellAge database (Supplementary Table S1). The expression of cellular senescence-related genes was extracted from the merged data. The workflow chart was visualized in Supplementary Figure S1.

Clusters based on cellular senescence-related gene expression

The "ConsensusClusterPlus" package was used to perform a Consensus Clustering analysis on the merged data. The principal component analysis (PCA) was performed to validate the accuracy of distinguishing different clusters based on cellular senescence-related gene expression. Kaplan-Meier survival analysis was performed with the "survival" and "survminer" packages. The heatmap was plotted by using the "pheatmap" package. The single-sample gene set enrichment analysis (ssGSEA) and gene set variation analysis (GSVA) were conducted based on the "GSEABase" and "GSVA" packages.

Differential analysis of clusters

Differential genes (DEGs) among clusters were identified (adjusted p-value = 0.001). The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed to explore enrichment pathways on DEGs.

Gene clusters based on differential genes

Prognostic genes were identified by performing survival analysis with the "survival" package (filtering condition: p-value < 0.05). The "ConsensusClusterPlus" package was applied to conduct the Consensus Clustering analysis on DEGs among clusters. PCA analysis was performed to validate the accuracy of distinguishing different gene clusters based on DEGs among clusters. Kaplan-Meier survival analysis was

applied with the "survival" and "survminer" packages. The heatmap was plotted with the "pheatmap" package.

Constructing a risk score signature

The colon cancer data were classified into training and testing sets. The selection operator (LASSO) Cox regression analysis with 10-fold cross-validation was performed to construct a prognostic risk score signature in the training set. The testing set was used to validate the accuracy of the signature. Formula:

$$\text{Risk score} = \sum i1\,(\text{Coefi} * \text{ExpGenei})$$

"Coef", regression coefficient; "ExpGene", the expression of genes. The training and testing sets were divided into high- and low-risk groups based on the medium value of risk scores. Principal component analysis (PCA) and t-distributed stochastic neighbor embedding analysis (t-SNE) were applied to confirm the signature's accuracy to distinguish between high- and low-risk score groups. The Kaplan-Meier survival curves and ROC curves were plotted with "survival", "survminer", and "timeROC" packages. The "ggplot2" and "pheatmap" packages were applied to plot risk status, survival status, risk histogram, and risk heatmap. Univariate and multivariate Cox regression analyses were conducted to validate whether the signature predicted colon cancer patients' prognosis independently of other clinical characteristics. C-index curves were plotted based on the "survival", "rms", and "pec" packages. The "timeROC" and "survcomp" packages were used to plot ROC curves and C-index histograms for comparison of signatures. The meta-analysis was performed to investigate the heterogeneity of the risk score signature in predicting the prognosis of colon cancer patients between training and testing sets with the "meta" package. The forest diagram of the meta-analysis was drawn by using the fixed-effects model.

Developing a nomogram

A nomogram was plotted with the "regplot", "rms", and "survivor" packages. The ROC curve of the nomogram was drawn based on the "timeROC" package.

Validating the risk score signature in clinical subgroups

The Sankey diagram was plotted to illustrate the construction process of the risk score signature with the "ggalluvial" package. The "ggpubr" and "ggplot2" packages were used to plot box plots to show the differences in different risk scores across clusters and gene clusters. The heatmap and box plot were drawn to investigate differences in patients' risk scores across clinical subgroups with the "ComplexHeatmap", "ggpubr", and "limma" packages. We also performed Kaplan-Meier survival analysis to further validate the application of the risk score signature in different clinical subgroups.

The landscape of gene mutation in different risk score groups

The "maftools" package was applied to visualize the gene mutation landscape in high- and low-risk score groups.

Exploring immunotherapy response in different risk score groups

Immune score files for colon cancer were downloaded from the Cancer Immunome Database (TCIA, https://tcia.at/). Immunotherapy analysis was performed to explore the therapy differences of IPS-CTLA4, PD1, PDL1, and PDL2 blockers in patients with different risk scores with the "ggpubr" package. The"pRRophetic_0.5. tar.gz" was acquired from the Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org/). Finally, the "pRRophetic" package was used to analyze the differences in half-maximal inhibitory concentration (IC50) values between different risk score groups and to identify potential drugs for colon cancer patients.

Validating the risk score signature

Differential expression of the signature between normal and tumor samples was investigated by performing differential analysis. Finally, we searched the Human Protein Atlas (HPA, https://www.proteinatlas.org/) database for immunohistochemical results of the signature genes. In addition, the gene mutation and copy number variant of the signature genes were analyzed.

Statistical analysis

All scripts were run in Strawberry-Perl-5.32.1.1 and all codes were run in R 4.1.2. The colon cancer data were classified into different clusters by Consensus Clustering analysis. Then, DEGs among clusters were identified. The prognostic DEGs were identified by survival analysis. The colon cancer data were again divided into different gene clusters by Consensus Clustering analysis. Next, LASSO regression analysis was performed to construct a risk score signature. Patients were classified into high- and low-risk score groups. PCA analysis

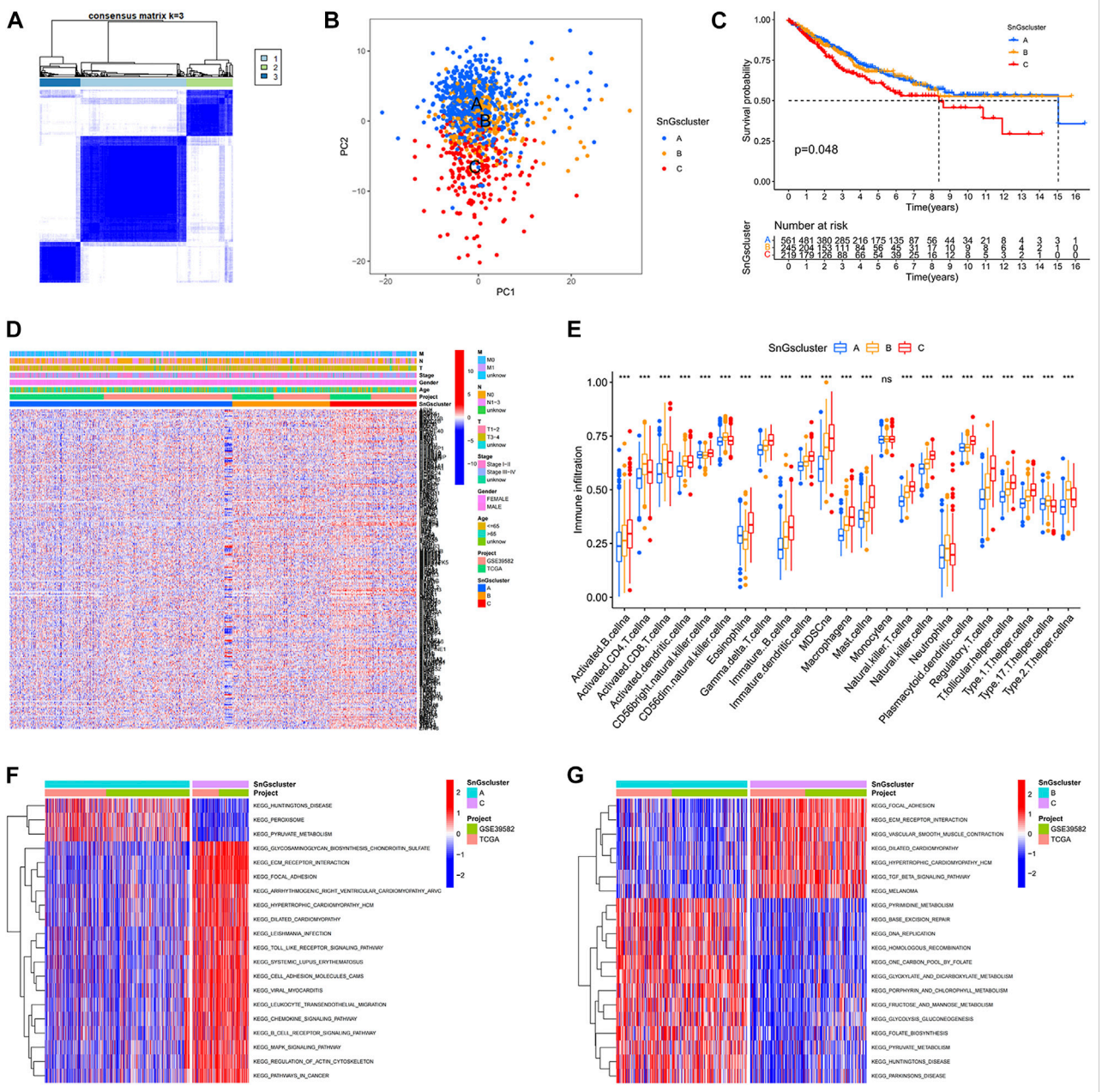


FIGURE 1
Three clusters based on cellular senescence-related gene expression levels. **(A)** Consensus Cluster Analysis. When cluster Num = 3, the relationship in the cluster was tight and the correlation between clusters was weak. **(B)** PCA plot. Blue dots represent patients in Cluster A; yellow dots represent patients in Cluster B; red dots represent patients in Cluster C. **(C)** Kaplan-Meier survival curves. The prognosis of patients was different among the three clusters, $p = 0.048$. **(D)** Heat map. Cellular senescence-related genes were upregulated in Cluster C. **(E)** Box plots. The horizontal coordinate represents immune infiltrating cells; the vertical coordinate represents immune scores; ns represents no difference in immune cell scores among the three clusters; $^{***}p < 0.001$. **(F)** Differential GSVA enrichment pathways between Cluster A and Cluster C. Red represents the high expression pathway and blue represents the low expression pathway. **(G)** Differential GSVA enrichment pathways between Cluster B and Cluster C. Red represents the high expression pathway and blue represents the low expression pathway.

and t-SNE analysis were applied to confirm the accuracy of signature in distinguishing high and low-risk score groups. Kaplan-Meier survival analysis, ROC analysis, univariate analysis, multivariate analysis, C-index method, and meta-analysis were performed to explore the role of the signature in the prognosis of colon cancer. And the TICA algorithm and "pRRophetic" package were used to investigate the significance of the signature in therapy for colon cancer. Finally, all signature

genes were performed for differential analysis and validated in the HPA database. *p*-values less than 0.05 were considered to be statistically significant.

Results

Clusters based on cellular senescence-related gene expression level

We acquired 41 normal samples and 473 colon cancer tissue samples from the TCGA database. Another 585 colon cancer samples were obtained from the GEO database. 279 cellular senescence-related genes were acquired from the CellAge database. When cluster Num = 3, the relationship in the cluster was tight and the correlation between clusters was weak (Figure 1A). So, all samples were classified into three clusters. Other classification results were visualized in Supplementary Figure S2. In the PCA plot, red points, yellow points, and blue points were separated, which indicated that cluster A, cluster B, and cluster C could be distinguished based on the expression of cellular senescence-related genes (Figure 1B). The prognosis of patients among three clusters showed differences, with patients in cluster A and cluster B having longer survival times than cluster C (Figure 1C). We observed that cellular senescence-related genes were expressed at the lowest level in cluster A and at the highest level in cluster C (Figure 1D). And the three clusters showed no difference in the different clinical subgroups. In addition, the differences in the level of immune infiltration among the three clusters were analyzed. Interestingly, cluster C not only contained high immune cell infiltration, but also many immunosuppressive cells, such as myeloid-derived suppressor cells (MDSCs), regulatory T cells (Tregs), and macrophages (Figure 1E). It might be associated with the worse prognosis of colon cancer patients in cluster C.

We also investigated the differences in enrichment pathways among the three clusters. The significantly enriched pathways in cluster A included "PEROXISOME", "PYRUVATE METABOLISM", and "HUNTINGTONS DISEASE" (Figure 1F). And the predominantly enriched pathways in cluster B were "BASE EXCISION REPAIR", "HOMOLOGOUS RECOMBINATION", "PYRUVATE METABOLISM", "PARKINSONS DISEASE", and "HUNTINGTONS DISEASE" (Figure 1G). While"GLYCOSAMINOGLYCAN BIOSYNTHESIS CHONDROITIN SULFATE", "ECM RECEPTOR INTERACTION", "FOCAL ADHESION", and "TGF BETA SIGNALING PATHWAY" were significantly enriched in cluster C. The KEGG pathways in cluster A and cluster B were mainly involved in tumor suppression processes, while the KEGG pathways in cluster C were associated with tumorigenesis and metastasis.

2334 differential genes among three clusters

To further investigate the differences among the three clusters, we identified 2334 DEGs in cluster A, cluster B, and cluster C (Supplementary Figure S3A and Supplementary Table S2). And the enrichment pathways for DEGs were visualized in Supplementary Figure S3B,C. The results of the GO enrichment analysis showed that "positive regulation of cell adhesion", "leukocyte migration" and "leukocyte cell–cell adhesion" were significantly enriched in biological processes (BP); "collagen–containing extracellular matrix", "cell–substrate junction", and "focal adhesion" were significantly enriched in molecular function (CC); "actin-binding" and "extracellular matrix structural constituent" were significantly enriched in the cellular component (MF). "PI3K–Akt signaling pathway", "Focal adhesion", "Osteoclast differentiation", "Rap1 signaling pathway", and "Proteoglycans in cancer" were significantly enriched in KEGG. The significantly enriched pathways in DEGs were associated with tumor development and metastasis.

Gene clusters based on prognostic DEGs

We further studied the association of DEGs among clusters with prognosis, and 681 prognostic DEGs were identified (Supplementary Table S3). Colon cancer samples were classified into three gene clusters based on the expression of prognostic DEGs. When cluster Num = 3, the relationship in the gene cluster was tight and the correlation between gene clusters was weak (Supplementary Figure S4). The prognostic DEGs were significantly down-regulated in gene Cluster A and up-regulated in gene Cluster C (Figure 2A). Gene cluster A had the best prognosis, while gene cluster C had the worst prognosis (Figure 2B). We further validated the accuracy of classifying colon cancer samples into three gene clusters based on the expression of prognostic DEGs (Figure 2C).

Construction and validation of a cellular senescence-related risk score signature

We constructed a risk score signature to predict the prognosis of colon cancer patients based on prognostic DEGs (Figures 3A,B). The training set was used to construct the risk score signature, while the testing set was applied to validate the accuracy of the signature. Formula: Risk score = FITM2 exp. * (-0.340377976707324) + APOL6 exp. * (-0.385962800820076) + VWF exp. * 0.348549751087245 + PRRX2 exp.* 0.222316119860682 + CCL22 exp. * (-0.505065627522616) + ALPL exp. * 0.427185163663466 + SON exp. * 0.403475482932549 + KIF7 exp. * 0.466822368846872 + ZEB1-AS1 exp. * 0.509278568824046. The colon cancer

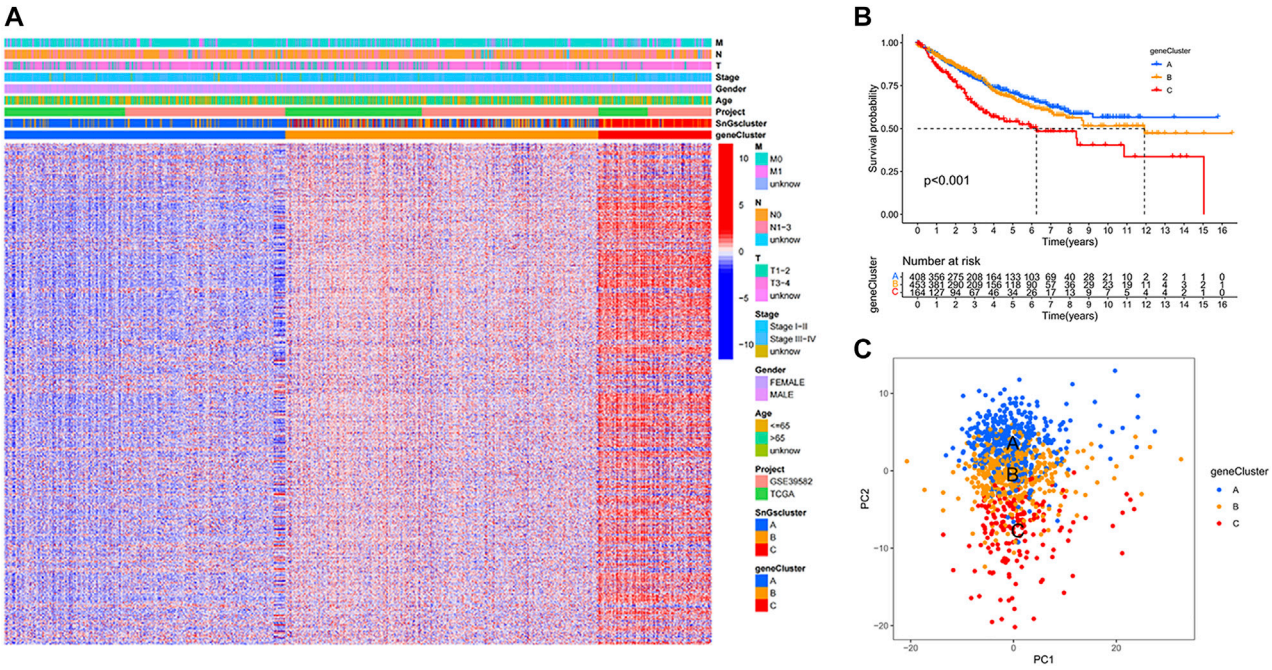


FIGURE 2
Three gene clusters based on prognostic differential gene expression levels. **(A)** Heat map. Prognostic differential genes were upregulated in gene Cluster C and down-regulated in gene Cluster A. **(B)** Kaplan-Meier survival curves. $p < 0.001$. **(C)** PCA plot. Blue dots represent patients in gene Cluster A; yellow dots represent patients in gene Cluster B; red dots represent patients in gene Cluster C.

sample was classified into high and low-risk score groups based on the medium value of the risk score. The statistics of clinical information for the training set and the testing set were visualized in Supplementary Table S4. Red points and blue points were significantly separated in the PCA plot and the t-SNE plot, which demonstrated the accuracy of distinguishing high- and low-risk score groups based on the risk score (Figures 3C–F). We observed a better prognosis for patients in the low-risk score group (Figures 3G–I). Moreover, the prediction of patient survival at 1, 3, and 5 years was more accurate based on the signature (Figure 3J-3L). The high-risk score group had higher mortality (Figures 4A–I). The expression levels of FITM2, APOL6, and CCL22 decreased significantly with increasing risk scores, which were low-risk genes. In contrast, the expression levels of VWF, PRRX2, ALPL, SON, KIF7, and ZEB1-AS1 increased significantly with increasing risk scores, which were high-risk genes.

We further confirmed the accuracy of the signature in predicting the prognosis of patients with colon cancer. *p*-values for the risk score were less than 0.001 in both univariate and multivariate Cox regression analyses, which indicated that the risk score could predict the prognosis of colon cancer patients independently of other clinical characteristics (Figures 5A–D). Moreover, the risk score predicted prognosis more accurately than other clinical characteristics, with the highest C-index value (Figure 5E). We searched online for four risk score signatures (Wang signature, Zhang signature, Zheng signature, and Ren signature) that predicted the prognosis of colon cancer (Ren et al., 2020; Zhang et al., 2020; Wang et al., 2021; Zheng et al., 2021). Surprisingly, the cellular senescence-related signature predicted the prognosis of colon cancer patients significantly better than the other four signatures, with the highest C-index value of 0.682 (Figure 5F). Supplementary Figure S5 visualized the predicted survival ROC curves and survival curves for other signatures. Furthermore, the meta-analysis showed less heterogeneity when using the signature to predict the prognosis of patients with colon cancer with I^2 <50% (Figure 5G).

Development of a nomogram

A nomogram was developed to benefit clinical work in predicting 1-year, 3-years, and 5-years survival probability in patients with colon cancer. For example, when the total point was 328, the 1-year survival probability of patients was more than 0.981, the 3-years survival probability was more than 0.937, and the 5-years survival probability was more than 0.898 (Figure 6A). Moreover, we found that predicting the survival probability of colon cancer patients was significantly better than other clinical characteristics based on the nomogram, with the highest AUC value of 0.823 (Figure 6B).

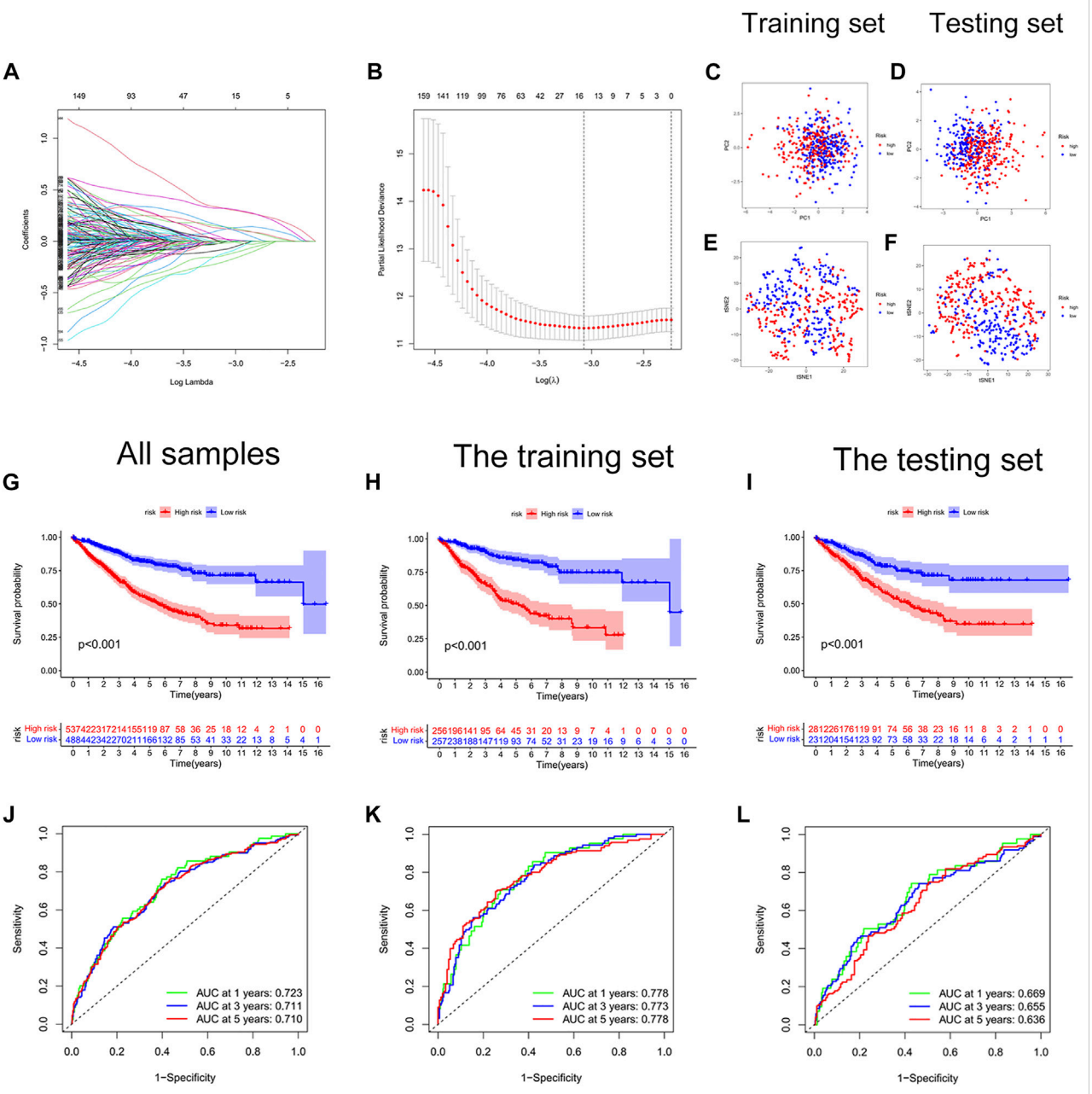


FIGURE 3
Construction of a cellular senescence-related risk score signature. **(A)** LASSO regression analysis. **(B)** Cross-validation for tuning the parameter selection. **(C)** The PCA plot of the training set. **(D)** The PCA plot of the test set. **(E)** The t-SEN plot of the training set. **(F)** The t-SEN plot of the testing set. **(G)** The K-M survival curve of all colon cancer samples, $p < 0.001$. **(H)** The K-M survival curve of the training set, $p < 0.001$. **(I)** The Kaplan-Meier survival curve of the testing set, $p < 0.001$. **(J)** The AUC values of 1-year, 3-years, and 5-years survival for all colon cancer patients were more than 0.700. **(K)** The AUC values of 1-year, 3-years, and 5-years survival for the training set were more than 0.750. **(L)** The AUC values of 1-year, 3-years, and 5-years survival for the testing set were more than 0.600.

Validation of risk score signature in clinical subgroups

The construction process of the signature was illustrated in Figure 7A. We further analyzed the differences in risk scores among the different clusters. Cluster A and gene Cluster A had the lowest risk scores, while cluster C and gene Cluster C had the highest risk scores (Figures 7B,C). All colon cancer patients were also classified into survival and death groups based on survival outcomes. Interestingly, patients in the survival group showed lower risk scores (Figure 7D). It further confirmed the above findings that patients in the low-risk score group had a better prognosis.

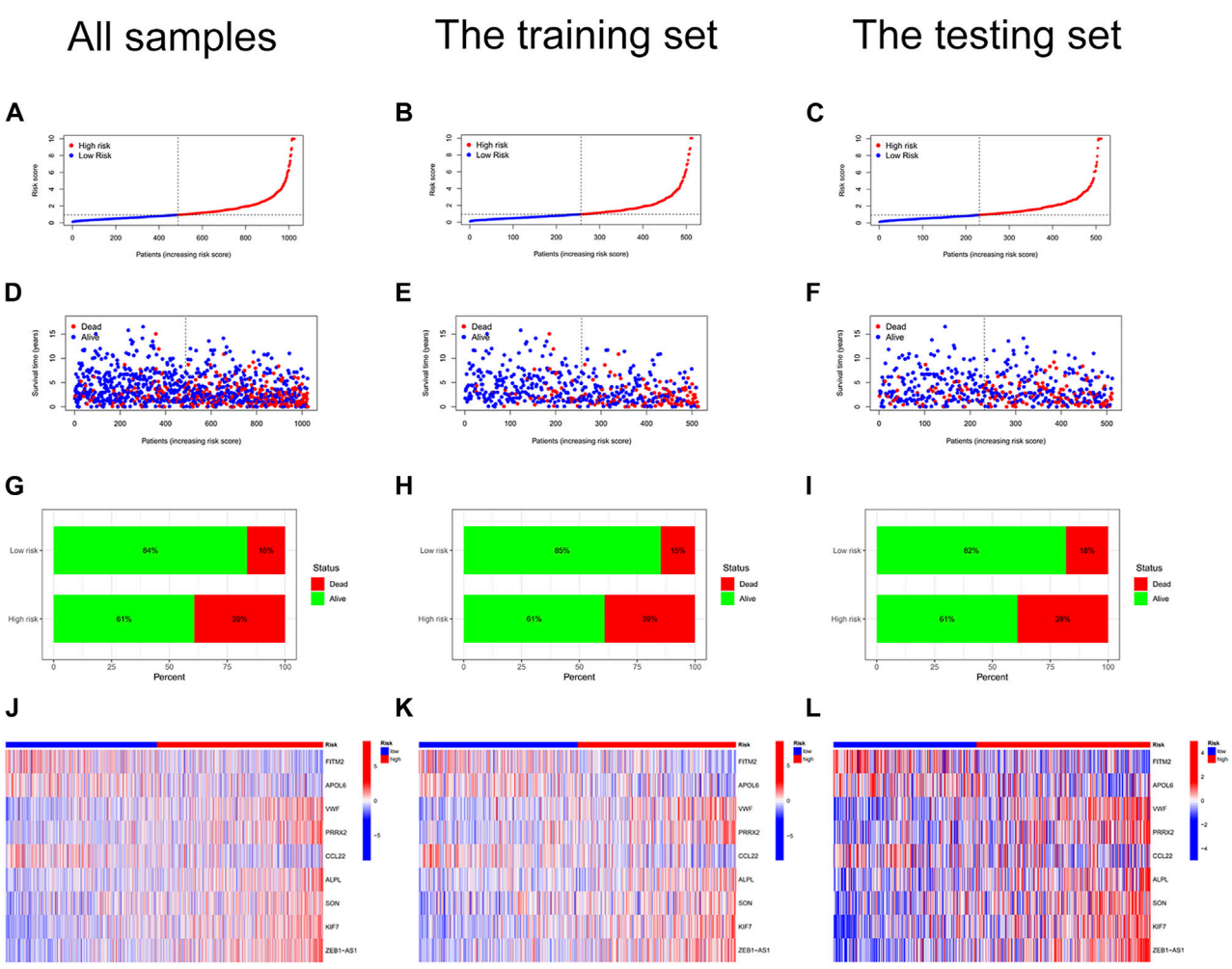


FIGURE 4
Risk curves. **(A)** Risk status plot of all colon cancer samples. The horizontal coordinate represented the ranked patients, and the risk scores of patients gradually increased from left to right; the vertical coordinate represented the risk scores. **(B)** Risk status of the training set. **(C)** Risk status of the testing set. **(D)** Survival status plot of all colon cancer samples. **(E)** Survival status plot of the training set. **(F)** Survival status plot of the testing set. **(G)** Risk histogram of all colon cancer samples. The percentage of survival patients in the low-risk score group was higher than that in the high-risk score group. **(H)** Risk histogram of the training set. **(I)** Risk histogram of the testing set. **(J)** Risk heat map of all colon cancer samples. The expression of FITM2, APOL6, and CCL22 decreased with increasing risk scores; the expression of VWF, PRRX2, ALPL, SON, KIF7, and ZEB1-AS1 increased with increasing risk scores. **(K)** Risk heat map of the training set. **(L)** Risk heat map of the testing set.

Next, we also explored whether there were differences in risk scores across clinical characteristics (Figure 7E). Unexpectedly, the risk scores showed no differences between men and women, nor between different age groups (≤65 and >65). In contrast, there were differences in risk scores among T stage (T1, T2, T3, T4), N stage (N0, N1, N2, N3), M stage (M0, M1), and pathological TNM stage (Stage I, Stage II, Stage III, Stage IV). The risk score increased gradually after T2 (Figure 7F). The risk score increased gradually after N0, except for N3 (Figure 7G). The risk score of M1 was significantly higher than that of M0 (Figure 7H). The risk score increased gradually after Stage I (Figure 7I). The immunotyping of colon cancer is classified into four subtypes, C1 (Wound Healing), C2 (IFN-gamma Dominant), C3 (Inflammatory), and C4 (Lymphocyte Depleted). Unexpectedly, there was no difference in risk scores among subtypes, except for the difference in risk scores between C1 and C2 (Figure 7J). And C2 had a lower risk score than C1.

We also observed that the signature was applicable to predict the prognosis of colon cancer patients in different clinical subgroups, including different ages, different gender, different T stages, different N stages, different M stages, and different pathological TNM stages (Figures 8A–L).

Gene mutation landscape in high- and low-risk score group

We also investigated gene mutations in different risk score groups. The gene mutation frequency of the low-risk score group

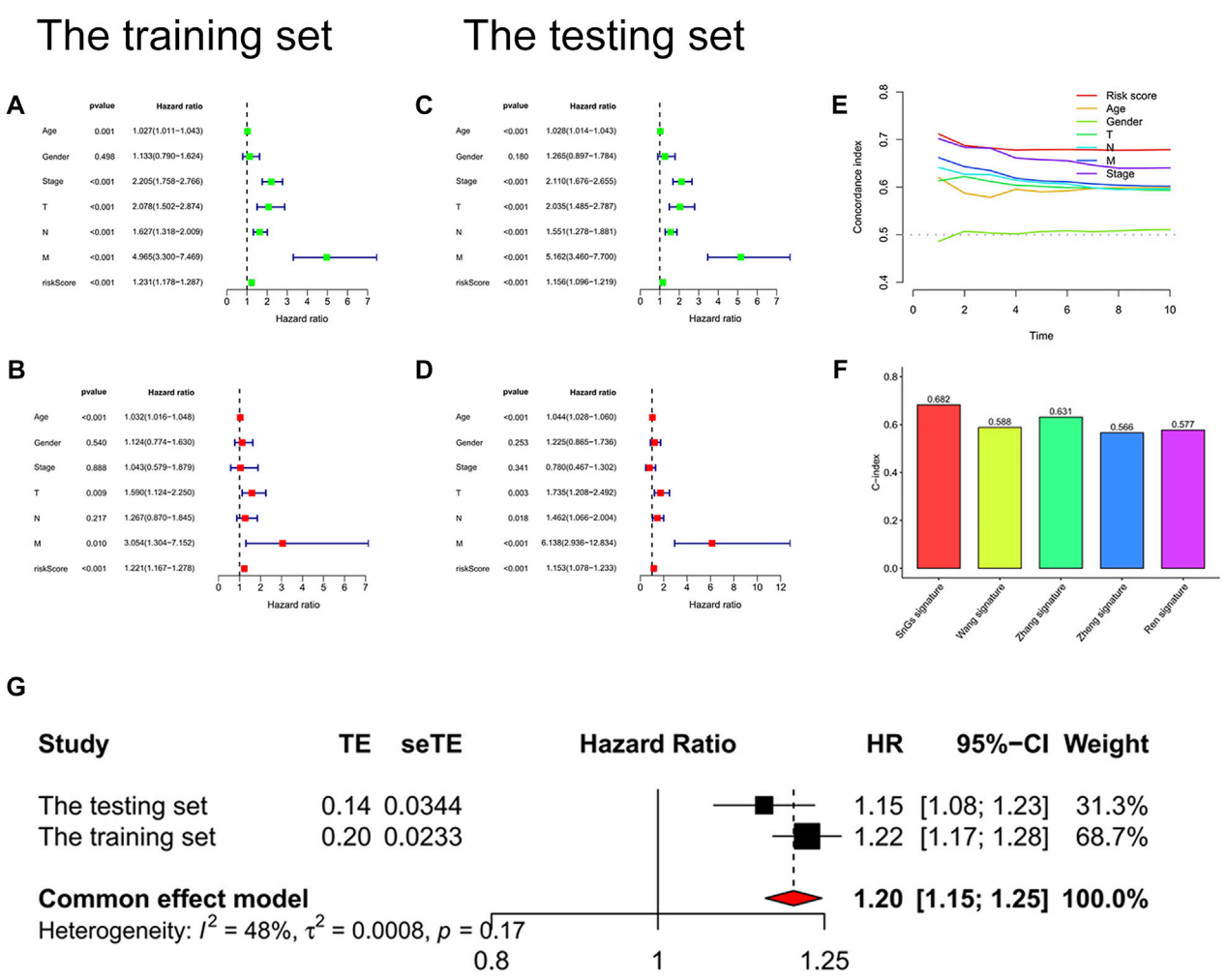


FIGURE 5
Validation of the risk score signature. **(A)** Univariate Cox regression analysis of the training set. Risk scores, $p < 0.001$. **(B)** Multivariate Cox regression analysis of the training set. Risk scores, $p < 0.001$. **(C)** Univariate Cox regression analysis of the testing set. Risk scores, $p < 0.001$. **(D)** Multivariate Cox regression analysis of the testing set. Risk scores, $p < 0.001$. **(E)** C-index curves. C-index value of risk scores was higher than other clinical characteristics (age, gender, pathological stage, T stage, N stage, and M stage). **(F)** Histogram for signature comparison with C-index method. SnGs signature had the highest C-index value compared to other signatures, AUC = 0.682. **(G)** Forest plot. The multivariate Cox regression analysis results of the training and testing sets were used to perform a meta-analysis, I2<50%.

was higher than the high-risk score group. The top 20 genes with mutation frequencies in the high-risk score group were visualized in Supplementary Figure S6A, while the low-risk score group was visualized in Supplementary Figure S6B.

Immunotherapy response in high- and low-risk score groups

In clinical work, patients with colon cancer have individual differences and develop different responses to different therapeutic drugs, resulting in different therapy outcomes. Immunotherapy and chemotherapy are currently the main tools in the systemic therapy of colon cancer. We observed that patients in the low-risk score group had higher immune scores (IPS) and better responses to immunotherapy drugs (CTLA4, PD1, PDL1, PDL2) (Figures 9A–D). We also identified 12 drugs suitable for colon cancer (Figures 9E–P). In particular, the IC50 values of four drugs (Erlotinib, Metformin, Methotrexate, and Mitomycin) were lower in the low-risk score group and more suitable for patients in the low-risk score group. In contrast, the IC50 values of eight drugs (Bexarotene, Bleomycin, Dasatinib, Docetaxel, Embelin, Imatinib, Pazopanib, and Shikonin) were lower in the high-risk score group and more applicable to patients in the high-risk score group.

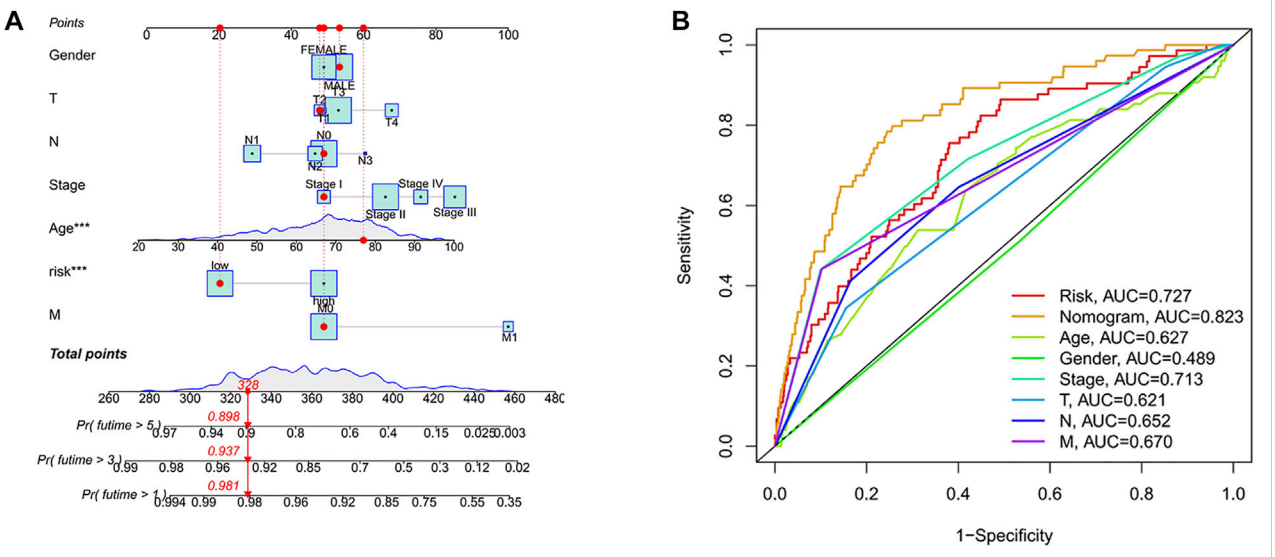


FIGURE 6
Developing a Nomogram. **(A)** Nomogram. When the total point was 328, the nomogram predicted that the 1-year survival probability of patients was more than 0.981, the 3-years survival probability was more than 0.937, and the 5-years survival probability was more than 0.898. **(B)** Nomogram ROC curve. The horizontal coordinate represented the false-positive rate expressed by 1-Specificity and the vertical coordinate represented the true-positive rate expressed by sensitivity.

Validation of the signature genes in normal and tumor tissues

ALPL, APOL6, SON, VWF, FITM2, and ZEB1-AS1 were significantly differentially expressed in normal and colon cancer tissues (Figures 10A–I). In particular, ALPL, APOL6, SON, and VWF were lowly expressed in tumor tissues. In contrast, FITM2 and ZEB1-AS1 were highly expressed in tumor tissues. We also confirmed the differential expression of ALPL, APOL6, SON, and VWF in the HPA database (Figure 10J). ALPL, APOL6, SON, and VWF proteins were significantly differentially expressed between normal and tumor tissues. In contrast, the expression levels of KIF7 and PRRX2 showed no difference between normal and tumor tissues.

Copy number variants and gene mutation in signature genes

We further investigated the gene mutation and copy number variants (CNV) for 10 signature genes in colon cancer. The gene with the highest mutation frequency was VWF at 8%, while ZEB1-AS1 had the lowest mutation frequency at 0% (Figure 11A). Interestingly, we observed that all genes showed amplification except PRRX2 and ALPL, which showed depletion (Figure 11B). And the chromosomal location of the CNV changes was visualized in Figure 11C. We also analyzed the expression of signature genes in different clusters. We observed that most of the signature genes were highly expressed in cluster C and gene Cluster C (Figures 11D,E).

Discussion

Cellular senescence has been revealed to inhibit the progression of colon cancer cells (Acosta et al., 2013; Cho et al., 2013). However, paradoxically, cellular senescence has also been found to promote the development of colon cancer (Guo et al., 2019). It might be related to the fact that it is highly heterogeneous (Sikora et al., 2021; Wang and Demaria, 2021). In view of the importance of cellular senescence in colon cancer, we would like to construct a cellular senescence-related risk score signature to predict prognosis and immunotherapy response.

In this article, firstly, all colon cancer samples were classified into three clusters based on cellular senescence-related gene expression. Clusters A and B had a better prognosis. In contrast, cluster C showed a worse prognosis since it contained high levels of immunosuppressive cell infiltration (MDSCs, Tregs, and macrophages) (Togashi and Nishikawa, 2017; Tian et al., 2019; Katopodi et al., 2021). The reasons for the differential prognosis of the three clusters were also revealed in the results of the GSVA analysis. The enrichment pathway in cluster A was associated with anti-tumor (Kim, 2020; Wenes et al., 2022). “PARKINSONS DISEASE” and “HUNTINGTONS DISEASE” were significantly enriched in cluster B. Patients with neurodegenerative diseases were considered to be at low risk of

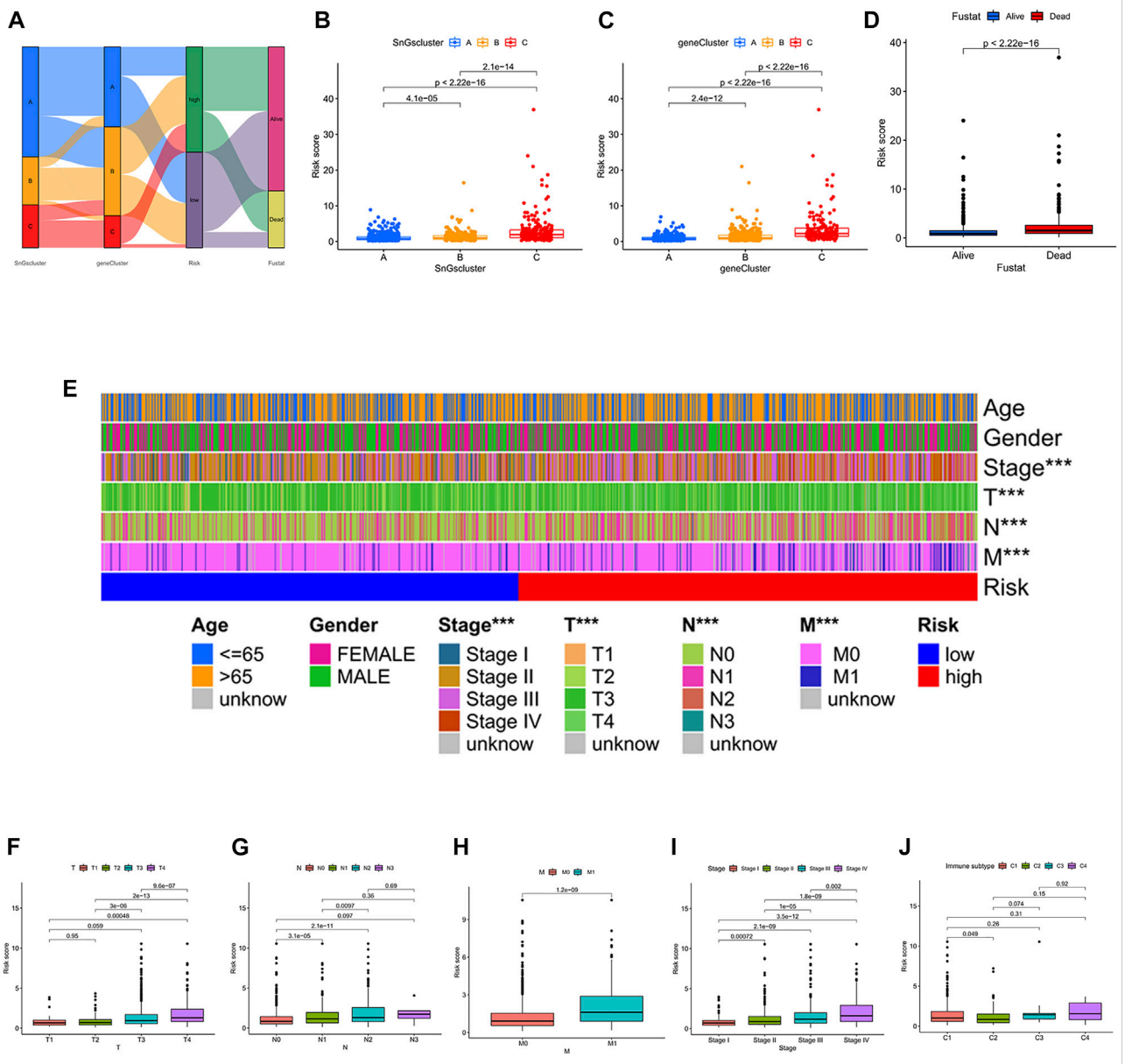


FIGURE 7
(A) Sankey diagram. The construction process of the risk score signature. **(B)** Box plot of risk scores for the three clusters. **(C)** Box plot of risk scores for the three gene clusters. **(D)** Box plot of risk scores for the survival status (Fustat). **(E)** Heatmap of correlation between risk and clinical characteristics. $^{***}p < 0.001$. **(F)** Box plot of the risk score for the T stage. **(G)** Box plot of risk scores for the N stage. **(H)** Box plot of risk scores for M stage. **(I)** Box plot of risk scores for pathological TNM stage. **(J)** A box plot of risk scores across the four immune subtypes. C1, Wound Healing; C2, IFN-gamma Dominant; C3, Inflammatory; C4, Lymphocyte Depleted.

developing cancer (Leong et al., 2021; Panegyres and Chen, 2021). In contrast, the enriched pathways in cluster C were involved in the development and metastasis of cancer (Zhao et al., 2018; Bao et al., 2019; Pudełko et al., 2019; Wu et al., 2021). Furthermore, the deficiencies of "BASE EXCISION REPAIR" and "HOMOLOGOUS RECOMBINATION" in cluster C were thought to be associated with a worse prognosis of cancer (Wallace et al., 2012; Toh and Ngeow, 2021). Next, 2334 DEGs among three clusters were identified. We observed that significantly enriched pathways in DEGs were associated with tumor development and metastasis (Bourboulia and Stetler-Stevenson, 2010; Huang et al., 2017; Izdebska et al., 2018; Li et al., 2021; Lin et al., 2022). It confirmed that cellular senescence played a crucial role in the prognosis of colon cancer. In order to further validate the above speculation, 2334 DEGs among clusters were performed consensus clustering analysis. All

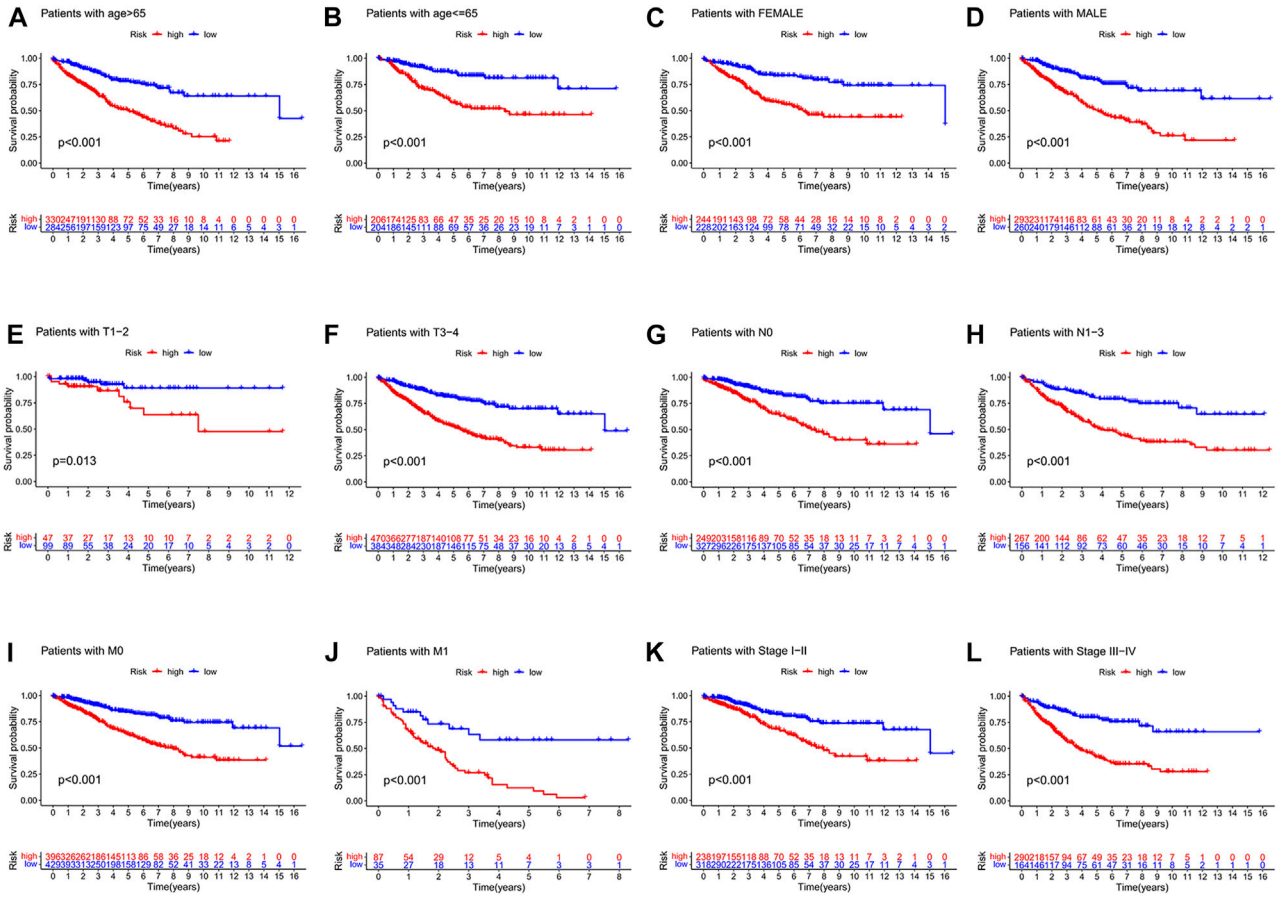


FIGURE 8
(A-L) visualized the risk score signature being applied to patients with different clinical subgroups, including different ages, different genders, different T-stages, different N-stages, different M-stages, and different pathological stages. **(A)** Patients with age > 65. **(B)** Patients with age ≦ 65. **(C)** Female patients. **(D)** Male patients. **(E)** Patients with stages T1-2. **(F)** Patients with stage T3-4. **(G)** Patients with stage N0. **(H)** Patients with stages N1-3. **(I)** Patients with stage M0. **(J)** Patients with stage M1. **(K)** Patients with stages I-II. **(L)** Patients with stages III-IV.

colon cancer patients were divided into three gene clusters. Gene cluster A had the best prognosis, while gene cluster C had the worst prognosis. It suggested that cellular senescence could also distinguish prognostic differences among gene clusters. Therefore, we considered that patients with different prognoses of colon cancer could be distinguished based on cellular senescence.

Next, according to 681 prognosis-related DEGs among clusters, a cellular senescence-related risk score signature (FITM2, APOL6, VWF, PRRX2, CCL22, ALPL, SON, KIF7, and ZEB1-AS1) was constructed to predict patients' prognosis. Low-risk score group showed longer survival and a lower percentage of deaths. The risk score could be used independently of other clinical features (age, gender, stage, T stage, N stage, and M stage) to predict patients' prognosis with the highest accuracy. Moreover, compared to other signatures, the cellular senescence-related risk score signature had the highest predictive accuracy with a C-index value of 0.682. Excitingly, we observed that the cellular senescence-related risk score signature predicted little heterogeneity in prognosis between the training set and testing set by performing the prognostic meta-analysis with $I^2 < 50\%$. It further confirmed the accuracy of the signature in predicting the prognosis of colon cancer patients. A nomogram predicting 1-year, 3-years, and 5-years survival probability in patients with colon cancer was constructed for the clinical work. It has the highest accuracy compared to other clinical characteristics, with an AUC value of 0.823. According to the comparison of risk scores in different subgroups, we observed the following phenomena: more advanced TNM stage was associated with higher risk scores; cluster C had a significantly higher risk score than cluster A and cluster B; gene cluster C had a significantly higher risk score than gene cluster A; the clinical outcome death group had a significantly higher risk score than the survival group. It further confirmed that the high-risk score group was associated with a worse prognosis, while the low-risk score group was associated with a better prognosis. There

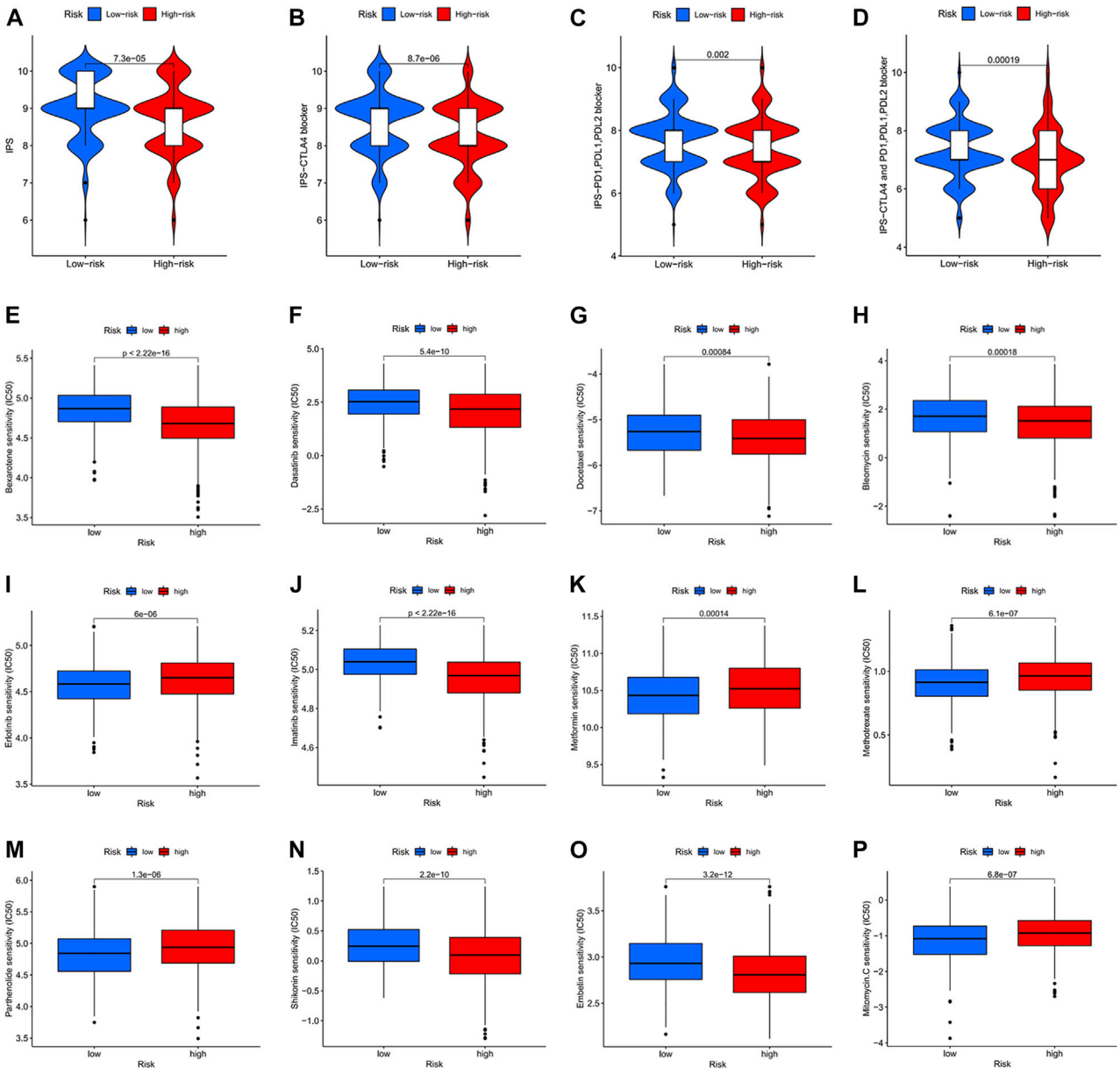


FIGURE 9
Immunotherapy response and chemotherapy response in colon cancer. **(A)** Immunotherapy response of different risk score groups for CTLA4-negative and PD1, PDL1, PDL2 negative. p = 7.3e-05. **(B)** Immunotherapy response of different risk score groups for CTLA4 positive and PD1, PDL1, PDL2 negative. p = 8.7e-06. **(C)** Immunotherapy response of different risk score groups for CTLA4 negative and PD1, PDL1, PDL2 positive. p = 0.002. **(D)** Immunotherapy response of different risk score groups to CTLA4 positivity and PD1, PDL1, PDL2 negativity. p = 0.00019. **(E–P)** Chemotherapy response of different risk score groups to 12 chemotherapy drugs.

was very high accuracy in distinguishing between high and low-risk score groups based on the risk score signature. We also demonstrated the suitability of the signature for predicting prognosis in different clinical subgroups, including different age groups (≤65 and >65), different gender (female and male), different T stages (T-2 and T3-4), different N stages (N0 and N1-3), different M stages (M0 and M1), and different TNM stages (stage I-II and stage III-IV). Since it is impossible to identify which colon cancer patients benefit from immunotherapy in clinical work, which often leads to misuse of immunotherapy drugs. Therefore, we performed further analysis to explore whether the signature could distinguish colon cancer patients who have better immunotherapy responses for targeted treatment. The low-risk score group had a better immunotherapy response. While the low-risk score group had

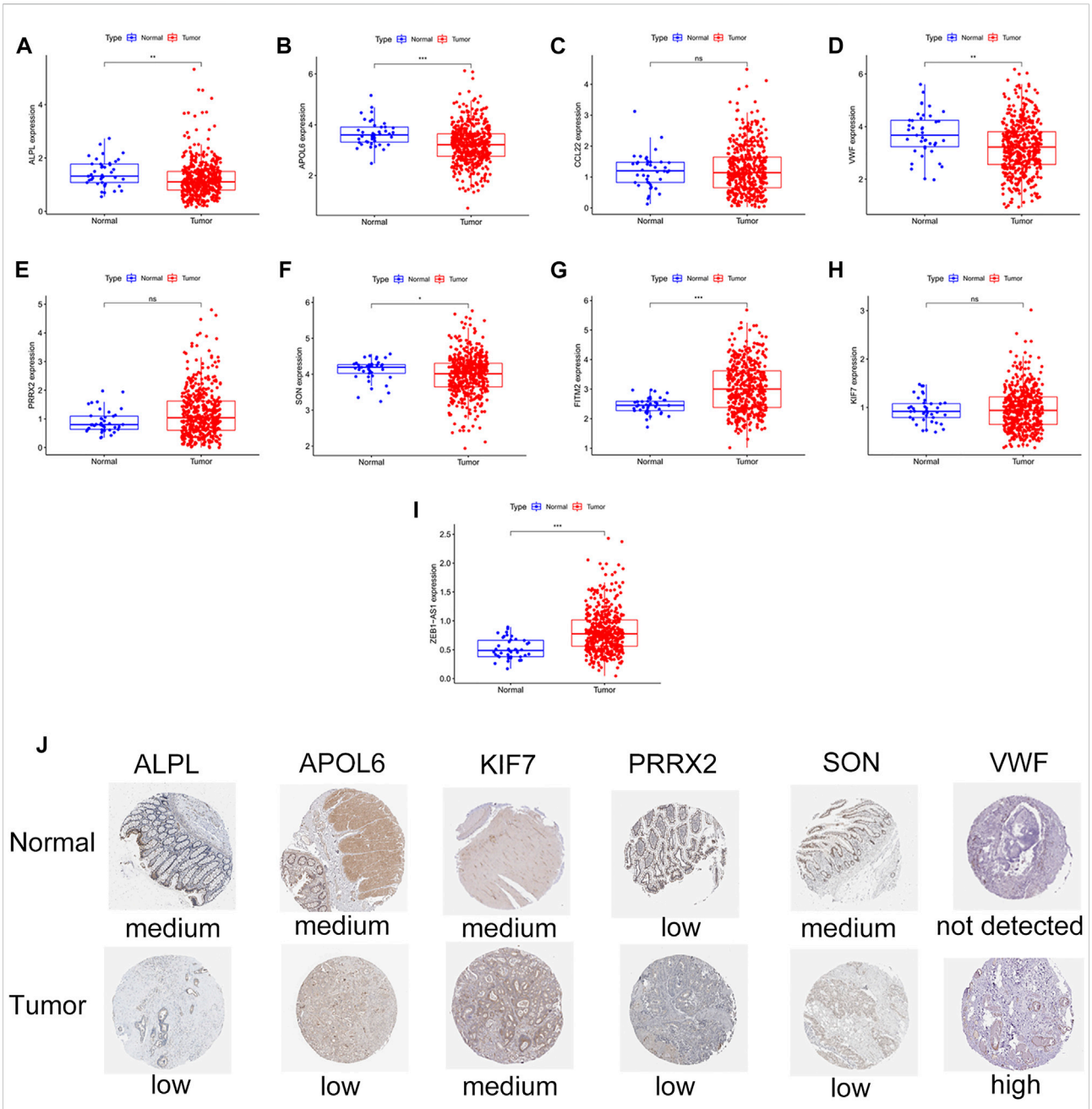


FIGURE 10
Validation of differential expression of FITM2, APOL6, VWF, PRRX2, CCL22, ALPL, SON, KIF7, and ZEB1-AS1 in normal samples and colon cancer samples. **(A–I)** Box plots. ns represents no difference; $^{*}p < 0.05$, $^{**}p < 0.001$, $^{***}p < 0.001$. **(J)** Immunohistochemical maps of ALPL, APOL6, KIF7, PRRX2, SON, and VWF expression proteins in the HPA database.

a worse immunotherapy response. Therefore, better benefits may be achieved when immunotherapeutic drugs (PD1/PDL-1/PD-L2/CTLA-4 blockers) are used for colon cancer patients in the low-risk score group.

Finally, FITM2, APOL6, VWF, PRRX2, CCL22, ALPL, SON, KIF7, and ZEB1-AS were further investigated. In our study, FITM2 was highly expressed in colon cancer tissues. It was consistent with the findings of Yang et al. (Yang et al., 2019). We demonstrated that TITM2 was a low-risk gene, which was associated with a better prognosis in patients with colon cancer. APOL6 showed a low expression level in colon cancer tissues and was a low-risk gene. It was due to the ability of APOL6 to induce apoptosis in colon cancer cells (Aryee et al., 2013). In our article, VWF was low expressed in colon cancer

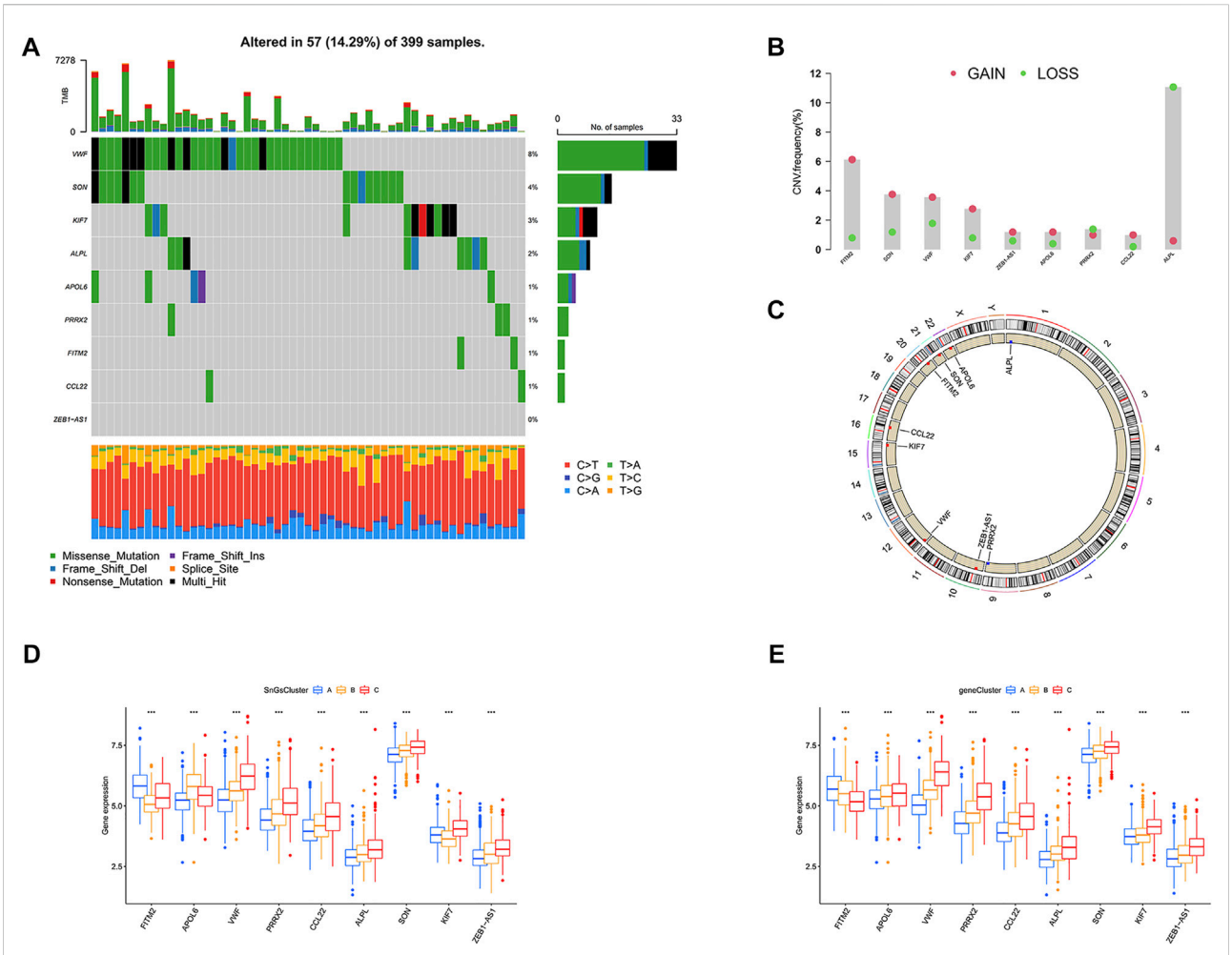


FIGURE 11
The landscape for mutation and expression of the 10 signature genes (FITM2, APOL6, VWF, PRRX2, CCL22, ALPL, SON, KIF7, and ZEB1-AS1). **(A)** The gene mutation frequency waterfall plot of 10 signature genes. **(B)** CNV alteration frequency of the 10 signature genes; Red, CNV gain; Green, CNV loss. **(C)** CNV changes in 10 signature genes on 23 chromosomes. Red, CNV increase; Blue, copy number loss. **(D)** Differential expression of 10 signature genes among three clusters. ***$p < 0.001$. **(E)** Differential expression of 10 signature genes among three gene clusters. ***$p < 0.001$.

tissue and was a high-risk gene. It was because VWF could promote a highly aggressive nature of colon cancer (Zanetta et al., 2000). Our study showed that PRRX2 was a high-risk gene. However, Chai et al. considered that PRRX2 inhibited distant metastasis of colon cancer cells and was a protective gene (Chai et al., 2019). This difference required further verification in subsequent experiments. Chen et al. revealed that high expression of CCL22 was associated with a better prognosis in patients with colon cancer (Chen et al., 2021). It was consistent with our findings. Luo et al. found that ALPL inhibited the aggressiveness of ovarian cancer (Luo et al., 2019). And Child et al. identified ALPL as a cancer suppressor gene for prostate cancer (Tong et al., 2019). However, the opposite was true for the role of ALPL in colon cancer. In our study, ALPL was a high-risk gene that was lowly expressed in colon cancer tissues. The significance of SON in colon cancer has not been studied by anyone. We first identified SON as a high-risk gene with low expression in colon cancer tissues. Hu et al. revealed that downregulation of KIF7 promoted antitumor activity in lung cancer and it was a cancer-promoting gene (Hu et al., 2020). Surprisingly, we also found KIF7 as a high-risk gene in colon cancer. In our article, ZEB1-AS1 was highly expressed in colon cancer tissues and was associated with a worse prognosis. This was associated with the ability of ZEB1-AS1 to cause the malignant progression of colon cancer (Ni et al., 2020). We also found that VWF had the highest mutation frequency, while ZEB1-AS1 was not mutated. All genes showed amplification except for PRRX2 and ALPL which showed depletion. Most of the signature genes were upregulated in cluster C and gene cluster C.

In conclusion, the cellular senescence-related risk score signature could be used to predict prognosis and immunotherapy response in colon cancer.

Data availability statement

Publicly available datasets were analyzed in this study. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

LD designed the study flow and wrote the manuscript. JL, TL, and BC performed the gastric cancer data acquisition and collation. XW, LD, TL, and TB drew all the figures. Finally, WY reviewed and revised the paper in detail.

Acknowledgments

Thanks to the following databases for supporting the article, including the TCGA database, GEO database, CellAge database, TCIA database, and HPA database.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.961554/full#supplementary-material

SUPPLEMENTARY FIGURE S1
The design flow chart for the research.

SUPPLEMENTARY FIGURE S2
Clustering results based on the expression of 279 cellular senescence-related genes. Cluster Num = 2, 4, 5, 6, 7, 8, 9. [Not Available in CrossRef].

SUPPLEMENTARY FIGURE S3
Identification of DEGs in clusters. **(A)** Venn diagram. 2334 DEGs were acquired by taking intersections of cluster AB, cluster AC, and cluster BC differential genes. **(B)** Bubble diagram of GO enrichment. BP, biological process; MF, molecular function; CC, cellular component; vertical coordinates represent GO names, horizontal coordinates represent gene proportions; circle size represents the number of genes enriched on GO; circle color represents enrichment significance. **(C)** KEGG enrichment bubble diagram. Vertical coordinates represent GO names, horizontal coordinates represent gene proportions; circle size represents the number of genes enriched on GO; circle color represents enrichment significance.

SUPPLEMENTARY FIGURE S4
Clustering results based on the expression of 681 prognostic DEGs among clusters. Cluster Num = 2, 3, 4, 5, 6, 7, 8, 9. When cluster Num = 3, the relationship in the gene cluster was tight and the correlation between gene clusters was weak.

SUPPLEMENTARY FIGURE S5
(A–D) ROC curves for the risk score signature of Ren, Wang, Zhang, and Zheng for predicting 1-, three- and 5-years survival in patients with colon cancer. **(E–H)** Patients with colon cancer were classified into high and low-risk score groups based on the risk score signature of Ren, Wang, Zhang, and Zheng. Patients in the low-risk score group had a better prognosis.

SUPPLEMENTARY FIGURE S6
Waterfall plots of gene mutations in different risk score groups. **(A)** High-risk score group. **(B)** Low-risk score group.

SUPPLEMENTARY TABLE S1
279 cellular senescence-related genes were acquired from the CellAge database.

SUPPLEMENTARY TABLE S2
2334 DEGs among Cluster A, Cluster B, and Cluster C.

SUPPLEMENTARY TABLE S3
681 prognostic DEGs among Cluster A, Cluster B, and Cluster C.

SUPPLEMENTARY TABLE S4
The statistics of clinical information for all colon cancer samples, the training set, and the testing set.

References

Acosta, J. C., Banito, A., Wuestefeld, T., Georgilis, A., Janich, P., Morton, J. P., et al. (2013). A complex secretory program orchestrated by the inflammasome controls paracrine senescence. *Nat. Cell. Biol.* 15 (8), 978–990. doi:10.1038/ncb2784

Aryee, D. N., Niedan, S., Ban, J., Schwentner, R., Muehlbacher, K., Kauer, M., et al. (2013). Variability in functional p53 reactivation by PRIMA-1(Met)/APR-246 in Ewing sarcoma. *Br. J. Cancer* 109 (10), 2696–2704. doi:10.1038/bjc.2013.635

Azwar, S., Seow, H. F., Abdullah, M., Faisal Jabar, M., and Mohtarrudin, N. (2021). Recent updates on mechanisms of resistance to 5-fluorouracil and reversal strategies in colon cancer treatment. *Biol. (Basel)* 10 (9), 854. doi:10.3390/biology10090854

Bao, Y., Wang, L., Shi, L., Yun, F., Liu, X., Chen, Y., et al. (2019). Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. *Cell. Mol. Biol. Lett.* 24, 38. doi:10.1186/s11658-019-0162-0

Ben, S., Zhu, Q., Chen, S., Li, S., Du, M., Xin, J., et al. (2021). Genetic variations in the CTLA-4 immune checkpoint pathway are associated with colon cancer risk,

prognosis, and immune infiltration via regulation of IQCB1 expression. *Arch. Toxicol.* 95 (6), 2053–2063. doi:10.1007/s00204-021-03040-0

Body, A., Prenen, H., Latham, S., Lam, M., Tipping-Smith, S., Raghunath, A., et al. (2021). The role of neoadjuvant chemotherapy in locally advanced colon cancer. *Cancer Manag. Res.* 13, 2567–2579. doi:10.2147/CMAR.S262870

Bourboulia, D., and Stetler-Stevenson, W. G. (2010). Matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs): Positive and negative regulators in tumor cell adhesion. *Semin. Cancer Biol.* 20 (3), 161–168. doi:10.1016/j.semcancer.2010.05.002

Calcinotto, A., Kohli, J., Zagato, E., Pellegrini, L., Demaria, M., Alimonti, A., et al. (2019). Cellular senescence: Aging, cancer, and injury. *Physiol. Rev.* 99 (2), 1047–1078. doi:10.1152/physrev.00020.2018

Chai, W. X., Sun, L. G., Dai, F. H., Shao, H. S., Zheng, N. G., Cai, H. Y., et al. (2019). Inhibition of PRRX2 suppressed colon cancer liver metastasis via inactivation of Wnt/β-catenin signaling pathway. *Pathol. Res. Pract.* 215 (10), 152593. doi:10.1016/j.prp.2019.152593

Chakrabarti, S., Peterson, C. Y., Sriram, D., and Mahipal, A. (2020). Early stage colon cancer: Current treatment standards, evolving paradigms, and future directions. *World J. Gastrointest. Oncol.* 12 (8), 808–832. doi:10.4251/wjgo.v12.i8.808

Chen, W., Huang, J., Xiong, J., Fu, P., Chen, C., Liu, Y., et al. (2021). Identification of a tumor microenvironment-related gene signature indicative of disease prognosis and treatment response in colon cancer. *Oxid. Med. Cell. Longev.* 2021, 6290261. doi:10.1155/2021/6290261

Cho, Y. Y., Kim, D. J., Lee, H. S., Jeong, C. H., Cho, E. J., Kim, M. O., et al. (2013). Autophagy and cellular senescence mediated by Sox2 suppress malignancy of cancer cells. *PLoS One* 8 (2), e57172. doi:10.1371/journal.pone.0057172

Demirci, D., Dayanc, B., Mazi, F. A., and Senturk, S. (2021). The Jekyll and Hyde of cellular senescence in cancer. *Cells* 10 (2), 208. doi:10.3390/cells10020208

Di Micco, R., Krizhanovsky, V., Baker, D., and d'Adda di Fagagna, F. (2021). Cellular senescence in ageing: From mechanisms to therapeutic opportunities. *Nat. Rev. Mol. Cell. Biol.* 22 (2), 75–95. doi:10.1038/s41580-020-00314-w

Guo, J. N., Chen, D., Deng, S. H., Huang, J. R., Song, J. X., Li, X. Y., et al. (2022). Identification and quantification of immune infiltration landscape on therapy and prognosis in left- and right-sided colon cancer. *Cancer Immunol. Immunother.* 71 (6), 1313–1330. doi:10.1007/s00262-021-03076-2

Guo, Y., Ayers, J. L., Carter, K. T., Wang, T., Maden, S. K., Edmond, D., et al. (2019). Senescence-associated tissue microenvironment promotes colon cancer formation through the secretory factor GDF15. *Aging Cell.* 18 (6), e13013. doi:10.1111/acel.13013

Hu, Y., Wu, M. Z., Gu, N. J., Xu, H. T., Li, Q. C., Wu, G. P., et al. (2020). Human papillomavirus 16 (HPV 16) E6 but not E7 inhibits the antitumor activity of LKB1 in lung cancer cells by downregulating the expression of KIF7. *Thorac. Cancer* 11 (11), 3175–3180. doi:10.1111/1759-7714.13640

Huang, Y., Zhang, J., Hou, L., Wang, G., Liu, H., Zhang, R., et al. (2017). LncRNA AK023391 promotes tumorigenesis and invasion of gastric cancer through activation of the PI3K/Akt signaling pathway. *J. Exp. Clin. Cancer Res.* 28 (1), 36. doi:10.1186/s13046-017-0666-2

Izdebska, M., Zielińska, W., Grzanka, D., and Gagat, M. (2018). The role of actin dynamics and actin-binding proteins expression in epithelial-to-mesenchymal transition and its association with cancer progression and evaluation of possible therapeutic targets. *Biomed. Res. Int.*, 2018, 4578373. doi:10.1155/2018/4578373

Katopodi, T., Petanidis, S., Domvri, K., Zarogoulidis, P., Anestakis, D., Charalampidis, C., et al. (2021). Kras-driven intratumoral heterogeneity triggers infiltration of M2 polarized macrophages via the circHIPK3/PTK2 immunosuppressive circuit. *Sci. Rep.* 11 (1), 15455. doi:10.1038/s41598-021-94671-x

Kim, J. A. (2020). Peroxisome metabolism in cancer. *Cells* 9 (7), 1692. doi:10.3390/cells9071692

Leong, Y. Q., Lee, S. W. H., and Ng, K. Y. (2021). Cancer risk in Parkinson disease: An updated systematic review and meta-analysis. *Eur. J. Neurol.* 28 (12), 4219–4237. doi:10.1111/ene.15069

Li, H., Liang, J., Wang, J., Han, J., Li, S., Huang, K., et al. (2021). Mex3a promotes oncogenesis through the RAP1/MAPK signaling pathway in colorectal cancer and is inhibited by hsa-miR-6887-3p. *Cancer Commun.* 41 (6), 472–491. doi:10.1002/cac2.12149

Lichtenstern, C. R., Ngu, R. K., Shalapour, S., and Karin, M. (2020). Immunotherapy, inflammation and colorectal cancer. *Cells* 9 (3), 618. doi:10.3390/cells9030618

Lin, W., Wang, X., Wang, Z., Shao, F., Yang, Y., Cao, Z., et al. (2021). Comprehensive analysis uncovers prognostic and immunogenic characteristics of cellular senescence for lung adenocarcinoma. *Front. Cell. Dev. Biol.* 9, 780461. doi:10.3389/fcell.2021.780461

Lin, X., Zhuang, S., Chen, X., Du, J., Zhong, L., Ding, J., et al. (2022). lncRNA ITGB8-AS1 functions as a ceRNA to promote colorectal cancer growth and migration through integrin-mediated focal adhesion signaling. *Mol. Ther.* 30 (2), 688–702. doi:10.1016/j.ymthe.2021.08.011

Luo, M., Zhou, L., Zhan, S. J., Cheng, L. J., Li, R. N., Wang, H., et al. (2019). ALPL regulates the aggressive potential of high grade serous ovarian cancer cells via a non-canonical WNT pathway. *Biochem. Biophys. Res. Commun.* 513 (2), 528–533. doi:10.1016/j.bbrc.2019.04.016

Marisa, L., Blum, Y., Taieb, J., Ayadi, M., Pilati, C., Le Malicot, K., et al. (2021). Intratumor CMS heterogeneity impacts patient prognosis in localized colon cancer. *Clin. Cancer Res.* 27 (17), 4768–4780. doi:10.1158/1078-0432.CCR-21-0529

Ni, X., Ding, Y., Yuan, H., Shao, J., Yan, Y., Guo, R., et al. (2020). Long non-coding RNA ZEB1-AS1 promotes colon adenocarcinoma malignant progression via miR-455-3p/PAK2 axis. *Cell. Prolif.* 53 (1), e12723. doi:10.1111/cpr.12723

Panegyres, P. K., and Chen, H. Y. (2021). Alzheimer's disease, Huntington's disease and cancer. *J. Clin. Neurosci.* 93, 103–105. doi:10.1016/j.jocn.2021.09.012

Prasanna, P. G., Citrin, D. E., Hildesheim, J., Ahmed, M. M., Venkatachalam, S., Riscuta, G., et al. (2021). Therapy-induced senescence: Opportunities to improve anticancer therapy. *J. Natl. Cancer Inst.* 113 (10), 1285–1298. doi:10.1093/jnci/djab064

Prieto, L. I., and Baker, D. J. (2019). Cellular senescence and the immune system in cancer. *Gerontology* 65 (5), 505–512. doi:10.1159/000500683

Pudełko, A., Wisowski, G., Olczyk, K., and Koźma, E. M. (2019). The dual role of the glycosaminoglycan chondroitin-6-sulfate in the development, progression and metastasis of cancer. *FEBS J.* 286 (10), 1815–1837. doi:10.1111/febs.14748

Ren, J., Feng, J., Song, W., Wang, C., Ge, Y., Fu, T., et al. (2020). Development and validation of a metabolic gene signature for predicting overall survival in patients with colon cancer. *Clin. Exp. Med.* 20 (4), 535–544. doi:10.1007/s10238-020-00652-1

Sikora, E., Bielak-Zmijewska, A., and Mosieniak, G. (2021). A common signature of cellular senescence; does it exist? *Ageing Res. Rev.* 71, 101458. doi:10.1016/j.arr.2021.101458

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Tian, X., Shen, H., Li, Z., Wang, T., and Wang, S. (2019). Tumor-derived exosomes, myeloid-derived suppressor cells, and tumor microenvironment. *J. Hematol. Oncol.* 12 (1), 84. doi:10.1186/s13045-019-0772-z

Togashi, Y., and Nishikawa, H. (2017). Regulatory T cells: Molecular and cellular basis for immunoregulation. *Curr. Top. Microbiol. Immunol.* 410, 3–27. doi:10.1007/82_2017_58

Toh, M., and Ngeow, J. (2021). Homologous recombination deficiency: Cancer predispositions and treatment implications. *Oncologist* 26 (9), e1526–e1537. doi:10.1002/onco.13829

Tong, Y., Song, Y., and Deng, S. (2019). Combined analysis and validation for DNA methylation and gene expression profiles associated with prostate cancer. *Cancer Cell. Int.* 19, 50. doi:10.1186/s12935-019-0753-x

Wallace, S. S., Murphy, D. L., and Sweasy, J. B. (2012). Base excision repair and cancer. *Cancer Lett.* 327 (1-2), 73–89. doi:10.1016/j.canlet.2011.12.038

Wang, B., and Demaria, M. (2021). The quest to define and target cellular senescence in cancer. *Cancer Res.* 81 (24), 6087–6089. doi:10.1158/0008-5472.CAN-21-2032

Wang, L., Lankhorst, L., and Bernards, R. (2022). Exploiting senescence for the treatment of cancer. *Nat. Rev. Cancer* 22 (6), 340–355. doi:10.1038/s41568-022-00450-9

Wang, Y., Xia, H. B., Chen, Z. M., Meng, L., and Xu, A. M. (2021). Identification of a ferroptosis-related gene signature predictive model in colon cancer. *World J. Surg. Oncol.* 19 (1), 135. doi:10.1186/s12957-021-02244-z

Wenes, M., Jaccard, A., Wyss, T., Maldonado-Pérez, N., Teoh, S. T., Lepez, A., et al. (2022). The mitochondrial pyruvate carrier regulates memory T cell differentiation and antitumor function. *Cell. Metab.* 34 (5), 731–746.e9. e9. doi:10.1016/j.cmet.2022.03.013

Wu, S., Chen, M., Huang, J., Zhang, F., Lv, Z., Jia, Y., et al. (2021). ORAI2 promotes gastric cancer tumorigenicity and metastasis through PI3K/akt signaling and MAPK-dependent focal adhesion disassembly. *Cancer Res.* 81, 986–1000. doi:10.1158/0008-5472.CAN-20-0049

Yaghoubi, N., Soltani, A., Ghazvini, K., Hassanian, S. M., and Hashemy, S. I. (2019). PD-1/PD-L1 blockade as a novel treatment for colorectal cancer. *Biomed. Pharmacother.* 110, 312–318. doi:10.1016/j.biopha.2018.11.105

Yang, L., Li, L., Ma, J., Yang, S., Zou, C., Yu, X., et al. (2019). miRNA and mRNA integration network construction reveals novel key regulators in left-sided and right-sided colon adenocarcinoma. *Biomed. Res. Int.* 2019, 7149296. doi:10.1155/2019/7149296

Zanetta, L., Marcus, S. G., Vasile, J., Dobryansky, M., Cohen, H., Eng, K., et al. (2000). Expression of von willebrand factor, an endothelial cell marker, is up-regulated by angiogenesis factors: A potential method for objective assessment of tumor angiogenesis. *Int. J. Cancer* 85 (2), 281–288. doi:10.1002/(sici)1097-0215(20000115)85:2<281::aid-ijc21>3.0.co;2-3

Zhang, X., Zhao, H., Shi, X., Jia, X., and Yang, Y. (2020). Identification and validation of an immune-related gene signature predictive of overall survival in colon cancer. *Aging (Albany NY)* 12 (24), 26095–26120. doi:10.18632/aging.202317

Zhao, M., Mishra, L., and Deng, C. X. (2018). The role of TGF-β/SMAD4 signaling in cancer. *Int. J. Biol. Sci.* 14 (2), 111–123. doi:10.7150/ijbs.23230

Zheng, L., Yang, Y., and Cui, X. (2021). Establishing and validating an aging-related prognostic four-gene signature in colon adenocarcinoma. *Biomed. Res. Int.* 2021, 4682589. doi:10.1155/2021/4682589

Zhou, L., Niu, Z., Wang, Y., Zheng, Y., Zhu, Y., Wang, C., et al. (2022). Senescence as a dictator of patient outcomes and therapeutic efficacies in human gastric cancer. *Cell. Death Discov.* 8 (1), 13. doi:10.1038/s41420-021-00769-6

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fgene.2022.956866

OPEN ACCESS

EDITED BY
Jaspreet Kaur Dhanjal,
Indraprastha Institute of Information Technology Delhi, India

REVIEWED BY
Aditya Yashwant Sarode,
Columbia University, United States
Konstantinos Spanos,
University of Thessaly, Greece

*CORRESPONDENCE
Yi Si,
sysiy@yahoo.com
Weiguo Fu,
fu.weiguo@zs-hospital.sh.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 30 May 2022
ACCEPTED 28 July 2022
PUBLISHED 02 September 2022

CITATION
Huang L, Tang J, Lin L, Wang R, Chen F, Wei Y, Si Y and Fu W (2022), Association of genetic variants in ULK4 with the age of first onset of type B aortic dissection. *Front. Genet.* 13:956866. doi: 10.3389/fgene.2022.956866

Association of genetic variants in *ULK4* with the age of first onset of type B aortic dissection

Lihong Huang[1,2,3†], Jiaqi Tang[4†], Lijuan Lin[4], Ruihan Wang[1], Feng Chen[4], Yongyue Wei[4], Yi Si[1*] and Weiguo Fu[1*]

[1]Department of Vascular Surgery, Zhongshan Hospital, Fudan University, Shanghai, China, [2]Department of Biostatistics, Zhongshan Hospital, Fudan University, Shanghai, China, [3]Clinical Research Unit, Institute of Clinical Science, Zhongshan Hospital of Fudan University, Shanghai, China, [4]Department of Biostatistics, School of Public Health, Nanjing Medical University, Nanjing, China

Background: The association between autophagy, structural alterations of the aortic wall, and endothelial dysfunction in humans has yet to be fully elucidated. The family of *ULK* (UNC51-like) enzymes plays critical roles in autophagy and development. This study aimed to evaluate the association between *ULK* gene family members and patient age of first type B aortic dissection (TBAD) onset.

Methods: The genotype data in a TBAD cohort from China and the related summary-level datasets were analyzed. We applied the sequence kernel association test (SKAT) to test the association between single-nucleotide polymorphisms (SNPs) and age of first onset of TBAD controlling for gender, hypertension, and renal function. Next, we performed a 2-sample Mendelian randomization (MR) to explore the potential causal relationship between *ULK4* and early onset of TBAD at the level of gene expression coupled with DNA methylation with genetic variants as instrumental variables.

Results: A total of 159 TBAD patients with 1,180,097 SNPs were included. Concerning the association between the *ULK* gene family and the age of first onset of the TBAD, only *ULK4* was found to be significant according to SKAT analysis (q-*FDR* = 0.0088). From 2-sample MR, the high level of *ULK4* gene expression was related to a later age of first onset of TBAD (β = 4.58, p = 0.0214).

Conclusion: This is the first study of the *ULK* gene family in TBAD, regarding the association with the first onset age. We demonstrated that the *ULK4* gene is associated with the time of onset of TBAD based on both the SKAT and 2-sample MR analyses.

KEYWORDS

type B aortic dissection, *ULK4*, age first onset, genetic variants, gene expression

Abbreviations: eQTL, expression quantitative trait loci; KAT, sequence kernel association test; MAF, minor allele frequency; MFS, Marfan syndrome; mQTL, methylation quantitative trait loci; MR, Mendelian randomization; STBAD, type B aortic dissection; VSMCs, vascular smooth muscle cells.

Introduction

Type B aortic dissection (TBAD) is a rare while life-threatening condition, in which a tear occurs in the descending part of the aorta and may extend into the abdomen (Prêtre and Von Segesser, 1997; Hagan et al., 2000; Nienaber and Clough, 2015). Prevention of premature death from TBAD depends on the early identification of high-risk individuals, careful monitoring of the dissected aorta for aneurysmal dilations, medications to slow the rate of growth of aneurysms, and timely surgical repair of aneurysms (Mokashi and Svensson, 2019).

Aortic expansion is one of the risk factors associated with the need for intervention or adverse outcomes in patients with TBAD. It was reported that younger age at presentation was a clinical predictor of aortic expansion. Patient age <60 years was significantly associated with increasing aortic diameter, which was thought to be due to a less rigid aortic wall, making the aorta more prone to dilation in younger patients (Kamman et al., 2017). However, the essential reasons for this finding deserve further research, and the association of genetic variants with the age of first onset of type B aortic dissections is a valuable research direction.

Autophagy is a process in which intracellular components and dysfunctional organelles are delivered to the lysosome for degradation and recycling. Therefore, autophagy has various connections to many human diseases, as its functions are essential for cell survival, bioenergetic homeostasis, organism development, and cell death regulation. The association between autophagy and structural alterations of the aortic wall and endothelial dysfunction in humans has yet to be fully elucidated (Yang and Klionsky, 2020). Previous studies have shown that more than 20% of individuals with thoracic aortic aneurysms and dissections have a family history of disease that may be caused by a genetic syndrome, resulting from a single-gene mutation such as Marfan syndrome (MFS [MIM:154700]) arising from a fibrillin-1 (FBN1 [MIM:134797]) mutation (Biddinger et al., 1997). However, autophagy-related biomarkers studies for aortic dissection diseases are still rare. According to existing research, AMPK increases the process of autophagy after its activation. Although the mechanisms of AD and autophagy have not been fully elucidated, autophagy has been observed to be activated in impaired vascular smooth muscle cells (VSMCs). Excessive or impaired autophagy may lead to VSMC death or dysfunction, which is thought to promote aneurysm and AD (Clément et al., 2019).

It is known that the *ULK* (UNC51-like) enzymes are a family of mammalian kinases and play critical roles in autophagy and development. The mammalian *ULK* family of kinases comprises 5 genes: *ULK1* to *ULK4* and *STK36* (Chan and Tooze, 2009). These enzymes share a conserved N-terminal kinase domain, which is homologous to *C. elegans* UNC51 and yeast Atg1, the original kinase identified in the autophagy pathway. *ULK* kinases can be found in all observed eukaryotes. *ULK1* kinases are involved in autophagy (Young et al., 2009), *ULK3* is also implicated in hedgehog signaling and in autophagy-mediated senescence (Maloverjan et al., 2010), and several genome-wide association studies (GWASs) show linkage to blood pressure (Levy et al., 2009). *ULK4* is a pseudo kinase in all species and may be linked to neurogenesis, brain function (Liu et al., 2017), and blood pressure (Levy et al., 2009). It has also been reported that *ULK4* is potentially associated with acute aortic dissections (Guo et al., 2016). As can be seen, there is demonstrable link between the *ULK* family, autophagy, and acute aortic dissection, and thus an assessment of the effect of the *ULK* family on aortic dissection is warranted.

We performed an association analysis between *ULK* gene family members and the age of first onset of Chinese TBAD patients. The association study was performed in Chinese TBAD patients. The causal effect of *ULK* genes was subsequently verified by 2-sample Mendelian randomization (MR) at the gene expression and DNA methylation levels.

Materials and methods

Study population and data source

We obtained genotype data from a TBAD cohort enrolled through the Vascular Surgery Department of Zhongshan Hospital Fudan University, which included 162 Chinese patients with TBAD from January 2018 to June 2019. Each participant signed a consent form. The study was approved by the relevant ethics committees (Ethical approval No. B2019-110R) and was administered by trained personnel.

Genotyping and quality control of whole genome sequencing data

In association analysis, we obtained genotype data using whole genome sequencing, using the Illumina NovaSeq platform (Illumina, San Diego, CA, United States) in a paired-end 150 bp mode on 162 TBAD patients. Briefly, the samples were excluded if they met any of the following quality control (QC) criteria (Supplementary Figure S1): 1) overall genotype completion rate <95%, 2) unexpected duplicates or probable relatives, 3) heterozygosity rates more than six times the SD from the mean, or 4) gender discrepancies. SNPs were excluded if they met any of the following QC criteria: 1) SNPs had a low call rate of <95% in all samples, 2) the genotype distributions of SNPs deviated from those expected by the Hardy–Weinberg equilibrium ($p < 0.000001$), or 3) single-nucleotide variants (SNVs) with minor allele frequencies were less than 1%.

Summary-level data of expression quantitative trait loci and methylation quantitative trait loci

Methylation quantitative trait loci (mQTL) data were obtained from the Brisbane Systems Genetics Study (n = 614) and Lothian Birth Cohorts of 1921 and 1936 (n = 1366) (Wu et al., 2018). Details of the QC procedures were described in a previous study. Briefly, all the individuals were of European descent. Only the DNA methylation probes with at least a cis-mQTL at $p < 0.0001$ and only SNPs within 500 kb distance from each probe were included in the analysis. As for the summary statistics of expression quantitative trait loci (eQTL), we used the cis-eQTL in the prefrontal cortex from the PsychENCODE project (n = 1387) (Gandal et al., 2018; Wang et al., 2018). The eQTL analyses of PsychENCODE were performed by including 100 hidden covariate factors as covariates. Only the data of SNPs in a 500 kb window around the *ULK4* gene were included in the subsequent analysis.

Statistical analysis

Continuous variables are summarized as mean ± SD and categorical variables are described as numbers and percentages. The sequence kernel association test (SKAT), which is a supervised, flexible, and computationally efficient regression method was used to test for the association between a set of genetic variants and a continuous or dichotomous trait with adjustments made for relevant covariates (Ionita-Laza et al., 2013). A total of 421 variants from the *ULK* family passed QC and were included in SKAT analysis with controlling for gender, hypertension, and smoking status and renal function (Wu et al., 2011). Furthermore, separate analyses were conducted for all variants (n = 421) and rare variants with minor allele frequency (MAF) < 0.05 (n = 290). We used COXPRESdb (http://coxpresdb.jp) to drawing the co-expressed gene network with pathway and protein–protein interaction information. COXPRESdb was first released for human and mouse models in 2007 (Takeshi et al., 2007). One characteristic feature of COXPRESdb is its ability to compare multiple co-expression data derived from different transcriptomics technologies and different species, which strongly reduces false-positive relationships in individual gene co-expression data (Takeshi et al., 2012; Takeshi et al., 2019). To clarify the molecular mechanisms, we used COXPRESdb to perform co-expression analysis on those genes that exerted significant effect on aortic dissection.

The analysis workflow is shown in Figure 1. In general, we adopted two analysis steps. First, we applied the SKAT to test the association between a set of SNP and patients' age of first onset of TBAD, controlling for gender, hypertension, smoking status, and renal function. Then, linear regression models were applied to further detect significant SNPs in significant genes. Multiple comparisons were adjusted with the false discovery rate method (FDR) to control the overall false-positive rate at a 5% level. Biomarkers measured by a q-FDR value ≤0.05 were included in further study (Benjamini and Hochberg, 1995).

To further explore the role of *ULK4* in different clinical features, we constructed a genetic score with the significant SNPs and conducted subgroup analysis to compare the differences between different categories of features. The genetic score (GS) was calculated based on a weighted linear combination of individual values of the significant SNPs, with weights derived from the stepwise linear regression model. As a result, three SNPs were finally included in the genetic score with the formula defined as follows: GS = (−25.326 × rs191792955) + (12.408 × rs142574024) + (−4.425 × rs74282513).

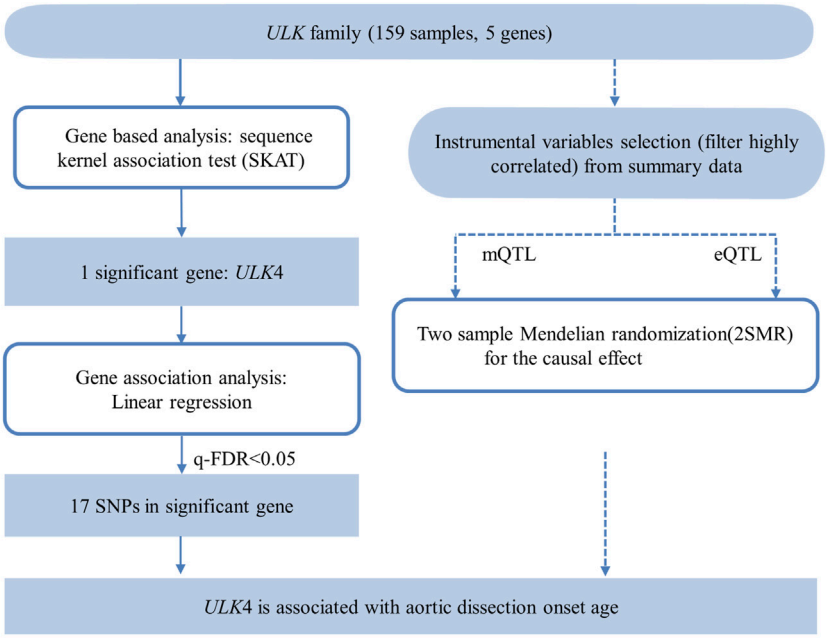
In the second step, we performed a 2-sample MR to explore the potential causal relationship between *ULK4* and the early onset of TBAD (Y) at the level of gene expression and DNA methylation (X) using genetic variants (G) as instrumental variables (Evans and Davey Smith, 2015). As shown in Figure 4, SNPs with a q-FDR < 0.05 were regarded as candidate instrumental variables and linkage disequilibrium (LD) clumping with a window of 1 MB, and an $r^2 < 0.2$ was applied to remove SNPs with high LD (Hemani et al., 2018). In addition, we performed sensitivity analyses using several approaches to investigate potential pleiotropic bias and verify the robustness of the results, including MR-Egger regression, weighted median MR, weighted mode MR, simple mode MR, funnel plots, and leave-1-variant-out analysis.

Statistical analyses were performed using R version 3.4.4 (The R Foundation of Statistical Computing, Vienna, Austria). A 2-sided p-value of less than 0.05 was considered to indicate a statistically significant difference.

Results

A total of 162 Chinese TBAD patients from Zhongshan Hospital were genotyped with two patients being excluded because of gender mismatches, and one patient being excluded because of familial relationships. Finally, 159 TBAD patients with 1,180,097 SNPs after QC were included in further analyses (Supplementary Figure S1). Demographics and clinical characteristics are shown in Supplemetary Table S1. The mean age of the enrolled TBAD inpatients was 56.11 years (SD 14.33), and 78.62% of the patients were male. The mean age of the first onset of TBAD was 54.89 years with a range of 18–85 years. Most of the enrolled TBAD inpatients (88.68%) also had hypertension, and 10.06% had diabetes.

Concerning on the association between the *ULK* gene family and age of first onset of TBAD, only *ULK4* was significant according to SKAT analysis (q-*FDR* = 0.0088) based on our TBAD samples. Among 361 variants in the *ULK4* gene, there



SNP indicates single nucleotide polymorphism; mQTL: methylation quantitative trait locus; eQTL: expression quantitative trait locus; q-FDR: False Discovery Rate q-values; ULK4: Unc-51 Like Kinase 4.

FIGURE 1
Flow chart of the study design and statistical analysis.

TABLE 1 Results of the ULK family regional gene-based SKAT analysis for aortic dissection.

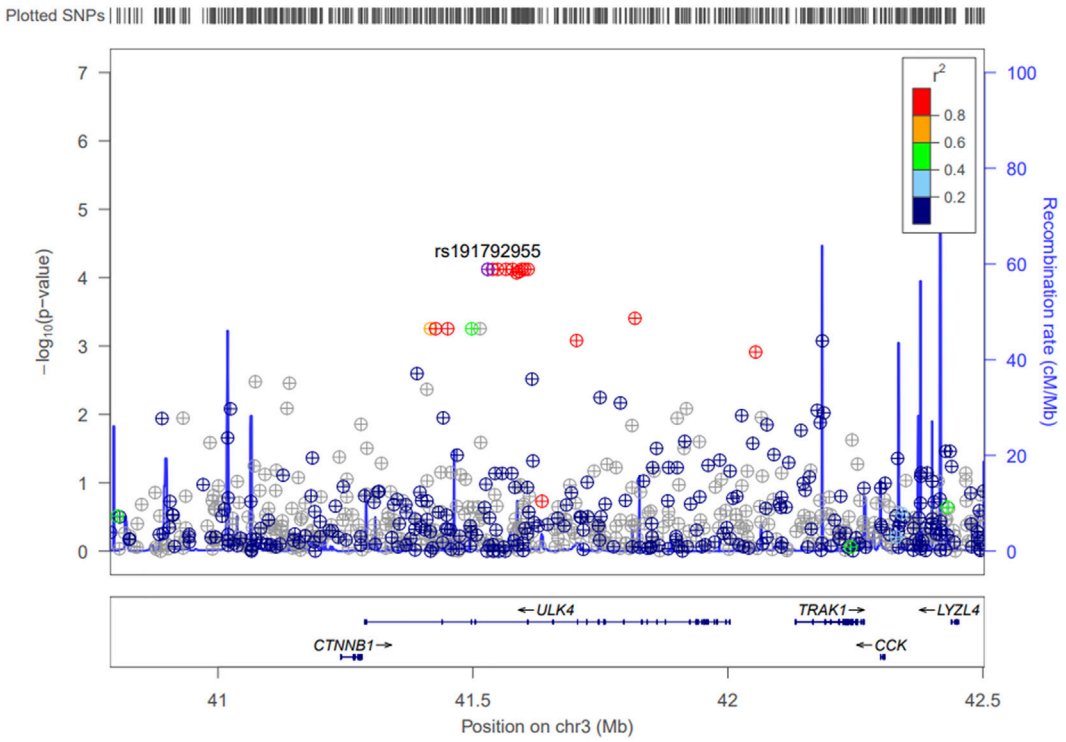
Candidate gene	All variants			MAF < 0.05		
	Number	*p*-value	q-FDR	Number	*p*-value	q-FDR
ULK1	18	0.7312	0.7312	13	0.6788	0.6788
ULK2	40	0.2331	0.5828	32	0.2127	0.5317
ULK3	2	0.6032	0.7312	2	0.6032	0.6788
ULK4	361	0.0018	0.0088	243	0.0015	0.0074
STK36	10	0.5855	0.7312	6	0.5156	0.6788

Models adjusted for sex, hypertension controlling, and renal function. Data source: TBAD cohort.

were 243 variants with MAF <0.05. In addition, SKAT analysis for variants with MAF <0.05 of *ULK4* was also significant (p = 0.0015) (Table 1). Further linear regression analysis of SNPs in *ULK4* revealed 17 SNPs, that were associated with age of first onset of TBAD, and the most strongly associated SNP in *ULK4* was rs191792955 (Figure 2; Supplementary Figure S2). The details of the association results of the SNPs sites of *ULK4* are shown in Supplementary Table S2.

Among the significant 17 SNPs, three SNPs were maintained for genetic score generation after stepwise regression with p < 0.05. As the genetic score increased, the age of first onset decreased for TBAD patients with well-controlled hypertension (coefficient 1.21, 95% CI: 0.677–1.751), and similar results were found for complex, male, and normal renal function TBAD patients. Heterogeneity details are shown in a forest plot (Supplementary Figure S3).

For 2-sample MR, the individual instrument-gene expression and instrument-age of first onset for TBAD are shown in Supplementary Table S3. Among the 19 SNPs associated with gene expression (q-FDR < 0.05), three were left after LD



Plots are produced in LocusZoom and show the most strongly associated SNP, rs191792955 (purple), intron of *ULK*4.

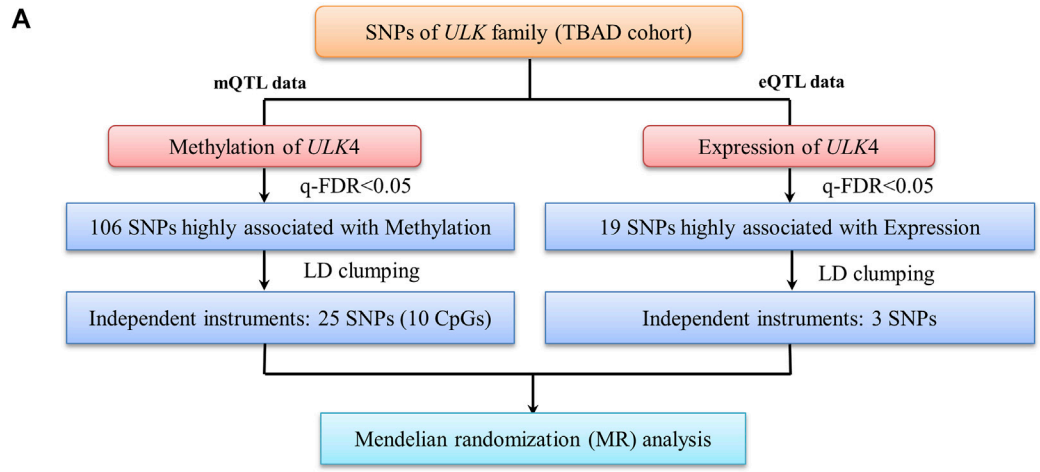
FIGURE 2
Regional association plots for ULK4.

clumping. A similar process was followed for the instrumental variable determination of *ULK4* methylation, and there were 10 CpG probes with 25 SNPs found (Supplementary Table S4). The details of the instrumental variable selections are shown in Figure 3A. The intercept of MR-Egger for *ULK4* gene expression and DNA methylation indicated that there was no potential horizontal pleiotropy, and the inverse variance weighted (IVW) analysis with fixed effects was conducted for two sample MR analyses. The high *ULK4* gene expression was related to a later age of onset for TBAD (β = 4.58, p = 0.0214) (Figures 3B, 4A). Among MR analyses of the 10 CpG probes, eight SNPs in cg25209153 were significant (p = 0.0155) (Figures 3B, 4A), and the higher cg25209153 was related to the earlier age of first onset (β = −4.02). In addition, several MR methods were used to evaluate the robustness of results, and the direct of the causal effects of both *ULK4* and cg25209153 were consistent with IVW, although the p values were not significant (Figures 4B, C). The details of sensitively analyses are shown in Supplementary Table S5.

According to the gene co-expression network of *ULK4* (Supplementary Figure S4), it was found that OSCP1 acts on smooth muscle cells in the tunica media layers of artery walls, participates in the regulation of the extracellular matrix (ECM) related to intimal proliferation after endothelial injury, and is related to restenosis after vascular injury (Martí-Pàmies et al., 2017). WDPCD is believed to be related to the development of arteries. In the process of coronary vasculature development, WDPCD participates in the regulation of epithelial–mesenchymal transition (EMT) to enable migration that gives rise to smooth muscle cells (Liu et al., 2018).

Discussion

Acute aortic dissection may be fatal without early diagnosis and appropriate management, and thus biomarker tests play an important role in preventing and diagnosing aortic dissection disease (Ranasinghe and Bonser, 2010). Several epigenetic studies of TBAD have identified potential biomarkers relevant to the etiology of TBAD (Wang et al., 2012; Erhart et al., 2020). Multiple GWAS studies have identified a significant association of the *ULK4* SNPs with hypertension. Genetic variants in *ULK4* have also been reported to be associated with the pathogenesis of sporadic thoracic aortic dissection (STAD) (Guo et al., 2016). To the best of our knowledge, this is the first



B

Exposure	Probe	nIV	SNPs	Method	Beta	SE	P
Expression	*ULK4*	3	rs1717022; rs7616123;rs60659132	Inverse variance weighted (fixed effects)	4.57	1.957	0.0196
DNA methylation	cg25209153	8	rs10510726; rs1256353; rs1495699; rs1716979; rs1716979; rs62257301; rs7629070; rs994439	Inverse variance weighted (fixed effects)	-5.43	2.642	0.0399

FIGURE 3
Two sample Mendelian randomization analyses. **(A)** Diagram of instrumental variable (IV) selection. **(B)** Results of Mendelian randomization (MR) between gene expression, DNA methylation, and onset age of TBAD.

study of the *ULK* gene family in TBAD to focus on the association with age of first onset.

The *ULK* (UNC51-like) enzymes play critical roles in autophagy and development. While *ULK1*, *ULK2*, and *ULK3* have been characterized, and very little is known about *ULK4*. Recently, deletions in *ULK4* have genetically linked to increased susceptibility to developing schizophrenia, which is a devastating neuropsychiatric disease with high heritability (Khamrui et al., 2019). Similarly, TBAD has also been identified as a suspected heritable characteristic. In their single-institute study, Shalhub et al. (2021) found that heritable TBAD was the cause of TBAD in one of four patients, and familial TBAD was presented at an early age. Finally, it has been established that hypertension is an essential component of both familial TBAD and sporadic TBAD (Shalhub et al., 2020).

Most patients in our TBAD cohort were male, most had hypertension, and few of them had diabetes. These population characteristics are consistent with the reported TBAD in the Chinese population (Huang et al., 2021). Based on whole genome sequencing data, we demonstrated that the *ULK4* gene is associated with the age of TBAD onset based on both the SKAT and 2-sample MR analyses. Furthermore, we found that DNA methylation of $cg25209153_{ULK4}$ and expression of *ULK4* were associated with the age of first onset. Furthermore, high DNA methylation of $cg25209153_{ULK4}$ was negatively correlated with age of first onset. Inversely, high expression of *ULK4* was positively correlated with age of first onset.

Several GWASs have reported significant associations of *ULK4* SNPs with hypertension in individuals of European (particularly those with high diastolic blood pressure), African American, and East Asian descent (Guo et al., 2016). It is also known that poorly controlled hypertension is a major risk factor for TBAD, and our study demonstrated that the *ULK4* gene, involved in endocytosis and axon growth (Ostberg et al., 2020), is associated with age of first onset of TBAD, suggesting that in addition to the association with the control of hypertension, genomic variants in *ULK4* have a potential mechanism for contributing to the early onset of TBAD. As autophagy is a highly conserved catabolic process and a major cellular pathway for the degradation of long-lived proteins and cytoplasmic organelles, *ULK4* plays critical roles in autophagy and dysregulation of autophagy may lead to the early onset of TBAD. Further studies are needed to validate the link between *ULK4* and the age of first onset of TBAD. *ULK4* may be an

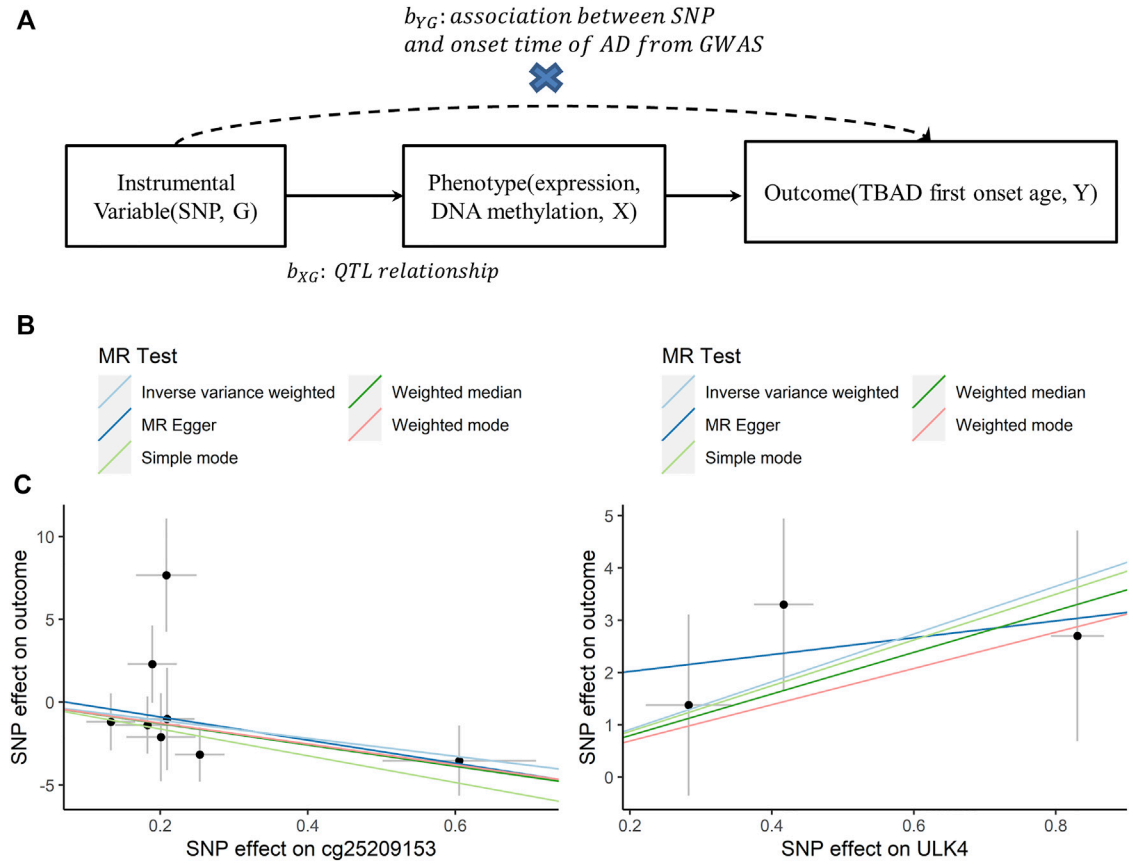


FIGURE 4
Association plots between ULK4 or cg25209153 and first onset age. **(A)** Diagram of Mendelian randomization (MR) analysis. **(B)** Scatter plots for association between cg25209153 and first onset age by different MR analytical methods. **(C)** Scatter plots for association between ULK4 and first onset age by different MR analytical methods.

effective companion diagnostic target in TBAD if it is confirmed by further fundamental and clinical studies.

In addition, our study will follow up the prognosis of the existing cases so as to suggest early detection and early treatments for at-risk patients. At the same time, we will also consider measuring different omics data such as proteomics from the same batch to further validate the importance of the ULK4 gene in TBAD.

Conclusion and limitations

Our study also has some limitations. First, the sample size was limited, which may impact the robustness of the 2-sample MR analyses. Second, because only TBAD patients were included in our study, the association of *ULK4* with type A dissections could not be investigated; moreover, as a case-only study, which did not involve the control group, the current study cannot conclude the effective companion diagnostic target in TBAD. Third, the validated DNA samples were not collected. In addition, the clinical information was not comprehensive, such as concomitant medication of antiplatelet and statin, the potential bias risk may impact the robustness of the results.

In conclusion, this is the first study of the *ULK* gene family in TBAD to focus on an analysis of the association with the first onset age. We demonstrated that the *ULK4* gene is associated with the age of first onset of TBAD based on both the SKAT and 2-sample MR analyses. *ULK4* may be an effective companion diagnostic target in TBAD.

Data availability statement

The data presented in the study are deposited through these following websites. The eQTL summary data can be found in http://www.psychENCODE.org/. The mQTL data from the meta-analyses of Brisbane Systems Genetics Study and the Lothian Birth Cohorts of 1921 and 1936 are available at http://cnsgenomics.com/software/smr/#Download. The whole genome sequencing data of TBAD cohort was provided by Yi Si, et al. Requests to access this dataset should be directed to sysiy@yahoo.com.

Ethics statement

The studies involving human participants were reviewed and approved by the relevant ethics committees of Zhongshan Hospital (Ethical approval No. B2019-110R) and was administered by trained personnel. The patients/participants provided their written informed consent to participate in this study.

Author contributions

LH, YS, and WF designed the research, wrote the manuscript, and responsible for final content. LH, YW, YS, and FC conceived and designed the research. LH, JT, LL, and RW conducted the research. LH, JT, and YW analyzed the data. FC and WF reviewed and edited the manuscript. All authors had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis and for the decision to submit the manuscript for publication.

Funding

This work was supported by the National Natural Science Foundation of China (Grant 81903407 to LH).

Acknowledgments

We are grateful to all patients who participated in this study.

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.956866/full#supplementary-material

References

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Biddinger, A., Rocklin, M., Coselli, J., and Milewicz, D. M. (1997). Familial thoracic aortic dilatations and dissections: A case control study. *J. Vasc. Surg.* 25 (3), 506–511. doi:10.1016/s0741-5214(97)70261-1

Chan, E. Y., and Tooze, S. A. (2009). Evolution of Atg1 function and regulation. *Autophagy* 5 (6), 758–765. doi:10.4161/auto.8709

Clément, M., Chappell, J., Raffort, J., Lareyre, F., Vandestienne, M., Taylor, A. L., et al. (2019). Vascular smooth muscle cell plasticity and autophagy in dissecting aortic aneurysms. *Arterioscler. Thromb. Vasc. Biol.* 39 (6), 1149–1159. doi:10.1161/ATVBAHA.118.311727

Erhart, P., Gieldon, L., Ante, M., Körfer, D., Strom, T., Grond-Ginsbach, C., et al. (2020). Acute stanford type B aortic dissection—Who benefits from genetic testing? *J. Thorac. Dis.* 12 (11), 6806–6812. doi:10.21037/jtd-20-2421

Evans, D. M., and Davey Smith, G. (2015). Mendelian randomization: New applications in the coming age of hypothesis-free causality. *Annu. Rev. Genomics Hum. Genet.* 16, 327–350. doi:10.1146/annurev-genom-090314-050016

Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., Chen, C., Liu, S., et al. (2018). Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* 362 (6420), eaat8127. doi:10.1126/science.aat8127

Guo, D-c., Grove, M. L., Prakash, S. K., Eriksson, P., Hostetler, E. M., LeMaire, S. A., et al. (2016). Genetic variants in LRP1 and ULK4 are associated with acute aortic dissections. *Am. J. Hum. Genet.* 99 (3), 762–769. doi:10.1016/j.ajhg.2016.06.034

Hagan, P. G., Nienaber, C. A., Isselbacher, E. M., Bruckman, D., Karavite, D. J., Russman, P. L., et al. (2000). The international registry of acute aortic dissection (IRAD): New insights into an old disease. *Jama* 283 (7), 897–903. doi:10.1001/jama.283.7.897

Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., et al. (2018). The MR-Base platform supports systematic causal inference across the human phenome. *elife* 7, e34408. doi:10.7554/eLife.34408

Huang, L., Kan, Y., Zhu, T., Chen, B., Xu, X., Dong, Z., et al. (2021). Ten-year clinical characteristics and early outcomes of type B aortic dissection patients with thoracic endovascular aortic repair. *Vasc. Endovasc. Surg.* 55 (4), 332–341. doi:10.1177/1538574420983652

Ionita-Laza, I., Lee, S., Makarov, V., Buxbaum, J. D., and Lin, X. (2013). Sequence kernel association tests for the combined effect of rare and common variants. *Am. J. Hum. Genet.* 92 (6), 841–853. doi:10.1016/j.ajhg.2013.04.015

Kamman, A. V., Brunkwall, V., Eric, L., Heijmen, R. H., and Trimarchi, S. (2017). Predictors of aortic growth in uncomplicated type B aortic dissection from the acute dissection stent grafting or best medical treatment (ADSORB) database. *J. Vasc. Surg.* 65 (4), 964–971. doi:10.1016/j.jvs.2016.09.033

Khamrui, S., Ung, P. M., Secor, C., Schlessinger, A., and Lazarus, M. B. (2019). High-resolution structure and inhibition of the schizophrenia-linked pseudokinase ULK4. *J. Am. Chem. Soc.* 142 (1), 33–37. doi:10.1021/jacs.9b10458

Levy, D., Ehret, G. B., Rice, K., Verwoert, G. C., Launer, L. J., Dehghan, A., et al. (2009). Genome-wide association study of blood pressure and hypertension. *Nat. Genet.* 41 (6), 677–687. doi:10.1038/ng.384

Liu, M., Xu, P., O'Brien, T., and Shen, S. (2017). Multiple roles of Ulk4 in neurogenesis and brain function. *Neurogenesis* 4 (1), e1313646. doi:10.1080/23262133.2017.1313646

Liu, X., Wang, Y., Liu, F., Zhang, M., Song, H., Zhou, B., et al. (2018). Wdpcp promotes epicardial EMT and epicardium-derived cell migration to facilitate coronary artery remodeling. *Sci. Signal.* 11 (519), eaah5770. doi:10.1126/scisignal.aah5770

Maloverjan, A., Piirsoo, M., Michelson, P., Kogerman, P., and Østerlund, T. (2010). Identification of a novel serine/threonine kinase ULK3 as a positive

regulator of Hedgehog pathway. *Exp. Cell Res.* 316 (4), 627–637. doi:10.1016/j.yexcr.2009.10.018

Martí-Pàmies, I., Cañes, L., Alonso, J., Rodríguez, C., and Martínez-González, J. (2017). The nuclear receptor NOR-1/NR4A3 regulates the multifunctional glycoprotein vitronectin in human vascular smooth muscle cells. *FASEB J.* 31 (10), 4588–4599. doi:10.1096/fj.201700136RR

Mokashi, S. A., and Svensson, L. G. (2019). Guidelines for the management of thoracic aortic disease in 2017. *Gen. Thorac. Cardiovasc. Surg.* 67 (1), 59–65. doi:10.1007/s11748-017-0831-8

Nienaber, C. A., and Clough, R. E. (2015). Management of acute aortic dissection. *Lancet* 385 (9970), 800–811. doi:10.1016/S0140-6736(14)61005-9

Ostberg, N. P., Zafar, M. A., Ziganshin, B. A., and Elefteriades, J. A. (2020). The genetics of thoracic aortic aneurysms and dissection: A clinical perspective. *Biomolecules* 10 (2), 182. doi:10.3390/biom10020182

Prêtre, R., and Von Segesser, L. K. (1997). Aortic dissection. *Lancet* 349 (9063), 1461–1464. doi:10.1016/S0140-6736(96)09372-5

Ranasinghe, A. M., and Bonser, R. S. (2010). Biomarkers in acute aortic dissection and other aortic syndromes. *J. Am. Coll. Cardiol.* 56 (19), 1535–1541. doi:10.1016/j.jacc.2010.01.076

Shalhub, S., Rah, J. Y., Campbell, R., Sweet, M. P., Quiroga, E., and Starnes, B. W. (2021). Characterization of syndromic, nonsyndromic familial, and sporadic type B aortic dissection. *J. Vasc. Surg.* 73 (6), 1906–1914.e2. doi:10.1016/j.jvs.2020.10.080

Shalhub, S., Roman, M. J., Eagle, K. A., LeMaire, S. A., Zhang, Q., Evangelista, A., et al. (2020). Type B aortic dissection in young individuals with confirmed and presumed heritable thoracic aortic disease. *Ann. Thorac. Surg.* 109 (2), 534–540. doi:10.1016/j.athoracsur.2019.07.004

Takeshi, O., Shinpei, H., Masayuki, S., Motoshi, S., Hiroyuki, O., and Kengo, K. (2007). COXPRESdb: A database of coexpressed gene networks in mammals. *Nucleic Acids Res.* 36, D77–D82. doi:10.1093/nar/gkm840

Takeshi, O., Yasunobu, O., Satoshi, I., and Shu, T. (2012). Motoike IN, kengo K: **COXPRESdb: A database of comparative gene coexpression networks of eleven species for mammals**. *Nucleic Acids Res.* 41 (D1), 1014–1020.

Takeshi, O., Yuki, K., Yuichi, A., Shu, T., and Kengo, K. (2019). COXPRESdb v7: A gene coexpression database for 11 animal species supported by 23 coexpression platforms for technical evaluation and evolutionary inference. *Nucleic Acids Res.* 47 (D1), D55-D62. doi:10.1093/nar/gky1155

Wang, D., Liu, S., Warrell, J., Won, H., Shi, X., Navarro, F. C., et al. (2018). Comprehensive functional genomic resource and integrative model for the human brain. *Science* 362 (6420), eaat8464. doi:10.1126/science.aat8464

Wang, L., Yao, L., Guo, D., Wang, C., Wan, B., Ji, G., et al. (2012). Gene expression profiling in acute Stanford type B aortic dissection. *Vasc. Endovasc. Surg.* 46 (4), 300–309. doi:10.1177/1538574412443315

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *Am. J. Hum. Genet.* 89 (1), 82–93. doi:10.1016/j.ajhg.2011.05.029

Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., et al. (2018). Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat. Commun.* 9 (1), 918–1014. doi:10.1038/s41467-018-03371-0

Yang, Y., and Klionsky, D. J. (2020). Autophagy and disease: Unanswered questions. *Cell Death Differ.* 27 (5), 858–871. doi:10.1038/s41418-019-0480-9

Young, A. R., Narita, M., Ferreira, M., Kirschner, K., Sadaie, M., Darot, J. F., et al. (2009). Autophagy mediates the mitotic senescence transition. *Genes Dev.* 23 (7), 798–803. doi:10.1101/gad.519709

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 06 September 2022
DOI 10.3389/fgene.2022.949110

OPEN ACCESS

EDITED BY
Rajkumar S. Kalra,
Okinawa Institute of Science and Technology Graduate University, Japan

REVIEWED BY
Zhiquan Li,
University of Copenhagen, Denmark
Aditya Yashwant Sarode,
Columbia University, United States

*CORRESPONDENCE
Huabing Zhang,
huabingzhang@ahmu.edu.cn
Wei Chen,
chenwei366@126.com,
chenwei366@ahmu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Genetics of Aging, a section of the journal Frontiers in Genetics

RECEIVED 20 May 2022
ACCEPTED 09 August 2022
PUBLISHED 06 September 2022

CITATION
Sun R, Wang X, Chen J, Teng D, Chan S, Tu X, Wang Z, Zuo X, Wei X, Lin L, Zhang Q, Zhang X, Tang K, Zhang H and Chen W (2022), Development and validation of a novel cellular senescence-related prognostic signature for predicting the survival and immune landscape in hepatocellular carcinoma. *Front. Genet.* 13:949110. doi: 10.3389/fgene.2022.949110

Development and validation of a novel cellular senescence-related prognostic signature for predicting the survival and immune landscape in hepatocellular carcinoma

Rui Sun[1†], Xu Wang[1†], Jiajie Chen[2†], Da Teng[3], Shixin Chan[1], Xucan Tu[1], Zhenglin Wang[1], Xiaomin Zuo[1], Xiang Wei[4], Li Lin[4], Qing Zhang[4], Xiaomin Zhang[4], Kechao Tang[4], Huabing Zhang[4,5]* and Wei Chen[1]*

[1]Department of General Surgery, First Affiliated Hospital of Anhui Medical University, Hefei, China, [2]Department of Dermatology, First Affiliated Hospital of Anhui Medical University, Hefei, China, [3]Department of Hepatopancreatobiliary Surgery, Affiliated Chuzhou Hospital of Anhui Medical University, First People's Hospital of Chuzhou, Chuzhou, China, [4]Department of Biochemistry and Molecular Biology, Metabolic Disease Research Center, School of Basic Medicine, Anhui Medical University, Hefei, China, [5]Affiliated Chuzhou Hospital of Anhui Medical University, First People's Hospital of Chuzhou, Chuzhou, China

Background: Cellular senescence is a typical irreversible form of life stagnation, and recent studies have suggested that long non-coding ribonucleic acids (lncRNA) regulate the occurrence and development of various tumors. In the present study, we attempted to construct a novel signature for predicting the survival of patients with hepatocellular carcinoma (HCC) and the associated immune landscape based on senescence-related (sr) lncRNAs.

Method: Expression profiles of srlncRNAs in 424 patients with HCC were retrieved from The Cancer Genome Atlas database. Lasso and Cox regression analyses were performed to identify differentially expressed lncRNAs related to senescence. The prediction efficiency of the signature was checked using a receiver operating characteristic (ROC) curve, Kaplan–Meier analysis, Cox regression analyses, nomogram, and calibration. The risk groups of the gene set enrichment analysis, immune analysis, and prediction of the half-maximal inhibitory concentration (IC50) were also analyzed. Quantitative real-time polymerase chain reaction (qPCR) was used

Abbreviations: DElncRNAs, differentially expressed lncRNAs; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GSEA, Gene set enrichment analysis; HCC, Hepatocellular carcinoma; IRGPI, immune-related gene prognostic index; IS, immune subtypes; KM, Kaplan-Meier; LASSO, least absolute shrinkage and selection operator; lncRNA, long non-coding ribonucleic acids; multi-Cox, multivariate Cox; IC50, Half-maximal inhibitory concentration; qPCR, quantitative real-time PCR; ROC, receiver operating characteristic; SRGs, senescence-related genes; srlncRNAs, senescence-related lncRNAs; TCGA, The Cancer Genome Atlas; uni-Cox, univariate Cox.

to confirm the levels of *AC026412.3*, *AL451069.3*, and *AL031985.3* in normal hepatic and HCC cell lines.

Results: We identified 3 srlncRNAs (*AC026412.3*, *AL451069.3*, and *AL031985.3*) and constructed a new risk model. The results of the ROC curve and Kaplan–Meier analysis suggested that it was concordant with the prediction. Furthermore, a nomogram model was constructed to accurately predict patient prognosis. The risk score also correlated with immune cell infiltration status, immune checkpoint expression, and chemosensitivity. The results of qPCR revealed that AC026412.3 and AL451069.3 were significantly upregulated in hepatoma cell lines.

Conclusion: The novel srlncRNA (*AC026412.3*, *AL451069.3*, and *AL031985.3*) signatures may provide insights into new therapies and prognosis predictions for patients with HCC.

KEYWORDS

hepatocellular carcinoma, lncRNA, senescence, prognosis, immune landscape

1 Introduction

Cancer is one of the leading causes of death worldwide (Bray et al., 2021). More than 19 million new cancer cases and nearly 10 million cancer-related deaths have been reported in 2020, including over 900,000 new liver cancer cases and 800,000 related deaths (Sung et al., 2021). Liver cancer has the seventh highest incidence among all cancer types and the third highest mortality rate. Hepatocellular carcinoma (HCC) is the most common type of liver cancer. East Asia and Africa have the highest incidence rates of HCC, and its incidence and mortality rates are still increasing in Europe and other parts of the world (Llovet et al., 2021; Sung et al., 2021). Owing to the progress of surgery and chemotherapy, the prognosis of patients with HCC has greatly improved, and the progress of tumor immunotherapy and the use of immune checkpoint inhibitors have also improved the treatment strategies for HCC treatment (Bagchi et al., 2021). However, more efficient molecular biomarkers for the early diagnosis of HCC are crucial for improving the clinical outcomes of patients with HCC.

Cellular senescence is a typical irreversible form of life stagnation that helps inactivate and eliminate diseased, dysfunctional, and other unnecessary cells. It is usually induced by various conditions, such as microenvironmental stress, damage to organelles and cellular infrastructure, and an imbalance of cellular signal networks. However, all these conditions are related to the increase in senescent cell abundance in various organs observed during the aging process. It is considered to be one of the basic hallmarks of cancer (Hanahan, 2022).

Long non-coding ribonucleic acids (lncRNAs) are composed of >200 nucleotides that cannot be translated into functional proteins (Iyer et al., 2015). In the human genome, there are more than 100,000 identified lncRNAs, many of which have been characterized (Heery et al., 2017). lncRNAs are usually the main regulators of gene expressions and functions through post-transcriptional, transcriptional, and epigenetic regulation (Castro-Oropeza et al., 2018). Previous studies have shown that lncRNAs can influence the immune microenvironment; therefore, they may have a role in the occurrence and development of malignancy (Atianand et al., 2017). The HOX transcript antisense RNA was found to be upregulated in colon tumor tissues and correlated with the tumor stage, invasion, metastasis, and survival time of patients (Luo et al., 2016; Tatangelo et al., 2018; Wei et al., 2020); it is also associated with cancer growth and metastasis (Wei et al., 2020). Zhao et al. (2019) reported that the knockdown of lncRNA myocardial infarction–associated transcript significantly promoted cellular senescence and inhibited HCC progression. Montes et al. (2021) identified MIR31HG as a potential therapeutic target in the treatment of senescence-related pathologies. The effects of senescence-related (sr) lncRNAs on malignant tumors have not been well studied; therefore, obtaining more knowledge on srlncRNAs will help us better understand their roles in cancer therapy.

In recent years, many studies have developed signatures for predicting cancer prognosis based on coding genes or non-coding RNAs. Kandimalla et al. (2020) identified a signature for predicting the survival in pancreatic ductal adenocarcinoma. Chen et al. (2021) constructed a prognosis index for head and neck tumors using immune-related genes. Zhou et al. (2021) identified an immune-related lncRNA signature to predict the survival and the immune landscape in patients with HCC. However, only a few studies have focused on signature development using srlncRNAs.

This study aimed to determine the value of srlncRNAs in predicting the prognosis and immune landscape of HCC, thus contributing to this growing area of research. Our findings may

help improve our understanding of the role of cellular senescence in HCC and lead to progress in treatment strategies.

2 Materials and methods

2.1 Data collection

RNA-seq expression data derived from patients with HCC, including 374 tumors and 50 non-cancerous samples, were collected from The Cancer Genome Atlas (TCGA) database with the TCGA-Assembler. Based on the patients' IDs, the clinical data of the patients were compared to their transcriptome data, which were screened using the following inclusion criteria: [1] histological diagnosis of HCC, [2] available expression profiles, and [3] a minimum overall survival of 30 days (Song et al., 2021). The data satisfying the inclusion criteria were extracted from the TCGA dataset (344 patients) for subsequent analysis, and 279 senescence-related genes (explained in Supplementary Table S1) were retrieved from the literature search and the CellAge public database.

2.2 Identification of senescence-related long non-coding ribonucleic acids

The association between lncRNAs and senescence-related genes (SRGs) was assessed using Pearson's correlations to identify srlncRNAs. Using the Bioconductor limma package in R software (version 4.1.3), HCC and non-neoplastic samples were compared, and differentially expressed lncRNAs (DElncRNAs) were defined with the following criteria: |log2 (fold change, FC)| >1 and false discovery rate < 0.05 (Ritchie et al., 2015). A total of 279 senescence-related genes and those of DElncRNAs were identified by using the correlation analysis. Hence, 422 srlncRNAs were selected based on the following criteria: Pearson's correlation coefficients > 0.5 and $p < 0.001$.

2.3 Construction of the senescence-related lncRNA prognostic model

First, we randomly divided the patients from the entire sample ($n = 342$) into training or testing sets at a rate of 1:1. Second, srlncRNAs (related to survival) in the training set were screened using univariate Cox (uni-Cox) regression ($p < 0.05$). Third, least absolute shrinkage and selection operator (LASSO) and multivariate Cox (multi-Cox) regression analyses were used for further filtering. Finally, a prognostic model for srlncRNAs was established in HCC. We calculated the risk score for HCC as follows: risk score=$\sum$n k=1expression (lncRNAk) × coefficient (lncRNAk) (Hong et al., 2020). Using the median value, we divided the cases into two groups: high and low. Moreover, testing sets were employed for signature validation. The signature was associated with clinical variables using the chi-square test. The Wilcoxon signed-rank test was performed to identify differences in the risk scores between the groups for clinical characteristics. Furthermore, the R package "rms" was used to build a nomogram model that connected the signature risk score and clinical factors, and calibration curves were used to assess the model (Iasonos et al., 2008).

2.4 Gene set enrichment analysis

Using the curated gene set (kegg.v7.4.symbols.gmt), broad GSEA v.4.2.3 was applied to detect high- and low-risk group-correlation pathways with the criteria: NOM $p <$ 0.05 and | NES |> 1 (Subramanian et al., 2005).

2.5 Infiltrating immune cell analysis

The immune infiltration statuses calculated in the datasets (XCELL, TIMER, QUANTISEQ, MCPcounter, EPIC, CIBERSORT, and CIBERSORT) and the infiltration estimation downloaded in TIMER2.0 (http://timer.cistrome.org/) were used to analyze the differences in immune infiltrating cell content using the Wilcoxon signed-rank test. Using the profile of infiltration estimation for HCC on that website, a bubble chart was created showing the differences in immune infiltrating cell content using the Wilcoxon signed-rank test and the following R packages—"limma", "scales", "ggplot2", and "ggtext" (Bagchi et al., 2021).

2.6 The investigation of the immune checkpoints and immune-related gene prognostic index

The "ggpubr" R package was used to compare the expression of immune checkpoint-related genes between the two groups. The multi-Cox regression analysis was used to construct an IRGPI model to validate the impact of the prognostic model on immunotherapy.

2.7 The sensitivity of different subgroups to chemotherapeutic agents

We used the half-maximal inhibitory concentration (IC50) to evaluate the therapeutic effects of common chemotherapeutic drugs (paclitaxel, doxorubicin, bexarotene, bicalutamide, imatinib, and tipifarnib) using the R package "pRRophetic" with data collected from the Genomics of Drug Sensitivity in Cancer.

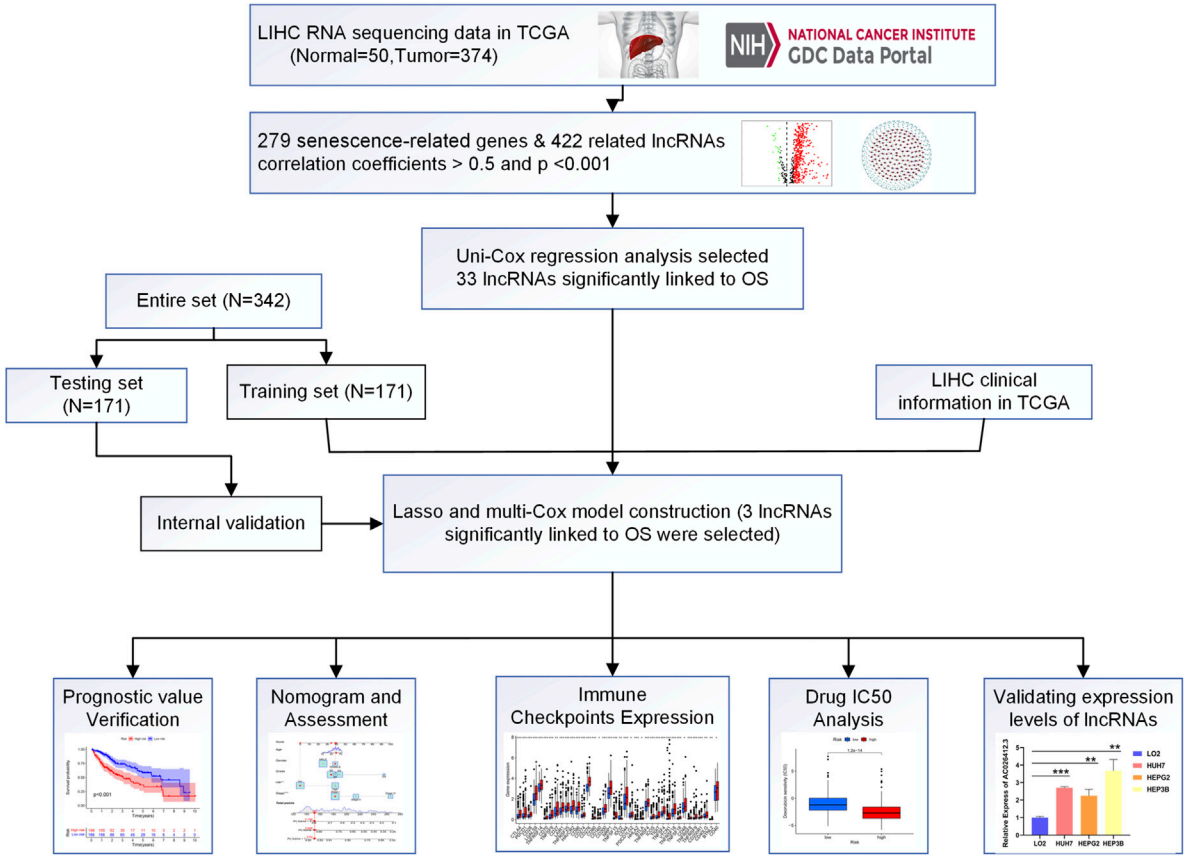


FIGURE 1
Flow diagram of the study (LIHC: Liver hepatocellular carcinoma; N: Number; OS: Overall suvival).

2.8 RNA isolation and quantitative real-time PCR

Total RNA was extracted from hepatoma cell lines (Huh7, HepG2, and Hep3B) and a normal hepatic cell line (LO2) using TRIzol reagent (Life, United States). NanoDrop 2000 (Thermo Scientific, America) was used to measure RNA purity and content. Complementary DNAs were synthesized using a RevertAid RT kit (Thermo, United States), and qPCR was performed on a Bio-Rad CFX system using qPCR Master mix (Universal, China). The sequences of the primers used for qPCR were as follows: *AC026412.3*, forward: 5′-TGTGAGGTGAGG GAGCGAT-3′, reverse: 5′-TGAGCCAAAGGGATCTACGC-3′; *AL451069.3*, forward: 5′-GGGACACGGACCTAGACACT-3′, reverse: 5′-CCTGCAAGACCGTAGCCTC-3′; *ALO31985.3*, forward: 5′-TCTCACTATGTTGCTGGACTGG-3′, reverse: 5′-CCACAGATCACTAACACGCC-3′. We used glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as the internal reference, and the data were analyzed using the $2^{-\Delta\Delta Ct}$ approach. The expression levels of the three lncRNAs were compared using an unpaired t-test.

3 Results

3.1 Defining senescence-related lncRNAs

The flow-diagram of our study is shown in Figure 1. We downloaded 50 normal samples and 374 tumor samples from the TCGA database to identify the srlnRNAs. Next, 422 srlncRNAs (Figure 2A) were obtained by using the co-expression analysis of 279 senescence-related genes and DElncRNAs between normal and tumor samples. Of these, 402 were up-regulated (Figures 2B, C).

3.2 Establishment and validation of the model

Using the univariate-Cox regression analysis (Figure 3A), 33 srlncRNAs that significantly associated with the overall survival were identified and are displayed in a heatmap (Figure 3B). LASSO and multi-Cox regression analyses were used to further screen these lncRNAs, and three lncRNAs related

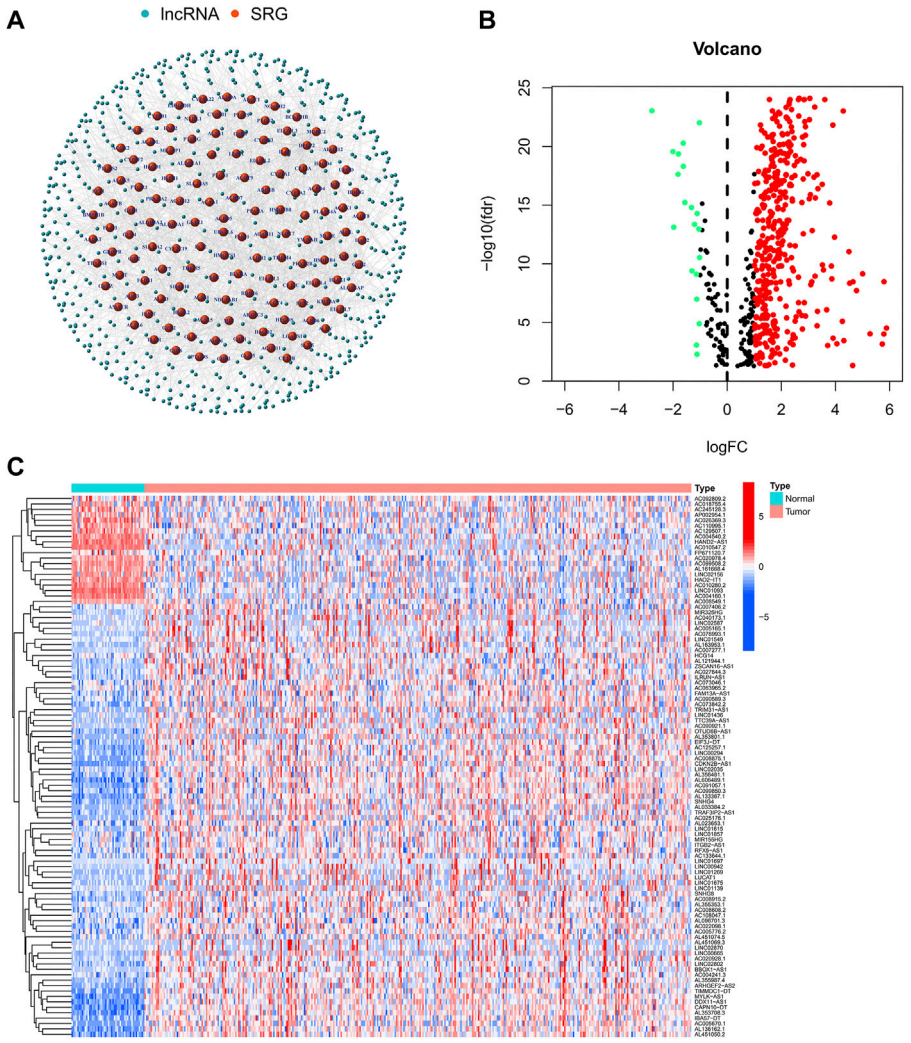


FIGURE 2
(A) The network showing the correlation between DEsrlncRNAs and mRNAs and **(B,C)** the heatmap and volcano plots of DEsrlncRNAs from the TCGA dataset.

to senescence were extracted in HCC (Figures 3C, D). In addition, all lncRNAs were up-regulated in the Sankey diagram (Figure 3E).

The risk score was calculated using the following formula: risk score = *AC026412.3* × (1.6474) + *AL451069.3* × (0.6620) + *AL031985.3* × (1.0340).

We then compared the distribution of the risk scores, survival status, survival time, and associated expression criteria of these lncRNAs for the low- and high-risk groups in the training, testing, and entire sets. These results suggested that the high-risk group had a poorer prognosis (Figures 4A–L).

According to chi-square tests (Figure 5A) and Wilcoxon signed-rank test, the risk score was significantly associated with the clinical grade (Figure 5B), American Joint Committee on Cancer stage (Figure 5C), and T stage (Figure 5D). In addition, conventional clinicopathological characteristics, including age, sex, and stage, also showed that the high-risk group had worse prognoses (Figures 5E–J). These results indicate that the risk model is highly consistent with the American Joint Committee on Cancer staging system and has a better ability to predict prognosis.

Prognostic factors were detected in the uni- and multi-Cox regression analyses (Figures 6A,B) and a nomogram was constructed using the risk scores and other clinical characteristics to better predict the survival of patients with HCC (Figure 6C). What's more, the nomogram correlated with the actual observations, as shown in the calibration curve (Figure 6D). The 1-, 3-, and 5-year areas under the ROC curve of the entire set were 0.754, 0.675, and 0.670, respectively (Figure 6E). Compared to other clinicopathological features,

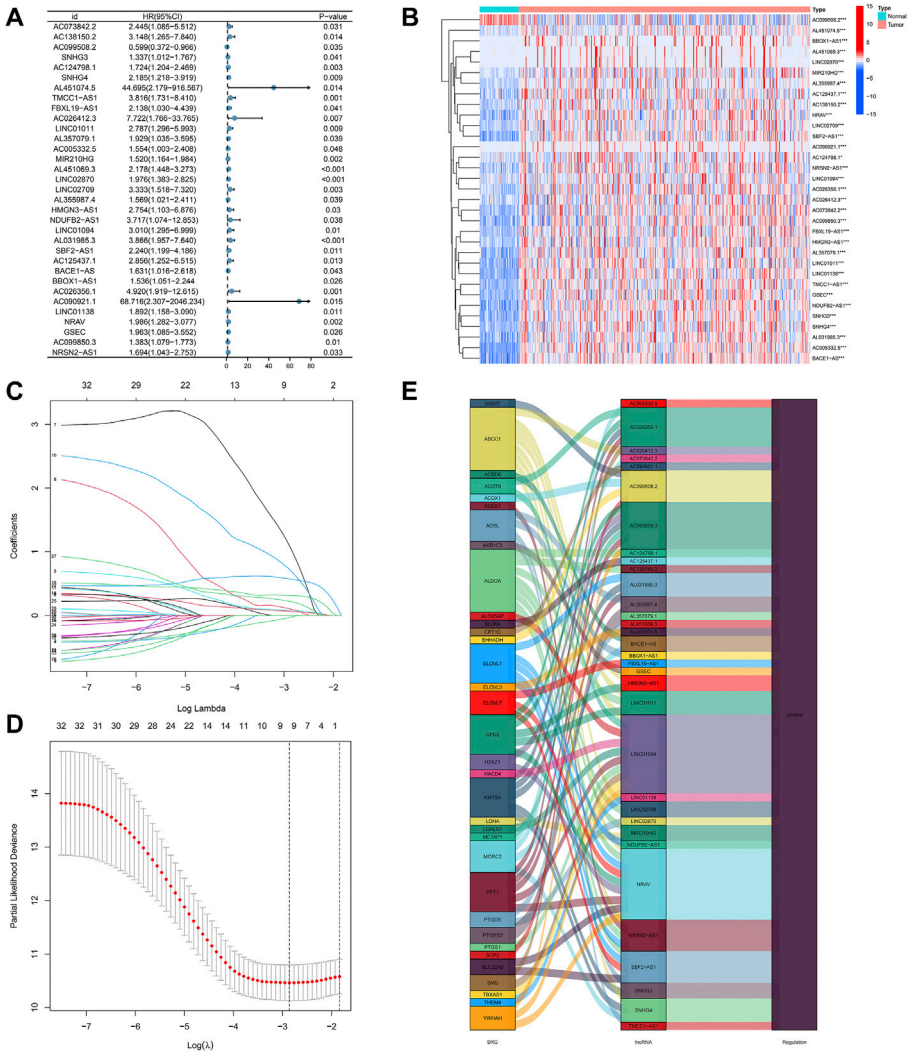


FIGURE 3
(A) 33 lncRNAs extracted by using the uni-Cox regression analysis, **(B)** the heat map of 33 prognostic lncRNAs, **(C,D)** senescence-related lncRNAs screened by the Lasso regression analysis, and **(E)** the Sankey diagram of 33 senescence genes and related lncRNAs.

the risk score had the largest area under the ROC curve (Figure 6F).

3.2.1 Gene set enrichment analysis

To explore the different biological functions in the two risk groups, the GSEA software was used to identify the top five pathways in the two risk groups with the criteria of false discovery rate < 0.25, |NES| >1.5, and $p < 0.05$. In fact, most of the pathways were associated with tumorigenesis or immunity, such as the "fatty acid metabolism", "peroxisome proliferator-activated receptors signaling pathway", and "complement and coagulation cascades" (Figure 7A). Therefore, we performed an immunity analysis of the model.

3.3 The exploration of the risk model for immunotherapy

Using Spearman's correlation and Wilcoxon signed-rank tests, the risk score was found to be related to several widely studied immune cells (such as B cells, $CD8^+$ T cells, and cancer-associated fibroblasts) on different platforms (Figures 7B,C). The expression of the immune checkpoint-related genes was higher in the high-risk group than in the low-risk group (Figure 8A). This implies that patients in the high-risk group could select checkpoint inhibitors that are more appropriate for immunotherapy (Kono et al., 2020). Moreover, the high-risk group had a larger proportion of immune subtypes (IS) 1 and 2 in the immunity landscape and a smaller proportion of 3

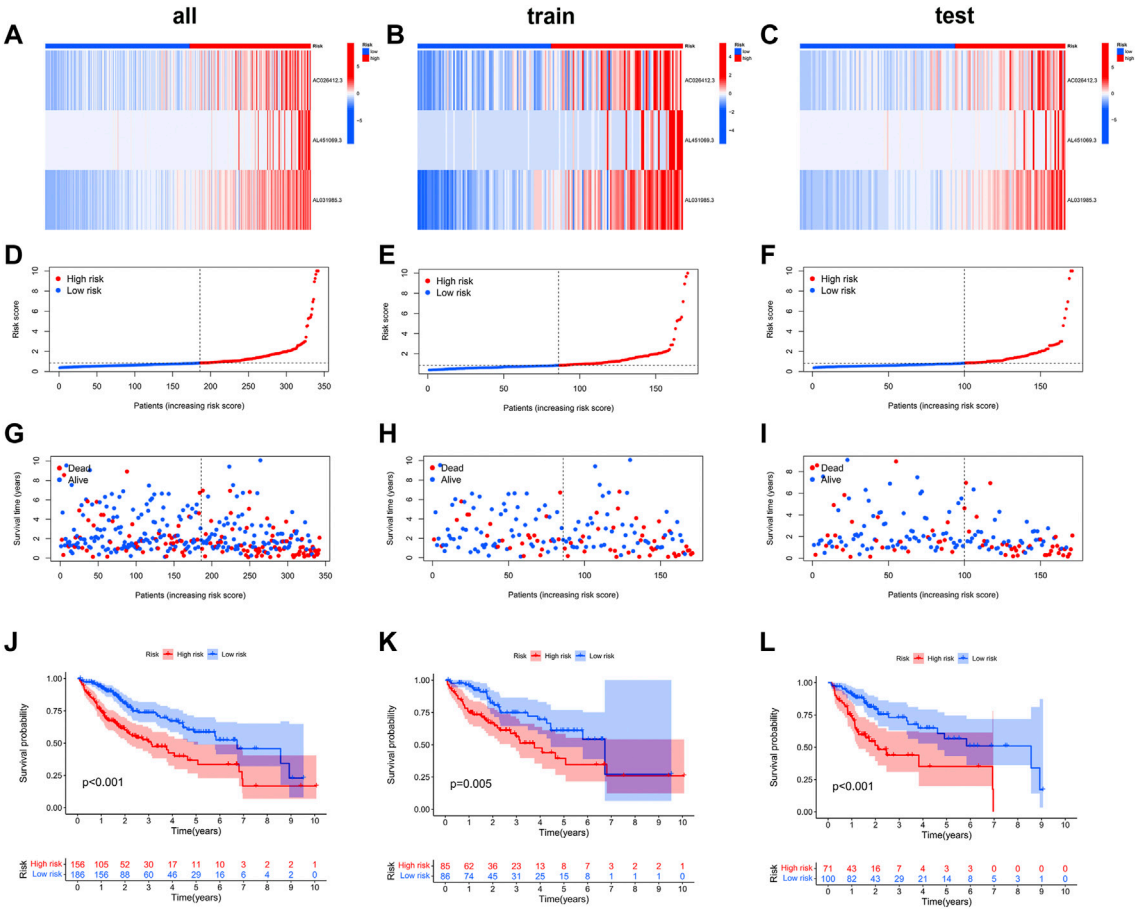


FIGURE 4
The heat map **(A–C)**, risk score **(D–L)**, survival status **(G–I),** and Kaplan–Meier curves **(J–L)** of the two groups in the training, testing, and entire sets, respectively.

(Figure 8B), which means that it had a poorer prognosis (the immune landscape of cancer). Consistent with previous reports, there were more chemotherapeutics with lower IC50 values in the high-risk group (Figure 8C), such as paclitaxel (Liu et al., 2020) (Figure 8C).

3.4 Validating the expression levels of *AC026412.3*, *AL451069.3*, and *AL031985.3*

To explore the expression levels of *AC026412.3*, *AL451069.3*, and *AL031985.3*, qPCR was performed to test the normal hepatic and hepatoma cell lines. The expression levels of *AL451069.3* and *AC026412.3* in hepatoma cell lines were much higher than those in a normal hepatic cell line (Figure 9). Furthermore, the expression levels of these srlncRNAs were different in diverse hepatoma cell lines (Supplementary Figure S1).

4 Discussion

4.1 Resource identification initiative

Cellular senescence has been found to play a role in the development and progression of various types of malignant tumors, including HCC, and is considered a barrier to the progression from a chronic liver disease to HCC. Xiang et al. (2021) reported that the lncRNA PINT87aa was upregulated in senescent HCC cells and could induce cell cycle arrest by blocking FOXM1-mediated PHB2. Mittermeier et al. (2020) described the characteristics and functions of cellular senescence in the development of novel drug targets for HCC therapies. Karakousis et al. (2020) suggested that hepatitis B is a link between cellular senescence and HCC development. A better understanding of the role of cellular senescence in HCC may provide a new perspective for HCC treatment and aid in the development of new therapeutic methods.

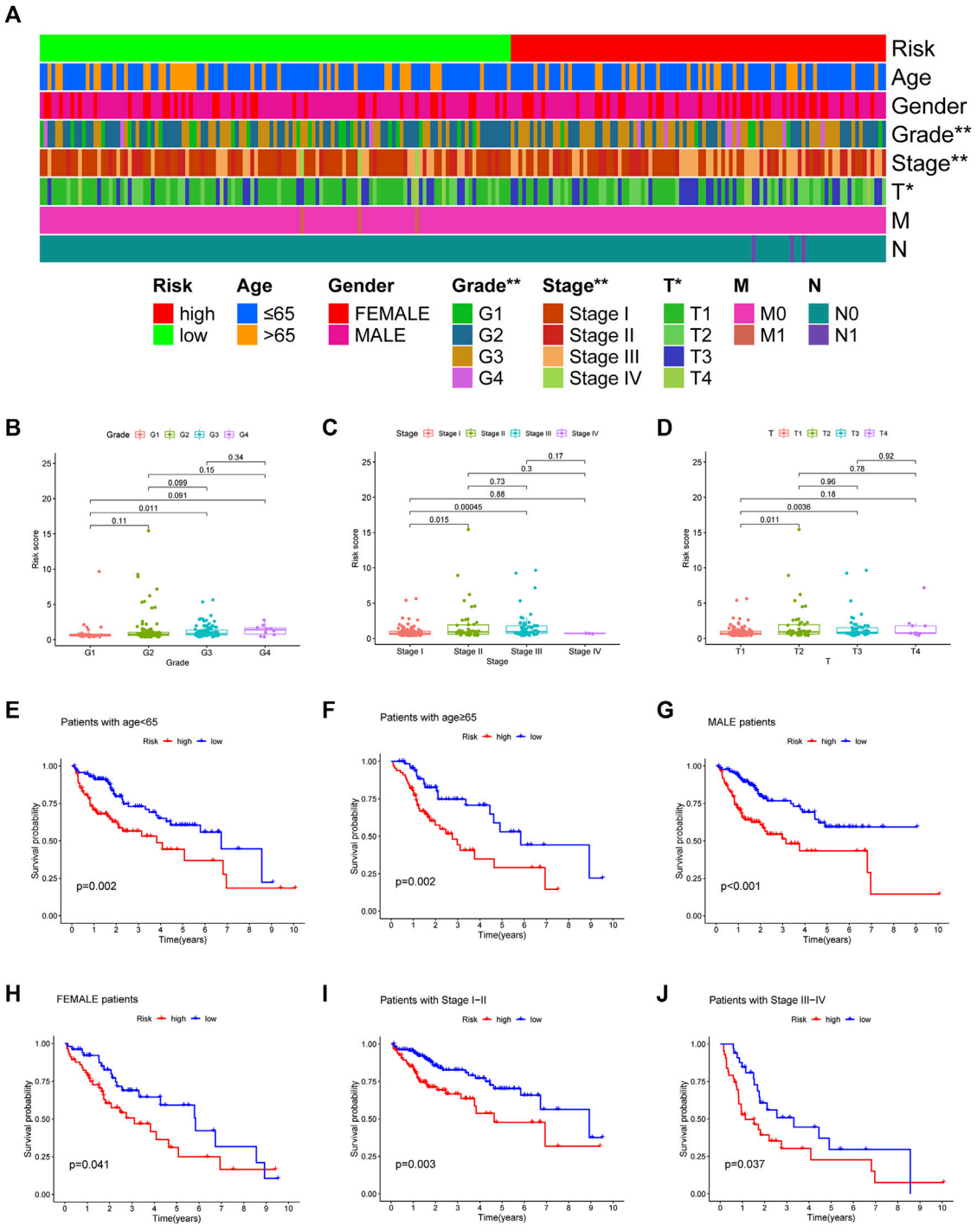


FIGURE 5
The strip chart **(A)** and Scatter diagram **(B–D)** showing significant correlation of the tumor grade, American Joint Committee on Cancer stage, and T stage with risk scores. *$p < 0.05$ and **$p < 0.01$. **(E–J)** The Kaplan–Meier analysis showing a longer survival time in low-risk group patients.

The expression patterns and clinical information of 377 patients with HCC were downloaded from the TCGA database, senescence-related genes were identified from the CellAge public database, and a co-expression analysis was performed to identify the genes potentially involved in HCC. Three prognosis-related DEsrlncRNAs were screened to construct a signature using

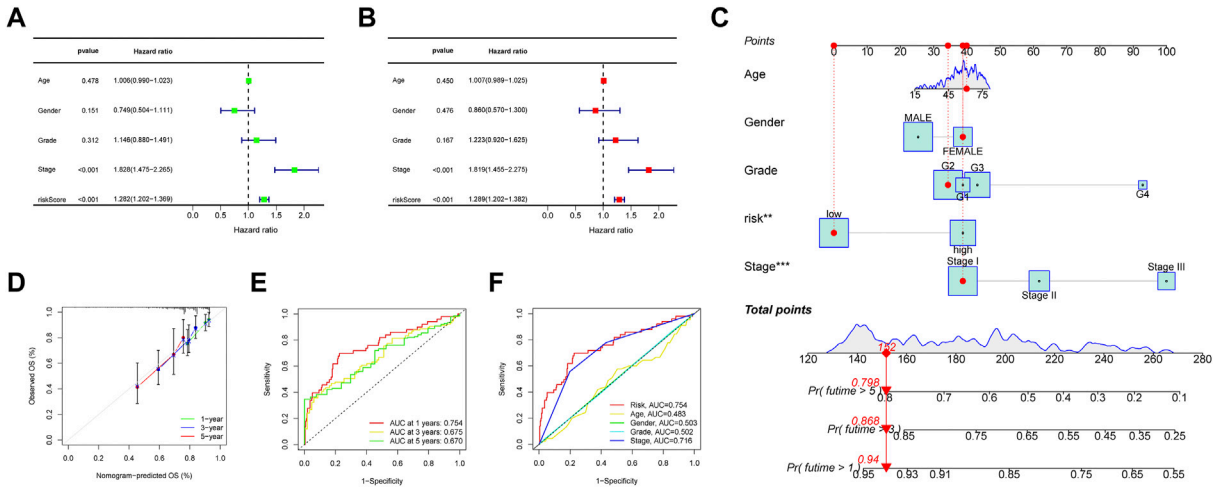


FIGURE 6
(A, B) Forest plots of the uni- and multi-Cox regression analyses in HCC, **(C)** the nomogram-combined risk score, age, and tumor stage to predict the 1-, 3-, and 5-year OS in HCC and **(D)** evaluation of the nomogram by correlating it with the calibration curves. **(E)** The ROC curves of the model for prognosis, and **(F)** the ROC curves of the risk score and clinicopathologic features.

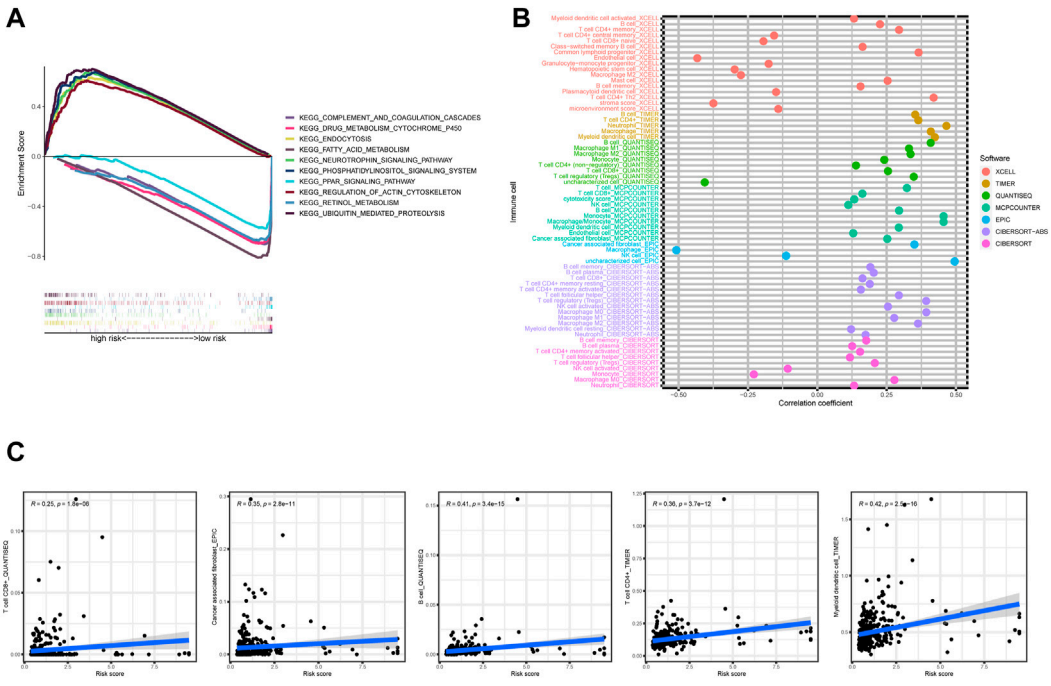


FIGURE 7
(A) The top five pathways with enrichment in the high- and low-risk groups with the GSEA analysis, **(B)** the bubble chart showing risk groups and immune cells, and **(C)** line graph demonstrating risk score and immune cells.

LASSO and uni-Cox regression analyses: *AC026412.3*, *AL451069.3*, and *AL031985.3*. Among the three srlncRNAs, *AL031985.3* has been identified as a potential therapeutic target in HCC in a previous study (Jia et al., 2020). Moreover, the Sankey diagram showed that the three srlncRNAs were associated with a few coding genes, including *PPT1*, *PTGDS*, and *ELOVL1*. High *PPT1* expression is

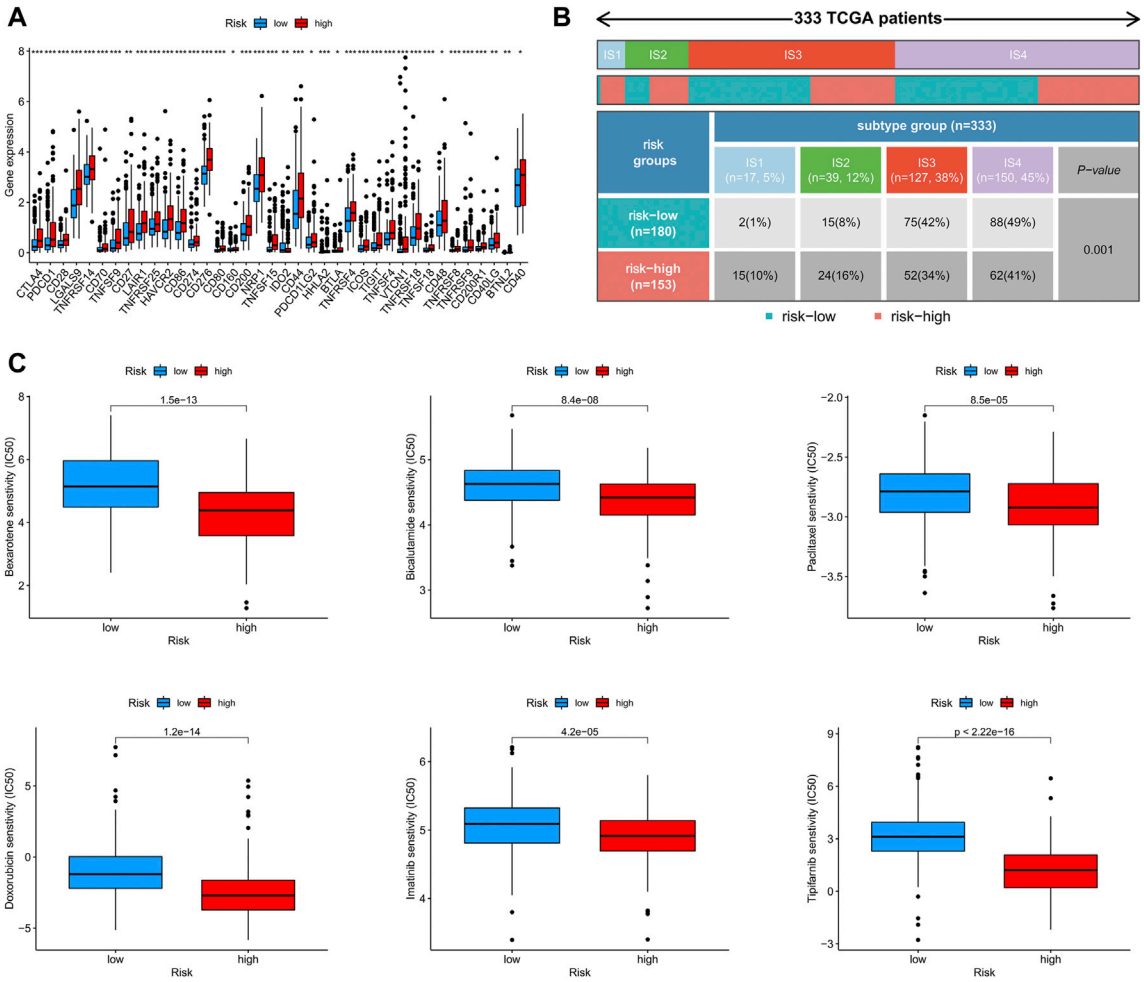


FIGURE 8
(A) The difference of 38 checkpoint expressions in the risk groups, **(B)** the immune subtype of high- and low-risk groups, and **(C)** the immunotherapy prediction of the risk groups.

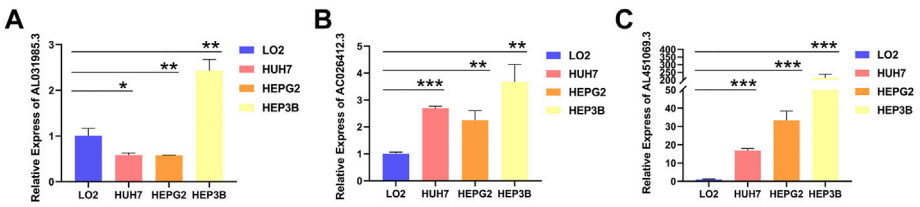


FIGURE 9
The relative RNA level of AL031985.3 **(A)**, AC026412.3 **(B)**, and AL451069.3 **(C)** in normal hepatic and hepatoma cell lines. Data is presented as Mean with SD, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

associated with poor prognosis in patients with HCC, and PPT1 inhibition could enhance the sensitivity to sorafenib therapy in HCC (Xu et al., 2022). *PTGDs* are prognostic biomarkers of breast cancer (Adekeye et al., 2022). Hama et al. (2021) demonstrated that the expression of *ELOVL1* was significantly higher in CRC tissues than in normal tissues. These

results suggest that the three identified srlncRNAs may serve as potential biomarkers for cancer diagnosis and treatment.

The risk score was calculated based on the expression levels of the three srlncRNAs, and patients in each cohort were separated into high- and low-risk groups according to the calculated risk score. The Kaplan–Meier curve showed that patients with a low risk score had a better prognosis. Based on the results of the uni- and multi-Cox regression analyses, the risk score could be an independent prognostic factor for patients with HCC. In addition, nomograms are widely used as tools in oncology, particularly for survival prediction (Iasonos et al., 2008; Balachandran et al., 2015). The nomogram model and calibration plot showed good prediction efficiency for HCC prognosis. Moreover, the correlation between risk scores and clinical features of HCC was also analyzed; the risk score was significantly related to the tumor grade, AJCC stage, and T stage, indicating that the risk score can be used for predicting the occurence and development of HCC. However, the results of the Wilcoxon signed-rank test showed that the advancing stages (G4, stage IV, and T4) were not significantly related to the calculated risk score. Because the sample content of the TCGA database is too small, we will have to collect more samples to re-validate.

Based on the results of GSEA, we focused our attention on the immunity factors. Previous research has suggested that tumor-infiltrating $CD4^+$ T cells can upregulate the immune checkpoint genes (Toor et al., 2019). We used TIMER2.0 to assess the relationship between the risk score and tumor-infiltrating immune cells (Van Veldhoven et al., 2011; Newman et al., 2015; Becht et al., 2016; Aran et al., 2017; Li et al., 2017; Finotello et al., 2019; Tamminga et al., 2020). The results revealed that the risk score was positively related to B cells, $CD8^+$ T cells, and cancer-associated fibroblasts. To further explore the potential of checkpoint blockade therapy and chemotherapy, we compared the two groups' expression levels of the immune checkpoint genes and found 38 checkpoint genes that were differentially expressed between the two groups in this study. Consistent with the alteration of the checkpoint genes, the IC50 values of six common chemotherapeutics were higher in the low-risk group. These findings suggest that patients with high-risk scores may be more suitable for immunotherapy and chemotherapy.

Thorsson et al. (2018) identified the ISs, including wound healing (IS1), IFN-γ dominant (IS2), inflammatory (IS3), lymphocyte-depleted (IS4), immunologically quiet (IS5), and TGF-β dominant (IS6) types of cancer. It was observed that IS1 and IS2 had worse outcomes, IS3 had a favorable prognosis, and IS3 was enriched in PBRM1 mutation. Moreover, patients with PBRM1 mutations were more responsive to immunotherapy (Miao et al., 2018). Our study indicated that patients with low-risk scores had a larger proportion of IS3, which comports with the Kaplan–Meier curve.

However, our study has few limitations. First, our analysis was based on public datasets and retrospectively collected samples, which may have caused an inherent case selection bias. Second, further experiments are required to confirm our findings. Finally, clinical features related to surgery, neoadjuvant chemotherapy, and tumor markers were not included in our study, and clinical cases are required to further validate our conclusions.

In conclusion, the cellular senescence-based prognostic signature constructed in this study may be useful for predicting the survival and guiding clinical therapies for HCC. Our findings may improve the understanding of cellular senescence in HCC and provide more effective treatment strategies. However, additional experiments and clinical cases are required to validate these findings.

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

RS, XuW, and JC are responsible for the writing and submission of the manuscript. DT, SC, ZW, and XZ contributed to the data analysis and visualization. XiW, LL, QZ, XZ, and KT reviewed and edited the manuscript. HZ and WC designed this study. All authors have read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (Nos.81670517 and 81870402) and the Research Fund of the Anhui Institute of Translational Medicine (2021zhyx-C30).

Acknowledgments

We thank the TCGA database providers for providing their platform and contributors for uploading their meaningful datasets.

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.949110/full#supplementary-material

SUPPLEMENTARY TABLE S1
279 senescence-related genes.

SUPPLEMENTARY TABLE S2
422 senescence-related lncRNAs.

SUPPLEMENTARY TABLE S3
33 lncRNAs extracted by univariate Cox regression analysis.

SUPPLEMENTARY TABLE S4
3 lncRNAs used to construct prognostic model.

SUPPLEMENTARY TABLE S5
Raw data of qRT-PCR.

SUPPLEMENTARY TABLE S6
Data and scripts used in the study are uploaded to Jianguoyun, and the links are included.

References

Adekeye, A., Agarwal, D., Nayak, A., and Tchou, J. (2022). PTGES3 is a putative prognostic marker in breast cancer. *J. Surg. Res.* 271, 154–162. doi:10.1016/j.jss.2021.08.033

Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18 (1), 220. doi:10.1186/s13059-017-1349-1

Atianand, M. K., Caffrey, D. R., and Fitzgerald, K. A. (2017). Immunobiology of long noncoding RNAs. *Annu. Rev. Immunol.* 35, 177–198. doi:10.1146/annurev-immunol-041015-055459

Bagchi, S., Yuan, R., and Engleman, E. G. (2021). Immune checkpoint inhibitors for the treatment of cancer: Clinical impact and mechanisms of response and resistance. *Annu. Rev. Pathol.* 16, 223–249. doi:10.1146/annurev-pathol-042020-042741

Balachandran, V. P., Gonen, M., Smith, J. J., and DeMatteo, R. P. (2015). Nomograms in oncology: More than meets the eye. *Lancet. Oncol.* 16 (4), e173–180. doi:10.1016/s1470-2045(14)71116-7

Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 17 (1), 218. doi:10.1186/s13059-016-1070-5

Bray, F., Laversanne, M., Weiderpass, E., and Soerjomataram, I. (2021). The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer* 127 (16), 3029–3030. doi:10.1002/cncr.33587

Castro-Oropeza, R., Melendez-Zajgla, J., Maldonado, V., and Vazquez-Santillan, K. (2018). The emerging role of lncRNAs in the regulation of cancer stem cells. *Cell. Oncol.* 41 (6), 585–603. doi:10.1007/s13402-018-0406-4

Chen, Y., Li, Z. Y., Zhou, G. Q., and Sun, Y. (2021). An immune-related gene prognostic index for head and neck squamous cell carcinoma. *Clin. Cancer Res.* 27 (1), 330–341. doi:10.1158/1078-0432.ccr-20-2166

Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., et al. (2019). Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med.* 11 (1), 34. doi:10.1186/s13073-019-0638-6

Hama, K., Fujiwara, Y., Hayama, T., Ozawa, T., Nozawa, K., Matsuda, K., et al. (2021). Very long-chain fatty acids are accumulated in triacylglycerol and nonesterified forms in colorectal cancer tissues. *Sci. Rep.* 11 (1), 6163. doi:10.1038/s41598-021-85603-w

Hanahan, D. (2022). Hallmarks of cancer: New dimensions. *Cancer Discov.* 12 (1), 31–46. doi:10.1158/2159-8290.cd-21-1059

Heery, R., Finn, S. P., Cuffe, S., and Gray, S. G. (2017). Long non-coding RNAs: Key regulators of epithelial-mesenchymal transition, tumour drug resistance and cancer stem cells. *Cancers (Basel)* 9 (4), E38. doi:10.3390/cancers9040038

Hong, W., Liang, L., Gu, Y., Qi, Z., Qiu, H., Yang, X., et al. (2020). Immune-related lncRNA to construct novel signature and predict the immune landscape of human hepatocellular carcinoma. *Mol. Ther. Nucleic Acids* 22, 937–947. doi:10.1016/j.omtn.2020.10.002

Iasonos, A., Schrag, D., Raj, G. V., and Panageas, K. S. (2008). How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* 26 (8), 1364–1370. doi:10.1200/jco.2007.12.9791

Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., et al. (2015). The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* 47 (3), 199–208. doi:10.1038/ng.3192

Jia, Y., Chen, Y., and Liu, J. (2020). Prognosis-predictive signature and nomogram based on autophagy-related long non-coding RNAs for hepatocellular carcinoma. *Front. Genet.* 11, 608668. doi:10.3389/fgene.2020.608668

Kandimalla, R., Tomihara, H., Banwait, J. K., Yamamura, K., Singh, G., Baba, H., et al. (2020). A 15-gene immune, stromal, and proliferation gene signature that significantly associates with poor survival in patients with pancreatic ductal adenocarcinoma. *Clin. Cancer Res.* 26 (14), 3641–3648. doi:10.1158/1078-0432.ccr-19-4044

Karakousis, N. D., Papatheodoridi, A., Chatzigeorgiou, A., and Papatheodoridis, G. (2020). Cellular senescence and hepatitis B-related hepatocellular carcinoma: An intriguing link. *Liver Int.* 40 (12), 2917–2927. doi:10.1111/liv.14659

Kono, K., Nakajima, S., and Mimura, K. (2020). Current status of immune checkpoint inhibitors for gastric cancer. *Gastric Cancer* 23 (4), 565–578. doi:10.1007/s10120-020-01090-4

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). Timer: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77 (21), e108–e110. doi:10.1158/0008-5472.can-17-0307

Liu, Y., Guo, J., Shen, K., Wang, R., Chen, C., Liao, Z., et al. (2020). Paclitaxel suppresses hepatocellular carcinoma tumorigenesis through regulating circ-BIRC6/miR-877-5p/YWHAZ Axis. *Onco. Targets. Ther.* 13, 9377–9388. doi:10.2147/ott.s261700

Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular carcinoma. *Nat. Rev. Dis. Prim.* 7 (1), 6. doi:10.1038/s41572-020-00240-3

Luo, Z. F., Zhao, D., Li, X. Q., Cui, Y. X., Ma, N., Lu, C. X., et al. (2016). Clinical significance of HOTAIR expression in colon cancer. *World J. Gastroenterol.* 22 (22), 5254–5259. doi:10.3748/wjg.v22.i22.5254

Miao, D., Margolis, C. A., Gao, W., Voss, M. H., Li, W., Martini, D. J., et al. (2018). Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* 359 (6377), 801–806. doi:10.1126/science.aan5951

Mittermeier, C., Konopa, A., and Muehlich, S. (2020). Molecular mechanisms to target cellular senescence in hepatocellular carcinoma. *Cells* 9 (12), E2540. doi:10.3390/cells9122540

Montes, M., Lubas, M., Arendrup, F. S., Mentz, B., Rohatgi, N., and Tumas, S. (2021). The long non-coding RNA MIR31HG regulates the senescence associated secretory phenotype. *Nat. Commun.* 12 (1), 2459. doi:10.1038/s41467-021-22746-4

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

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

Song, J., Sun, Y., Cao, H., Liu, Z., Xi, L., Dong, C., et al. (2021). A novel pyroptosis-related lncRNA signature for prognostic prediction in patients with lung adenocarcinoma. *Bioengineered* 12 (1), 5932–5949. doi:10.1080/21655979.2021.1972078

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Tamminga, M., Hiltermann, T. J. N., Schuuring, E., Timens, W., Fehrmann, R. S., and Groen, H. J. (2020). Immune microenvironment composition in non-small cell lung cancer and its association with survival. *Clin. Transl. Immunol.* 9 (6), e1142. doi:10.1002/cti2.1142

Tatangelo, F., Di Mauro, A., Scognamiglio, G., Aquino, G., Lettiero, A., Delrio, P., et al. (2018). Posterior HOX genes and HOTAIR expression in the proximal and distal colon cancer pathogenesis. *J. Transl. Med.* 16 (1), 350. doi:10.1186/s12967-018-1725-y

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830. e814. doi:10.1016/j.immuni.2018.03.023

Toor, S. M., Murshed, K., Al-Dhaheri, M., Khawar, M., Abu Nada, M., and Elkord, E. (2019). Immune checkpoints in circulating and tumor-infiltrating CD4(+) T cell subsets in colorectal cancer patients. *Front. Immunol.* 10, 2936. doi:10.3389/fimmu.2019.02936

Van Veldhoven, C. M., Khan, A. E., Teucher, B., Rohrmann, S., Raaschou-Nielsen, O., Tjønneland, A., et al. (2011). Physical activity and lymphoid neoplasms in the European Prospective Investigation into Cancer and nutrition (EPIC). *Eur. J. Cancer* 47 (5), 748–760. doi:10.1016/j.ejca.2010.11.010

Wei, Z., Chen, L., Meng, L., Han, W., Huang, L., and Xu, A. (2020). LncRNA HOTAIR promotes the growth and metastasis of gastric cancer by sponging miR-1277-5p and upregulating COL5A1. *Gastric Cancer* 23 (6), 1018–1032. doi:10.1007/s10120-020-01091-3

Xiang, X., Fu, Y., Zhao, K., Miao, R., Zhang, X., Ma, X., et al. (2021). Cellular senescence in hepatocellular carcinoma induced by a long non-coding RNA-encoded peptide PINT87aa by blocking FOXM1-mediated PHB2. *Theranostics* 11 (10), 4929–4944. doi:10.7150/thno.55672

Xu, J., Su, Z., Cheng, X., Hu, S., Wang, W., Zou, T., et al. (2022). High PPT1 expression predicts poor clinical outcome and PPT1 inhibitor DC661 enhances sorafenib sensitivity in hepatocellular carcinoma. *Cancer Cell Int.* 22 (1), 115. doi:10.1186/s12935-022-02508-y

Zhao, L., Hu, K., Cao, J., Wang, P., Li, J., Zeng, K., et al. (2019). lncRNA miat functions as a ceRNA to upregulate sirt1 by sponging miR-22-3p in HCC cellular senescence. *Aging* 11 (17), 7098–7122. doi:10.18632/aging.102240

Zhou, P., Lu, Y., Zhang, Y., and Wang, L. (2021). Construction of an immune-related six-lncRNA signature to predict the outcomes, immune cell infiltration, and immunotherapy response in patients with hepatocellular carcinoma. *Front. Oncol.* 11, 661758. doi:10.3389/fonc.2021.661758

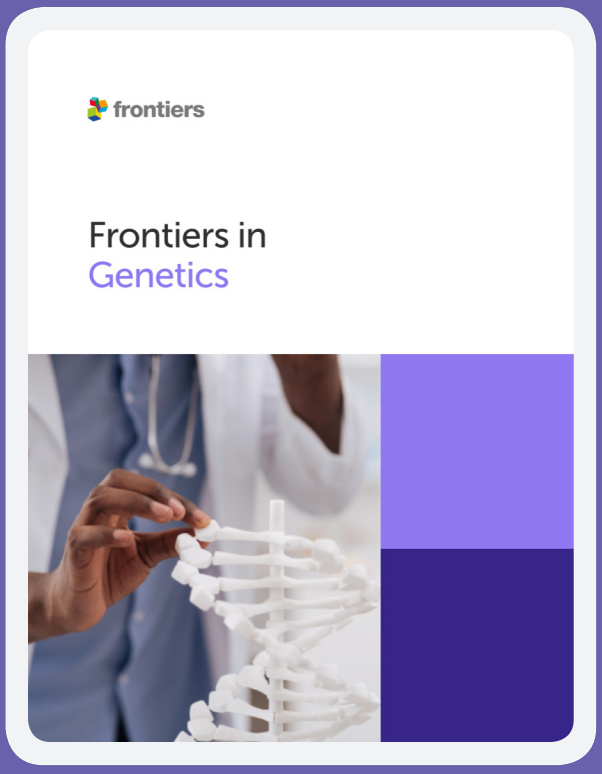
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact