Comprehensive analysis of a ceRNA network associated with TP53 mutations in liver hepatocellular carcinoma

Zongren Xu
Geneplus-Shenzhen

Hongyan Gan
Bayingolin Mongolian Autonomous Prefecture People's Hospital

Ruiqian Li
Department of Urology, the Third Affiliated Hospital of Kunming Medical University

Junhao Jia
Geneplus-Shenzhen

Lin Gao
Geneplus-Shenzhen

Zicheng Yu
Geneplus-Shenzhen

Jing Peng
Zhuzhou Central Hospital

Jie Yuan (✉ yuanjie@geneplus.org.cn)
Geneplus-Shenzhen https://orcid.org/0000-0003-4561-9117

Research Article

Keywords: ceRNA, TP53, hepatocellular carcinoma, PTBP1

Posted Date: May 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1461119/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background

Liver cancer is one of the most common causes of cancer-related death and is quantitatively dominated by the liver hepatocellular carcinoma (HCC) subtype. Although advances have been made in recent years, further exploration on new HCC carcinogenic mechanisms as well as diagnostic and therapeutic methods in the context of detrimental TP53 abnormalities is indispensable on the way to conquer HCC.

Methods

In this study, competing endogenous RNA (ceRNA) network was constructed on tumor samples stratified by the TP53 mutational status. Prognostic and therapeutic analyses were conducted on the identified network nodes. Multi-source databases were used for validations.

Results

The constructed ceRNA network demonstrated a TP53 mutant-specificity. Additionally, the ceRNA network nodes including AC017104.6, RP5-1092A11.5, RP11-365016.6 and PTBP1 were all prognosis-related. Further therapeutic correlation analysis unveiled the connection between their expressions and tumor microenvironment fluctuation as well as possible chemotherapy resistance. We also explored the drug sensitivity difference between TP53 wild type and mutant samples, complementing the therapeutic arsenal of HCC. The universality of the key node PTBP1 was verified by transcription and protein level expressions.

Conclusions

This is the first study focusing on the ceRNA network mediated by TP53 mutations and its provoked therapeutic associations in HCC. Our results deepened the understanding of the concomitant alternations brought by TP53 mutations and could benefit future therapeutic development in HCC.

Introduction

Being the second most archetypal deadly cancer worldwide\(^1\), liver cancer prevailed in underdeveloped areas and emerged as the fifth and ninth most common cancer in males and females respectively. When narrowing to Asian countries, the severity of the disease exacerbated due to the frequent chronic viral hepatitis infection or exposure to toxins. For example, the 5-year survival rate is as low as 12% in China\(^2\). Hepatocellular carcinoma (HCC), one cancer type originated from liver cells, accounted for more than 80% of global primary liver cancer amount\(^3\) and standards of its diagnosis and treatment have been
increasingly improved. For early-stage HCC, liver resection and transplantation were proven to be the surpassing treatment option. Disastrously, HCC early-stage surveillance and diagnosis still remains a challenge, resulting in the majority of HCC patients at initial diagnosis was at the intermediate stage. With the presence of vascular invasion and extrahepatic spread, curative treatment became impotent in late stage HCC and systemic therapy was proven an indispensable surrogate treatment option. Comparing to treatments like radiotherapy or chemotherapy, targeted therapy is now a more promising option for HCC regarding its improved efficacy. Multiple molecular therapies including tyrosine kinase inhibitors (TKIs), monoclonal antibodies and antisense oligonucleotides have been developed and researchers are paying indomitable efforts in continuous targeted drug development for live cancer.

Encoded by the human \( TP53 \) gene, P53 protein has long been known as the guardian of the genome, regarding its role in DNA repair and recombination. Unfortunately, somatic mutations in \( TP53 \) are one of the most common alternations in human cancer, which could demolish the genome stability, enhance tumor growth, affect the metabolic function of mitochondria and cause drug resistance. Unlike other tumor suppressor genes (TSGs) that were often debilitated by truncated mutations or deletions, missense mutations dominated the \( TP53 \) alteration landscape, producing full-length protein with a single amino acid substitution. In HCC, multiple oncogenes or TSGs including \( TERT \), \( TP53 \), \( CTNNB1 \), etc were reported to harbor tumorgenetic mutations, among which the association between \( TP53 \) mutated phenotype and shorter overall survival (OS) and recurrence-free survival (RFS) was previously revealed. Considering the ubiquitousness and severeness of \( TP53 \) mutations, researchers have been dedicated to developing drugs targeting \( TP53 \) aberrance. However, cancer therapies targeting mutant \( TP53 \) are mostly being tested in early phase clinical trials, hampering their extensive application in large and diverse populations. Owing to the inherent functional diversity of \( TP53 \) and its multi-faceted impact on tumor progression, the deriving of \( TP53 \) mutation concomitant biomarkers could undeniably facilitate the development of combination cancer therapy and benefit patients.

Recently, we have witnessed a surge of investigation into novel biomarker prediction for precision medicine, especially through bioinformatics approaches. Among multifarious biomarker identification methods based on mathematics, network topology and machine learning theories, the construction of competing endogenous RNA (ceRNA) networks has received particular attention due to its high mechanistic consistency with biological hypothesis and the affluence in knowledge discovery. Through the establishment of ceRNA network, the crosstalk between microRNA (miRNA), long non-coding RNA (lncRNA) and messenger RNA (mRNA) could be discovered, providing an unprecedented layer of miRNA regulatory network. Increasing evidence have shown that the ceRNA crosstalk participates in indispensable regulatory processes and its perturbation will scramble the balance of cell regulatory system, resulting in disease initiation and tumor occurrence. In HCC, ceRNA networks aided the prioritization of RNA biomarkers including miRNA\(^{12}\) and lncRNA\(^{13}\) by comparing normal controls with tumor patients. As previously mentioned, the regulatory peculiarity of \( TP53 \) granted its capability in orchestrating complicated biological processes. By incorporating \( TP53 \) mutational information in ceRNA network construction, the consequent genetic fluctuations caused by \( TP53 \) could be uncovered and
compared, which will provide unseen therapeutic targets and pave the way to systematic therapy advancement.

In this study, HCC expressional data was collected from The Cancer Genome Atlas (TCGA) database and stratified by the \( TP53 \) gene mutational status. By using these profiling data, a ceRNA network specifically existed in \( TP53 \) mutant samples was primarily established. After confirming the prognostic relevance of the network nodes, analyses on immune cell infiltration, chemotherapy resistance and drug response were further conducted. Generality of the key mRNA node was guaranteed by multi-source database validations. Clinically, these gene nodes could possibly benefit the development of systemic therapy as well as aid the treatment outcome prediction in HCC and improve the unsatisfactory prognosis of HCC patients.

**Materials And Methods**

**TCGA data preparation**

Firstly, the liver hepatocellular carcinoma (HCC) HTSeq-Count file from RNA-seq and isoform expression quantifications from miRNA-seq were retrieved from TCGA GDC Data Portal (https://portal.gdc.cancer.gov/) while the clinical information was downloaded from cBioPortal (https://www.cbioportal.org/) database. Patients met the following criteria were further discarded: i) harboring \( TP53 \) synonymous mutations; ii) lacking RNA and miRNA sequencing data. Additionally, we used miRBase\(^{14}\) MIMAT ID for miRNA quantifications and only mature miRNAs were kept and the summed read counts were calculated for further use.

**Differentially-expressed gene identification procedure**

The derivation of differentially-expressed mRNAs (DEMs), IncRNAs (DELs) and miRNAs (DEmis) were separately conducted between 365 primary solid tumor patients with two \( TP53 \) mutational status (111 \( TP53 \) mutant and 254 wild type) and 50 normal samples. More specifically, GENCODE v22 annotation files were primarily downloaded from GENCODE database (https://www.gencodegenes.org/) for gene type, length and gene symbol information retrieval. After eliminating IncRNAs, miRNAs and mRNAs with sample-wise average read count<1, a total of 8782 mRNAs, 3609 IncRNAs and 2166 miRNAs were retained for further analyses. For RNA-seq data, read counts of IncRNAs and mRNAs were normalized by TCGAbioliinks R package\(^{15}\) and DESeq2, limma and edgeR tools\(^{16-18}\) were utilized for differentially-expressed gene (DEG) identification between \( TP53 \) mutant type (MT) and normal as well as \( TP53 \) wild type (WT) and normal samples. We kept the genes with \( p \)-value<0.05 and absolute value of log2(Fold Change)>1 as candidate DEGs. Regarding the intrinsic statistical differences between the algorithms, the DEG candidates from tools were intersected to get convergent lists for two \( TP53 \) mutational statuses. For miRNA-seq data, quantile normalization using TCGAanalyze_Filtering function was initially conducted and DESeq2, limma and edgeR tools similarly facilitated the DEmi derivation. Identical thresholds on p-
value and log2(Fold Change) were applied on miRNAs and gene lists were intersected for further use. Conversions between MIMAT IDs and miRBase identifiers were completed by miRBaseVersions.db package (https://github.com/StefanHaunsberger/mirbaseversions.db) and statistics on the gene intersection between the three tools were performed by VennDiagram R package\textsuperscript{19}.

**Functional annotation on gene sets**

Pathway enrichment analyses on the convergent DEM lists of two mutational statuses were conducted using Enrichr R package\textsuperscript{20}, which integrated comprehensive gene-set annotations. Resources including KEGG pathway\textsuperscript{21}, Gene Ontology (GO) Biological Processes\textsuperscript{22} and Reactome database\textsuperscript{23} were adopted in our enrichment analyses. Enriched GO terms with p-value<0.01 and enriched KEGG and Reactome pathways with p-value<0.05 were reported.

**TP53 mutation-specific ceRNA network construction**

In accordance with the definition of ceRNAs which regulates the expression of mRNAs by competing for shared miRNAs, the ceRNA network was constructed using following procedures: i) DEMs, DELs and DEmis from the two \textit{TP53} mutational statuses were collected and co-expression networks were firstly constructed based on the expressions of DEM/DEL, DEL/DEmi and DEM/DEmi pairs. Only pairs with Pearson Correlation Coefficient (PCC)>0.6 in DEM/DEL network and negatively-correlated pairs in DEL/DEmi or DEP/DEmi network were retained. ii) miRNA targets from miRNA interaction databases including miRDB\textsuperscript{24} version 6.0, miRTarBase\textsuperscript{25} release 8.0 and miRcode\textsuperscript{26} version 11 were curated. Interactions with predicted score>80 were retained in miRDB database for confidence pursuance. Considering the incompatibility of gene identifiers in the multi-source interactions, R packages miRBaseVersions.db and org.Hs.eg.db (https://bioconductor.org/packages/release/data/annotation/html/org.Hs.eg.db.html) further aided the gene identifier conversion. Only DEM/DEmi interactions shared by miRDB and miRTarBase were used for further network construction. iii) For each \textit{TP53} mutational status-specific DEmi, the DEM targets in the shared interaction set were firstly retrieved and such interactions were intercepted with DEM/DEmi co-expression network, forming DEM/DEmi pairs in the ceRNA network. Analogously, the DEL targets of these DEmis were obtained from miRcode and intercepted with DEL/DEmi co-expression networks. When incorporating the mRNA-miRNA-IncRNA triplet in final ceRNA networks, only pairs existed in DEM/DEL co-expression networks were maintained. The Cytoscape\textsuperscript{27} package was applied for ceRNA network visualization and the expressional variation of network nodes was illustrated by pheatmap R package (https://cran.r-project.org/web/packages/pheatmap/index.html).
Key prognosis-related node identification in ceRNA network

Survival analysis was conducted for the prioritization of key prognostic nodes in constructed ceRNA networks. To begin with, the normalized and log2-transformed counts of genes were integrated with survival information. Later univariate cox regression was conducted using survival R package\textsuperscript{28} to examine the associations between gene expression levels and Overall Survival (OS) and Disease-Specific Survival (DSS). Significant nodes with p-value<$0.05$ were considered to be prognosis-related and were further subjected to Kaplan-Meier (K-M) curve generation by survminer package\textsuperscript{29}. A p-value<$0.05$ in log-rank test was considered with statistical difference.

Immune infiltration analyses for ceRNA network nodes

To detect potential associations between ceRNA network prognosis-related nodes and immune cell infiltration levels, the immune infiltration analyses were conducted on curated tumor samples. More specifically, gene-level FPKM values were downloaded from Xena database\textsuperscript{30} and Ensembl gene IDs were converted to gene symbols using in-house scripts. Later R package Immunedeconv\textsuperscript{31} was used for the deconvolution of gene expressions and the derivation of the immune cell composition in tumor microenvironment. The xCell method\textsuperscript{32} implemented in Immunedeconv package was selected regarding its capability of portraying the enrichment landscape of the most immune and stroma cell types. Activities of cancer-immunity cycle\textsuperscript{33} steps were addedly assessed by single sample Gene Set Enrichment Analysis (ssGSEA) from GSVA R package\textsuperscript{34}. Associations between nodes in ceRNA networks and immune infiltration analysis results as well as GSVA enrichment scores were further assessed by Spearman's rank correlation.

The deduction of potential chemotherapy resistance for ceRNA network nodes

To elucidate the association between ceRNA network nodes and possible chemotherapy resistance in HCC with different $TP53$ mutational statuses, chemotherapy resistance-related genes were firstly collected from literatures\textsuperscript{35,36}. More detailedly, genes related to i) tumor microenvironment (TME), including TME remodeling, cytokine and cancer stem cell; ii) liver cancer specific events, including hepatitis B virus infection, HCC-related protein 1 (HCRP1) and HCC target therapy; iii) specific pathways and mechanisms, including transmembrane transporter, DNA damage and repair, telomerase, lipid metabolism, head shock protein, autophagy, apoptosis protein and chemotherapy critical pathway; iv) transcriptional and translational regulators, including topoisomerase, epigenetics and translational factor were incorporated in the analyses. Associations between the four gene categories and ceRNA network nodes were further quantified by Spearman's rank correlation coefficients (SCC).
Methylation analysis on key mRNAs in ceRNA network

To explicate the expressional alternations of ceRNA network nodes, Illumina 450k methylation data was downloaded from Xena database. Experimental variations were further eliminated by discarding i) probes with not available (NA) value in more than 5% of samples; ii) samples with NA value in more than 20% of probes. Later BMIQ algorithm implemented in watermelon package was used for the bias correction of probe-level methylation values and the normalized methylation intensities of ceRNA network nodes were scrutinized in samples with different TP53 mutational statuses. Finally, we conducted expressional comparisons on common 5-methylcytosine (5mC) methyltransferase (writer), demethylase (eraser) and methylation recognition (reader) proteins in TP53 MT, TP53 WT and normal sample groups. Wilcoxon rank sum test was used in all group-wise comparisons.

Drug response difference investigation in patients with distinct TP53 mutational statuses

Data of drug response was download from the Genomics of Drug Sensitivity in Cancer database (GDSC, https://www.cancerrxgene.org/) which contains drug sensitivity data measured by fluorescence-based cell viability assays on 805 cell lines subjected to 72 hour treatment of 198 drugs. The sensitivity data was used to construct a ridge regression model by oncoPredict R package, which further aided the imputation of drug response values on HCC expressional data. Group-specific drug candidates were selected by comparing the drug response differences between TP53 MT and TP53 WT groups. A p-value<0.05 in Wilcoxon rank sum test was considered significant.

Multi-omics validation on the expression levels of ceRNA network components

Multi-omics HCC datasets were collected from International Cancer Genome Consortium (ICGC, https://dcc.icgc.org/) and Cancer Therapeutics Response Portal (CTRP, https://portals.broadinstitute.org/ctrp) databases for ceRNA network component validation. After removing samples possessing TP53 synonymous mutations, RNA expression levels of 84 TP53 MT and 202 normal samples from ICGC database were used for verification. Similarly, the difference was measured by Wilcoxon rank sum test. As for the CTRP database, protein expression levels of ceRNA network key node in 9 TP53 MT and 5 TP53 WT samples were selected for comparison. In addition, immunohistochemistry data was retrieved from the Human Protein Atlas (HPA, https://www.proteinatlas.org/) database to validate the protein levels of key network nodes in normal and tumor liver tissues.

Results

Dysregulated genes were identified in patients with distinct TP53 mutational statuses
As mentioned in Materials and methods, after rounds of filtrations, a total of 365 tumoral and 50 normal HCC samples were involved and the tumor samples were further divided into TP53 MT and TP53 WT group. Notably, patients with distinct TP53 mutational status exhibited significantly different OS and disease-free survival (DFS) (Suppl. Fig. 1A-B), connoting the detrimental effects of TP53 mutations in HCC. By later DEG identification procedures, a number of DEM, DEL and DEmi were identified by DESeq2, limma and edgeR tools in samples of two mutational statuses (Suppl. Fig. 2A-F, Suppl. Fig. 3A-F). When further inspecting the tool-wise gene intersection, 666 DEMs, 440 DELs and 34 DEmis were recurrently found up-regulated in TP53 MT group than normal control (Suppl. Fig. 2A, C, E) while 419 DEMs, 335 DELs and 10 DEmis showed higher expression in TP53 WT samples that the control set (Suppl. Fig. 2B, D, F). Besides, 276 DEMs, 49 DELs and 15 DEmis were down-regulated in TP53 MT samples (Suppl. Fig. 3A, C, E) while 167 DEMs, 34 DELs and 17 DEmis presented an expressional demotion in TP53 WT patients than the normal samples (Suppl. Fig. 3B, D, F).

With the aim of gaining insights into the accompanying changes granted by TP53 abnormality, DEMs obtained from the two groups were subjected to functional annotations by GO, KEGG and Reactome databases. For the TP53 MT subgroup, most of the up-regulated genes were enriched in biological processes or pathways related to cell migration, angiogenesis and apoptosis (Fig. 2A-C) while activities of metal iron response and transport-related pathways were suppressed (Fig. 2A-C). Similar pathway enrichments for the dysregulated genes were observed in TP53 WT group (Suppl. Fig. 4A-C). Interestingly, immune-related terms were found up-regulated in two subgroups and PD-1 checkpoint-related pathway was concurrently impaired. Concerning the impotence of anti-PD-1/PD-L1 blockers in major HCC patients, investigation on the TME could possibly provide additional therapeutic benefits for HCC.

Construction of TP53 alternation-specific ceRNA network aided the elicit of novel prognostic markers in HCC

By combining the expressional correlation and miRNA targeting information, TP53 mutational status-specific ceRNA networks were constructed based on the ceRNA hypothesis. As shown in Figure 3A, multiple genes were found involved in the regulations provoked by TP53 mutation. Interestingly, no ceRNA network was identified using the same construction strategy and DEGs in TP53 WT group, which affirmed the causal effect of TP53 abnormality on tumorigenesis or tumor progression. More explicitly, for the TP53 MT ceRNA network, IncRNAs (AC017104.6, RP5-1092A11.5 and RP11-365016.6) and mRNA (PTBP1) both exhibited expressional elevation in tumor samples while the only miRNA (hsa-miR-133b) showed opposing trend (Fig. 3B). When further examining tumor samples stratified by TP53 status, significant expressional increment of the aforementioned IncRNA and mRNA was observed in TP53 MT subgroup (Fig. 3C-F), again emphasized the importance of our identified regulatory network in biological process decoding.
To better understand the functional importance of the ceRNA network arose from TP53 abnormality, we performed cox regression analysis for the screening of prognosis-related biomarkers. As illustrated by the forest plot in Figure 4A, the univariate cox regression analysis confirmed the independent prognostic relevance of PTBP1, AC017104.6, RP5-1092A11.5 and RP11-365016.6 with OS, all of which reached Hazard Ratio (HR) > 1 and p-value < 0.05. However, when inspecting the relationship between nodes’ expression and DSS, only PTBP1 showed significant relevance (Fig. 4B). Next, we conducted survival analysis on these independent prognostic factors to unveil their potential predictive power on patient survival and confirm their regulatory significance in HCC. As expected, the survival of TP53 MT patients presented a negative trend with the expression of PTBP1, both for OS (Fig. 4C) and DSS (Fig. 4D). In addition, reverse associations between expression and patient OS were analogously observed on lncRNAs including AC017104.6 (Suppl. Fig. 5A), RP5-1092A11.5 (Suppl. Fig. 5B) and RP11-365016.6 (Suppl. Fig. 5C). Since the activation the above four ceRNA network nodes led to poor prognosis, we hypothesized their functional collaboration in the context of TP53 mutations. Indeed, patients in TP53 MT group could be separated by expressional patterns (Suppl. Fig. 5D) and the patients with higher expressions exhibited worse OS (Suppl. Fig. 5E) and DSS (Suppl. Fig. 5F). Through the ceRNA network construction in TP53 MT group, novel HCC prognosis-related genes were identified, which could deepen our understanding of TP53 gene’s mechanical role in cancer.

The ceRNA network established by TP53 mutations was associated with tumor microenvironment alterations in HCC

In the previous functional annotations on TP53-related DEGs, immune-related terms were uncovered. We next scrutinized the relationship between TP53-provoked regulatory alternations and TME. To be more precise, the prediction as well as quantification of the immune cell types were initially performed by the xCell method on TP53 MT samples and correlations between ceRNA network nodes and immune infiltration levels were further analyzed. As shown in Figure 5A, infiltration level of common lymphoid progenitor significantly correlated with the expression of PTBP1, AC017104.6, RP5-1092A11.5 and RP11-365016.6. Likewise, infiltrations of T cell CD4+ T helper (Th) 2, T cell CD4+ memory and mast cell also formed a correlation module with ceRNA node genes (Fig. 5A). To be more precise, among the four cell types that reached statistical significance, PTBP1 (Fig. 5B-E) and RP5-1092A11.5 (Suppl. Fig. 6) showed the highest SCC with T cell CD4+ Th2 infiltrations while the most relevant immune cell types for AC017104.6 and RP11-365016.6 were common lymphoid progenitor (Suppl. Fig. 7) and mast cell (Suppl. Fig. 8). Such associations deepened the understanding of the crosstalk between TME and TP53 alternation-provoked ceRNA network and possibly could benefit the immunotherapy development in HCC.
We next investigated whether the association between HCC TME perturbation and ceRNA network was TP53 mutant-specific. By comparing the predicted levels of immune cell infiltration between TP53 MT and WT group, 15 immune cell types reached statistical significance (adjusted p-value<0.05, Wilcoxon rank sum test adjusted by Benjamini & Hochberg method) (Suppl. Fig. 9A). Interestingly, infiltration levels of the immune cell types associated with ceRNA network nodes including T cell CD4+ Th2, mast cell, and T cell CD4+ memory were remarkably different between the groups stratified by TP53 mutational status (Suppl. Fig. 9A) and the elevated infiltration of these cell types conferred pernicious effect on TP53 MT patient survival, both for OS (Suppl. Fig. 9B, D, F) and DSS (Suppl. Fig. 9C, E, G). To conclude, our analyses confirmed the TP53 mutant-specific association between ceRNA network and TME variations, which could pose a pernicious effect on patient survival.

Additionally, the association between ceRNA network nodes and cancer-immunity cycle activity was examined by ssGSEA. The co-expressed network nodes exhibited a high correlation with the activity of “infiltration of immune cells (T cells) into tumors” (Suppl. Fig. 10A). The subsequent evaluation confirmed the statistical significance of the relationships between the enriched step and network nodes’ expression levels (Suppl. Fig. 10B-E). However, when further comparing the activity of “infiltration of immune cells (T cells) into tumors” between TP53 MT and WT group, no obvious difference was observed. Combining with the discoveries above, TP53 mutant-specific network could possibly trigger the infiltration of specific T cell subtypes, which offered a boost to the development of precision immuno-medicine.

**TP53 mutant-specific ceRNA network genes were connected with chemotherapy resistance**

We next scrutinized the relationship between ceRNA network nodes and possible chemotherapy resistance. Through the correlation analysis between chemotherapy resistance-related genes and ceRNA network nodes, we found that PTBP1, RP5-1092A11.5 and AC017104.6 exhibited high correlation with respective sets of chemo-related genes (Suppl. Fig. 11A) while RP11-365016.6 exhibited a weak connection with chemotherapy resistance. More detailedly, PTBP1 co-expressed with multi-facet chemotherapy resistance-related genes (Fig. 6A), covering all four categories we defined. As for RP5-1092A11.5, it was closely related to chemotherapy resistance caused by transmembrane transporters, critical (specific) pathways, lipid metabolism and HCC target therapy (Fig. 6B) while AC017104.6 demonstrated associations with chemotherapy resistance factors related to TME, cytokine, cancer stem cell, heat shock proteins, specific pathways and epigenetics (Fig. 6C). Unsurprisingly, among other participants in ceRNA crosstalks, mRNA PTBP1 presented the strongest correlation with chemotherapy resistance factors. Interestingly, the expression of these resistant genes gradually elevated in normal, TP53 WT and TP53 MT groups (Suppl. Fig. 11B), again reflected the biological complexity prompted by the TP53 mutations and emphasized the difficulty as well as cruciality of treatment development in HCC.
Transcription of genes could be affected in multifarious ways. Intending to uncover the cause of PTBP1 expressional elevation in TP53 MT patients, we collected the DNA methylation levels of PTBP1-related probes and checked their dynamics in different patient groups. Among all PTBP1-related probes, 7 of them demonstrated an observable group-wise difference (Suppl. Fig. 12A) while only probe cg02086742 possessed methylation level that anti-correlated with the expressional trend of PTBP1 (Suppl. Fig. 12A). As expected, methylation levels of cg02086742 were significantly lower in TP53 MT group (Suppl. Fig. 12B), implying the potential of applying it as a target of epigenetic therapeutics in HCC.

Moreover, regarding their pivotal regulatory effects on DNA methylation, the expression of DNA methylation writer (DNMT1, DNMT3A and DNMT3B), eraser (TET1, TET2 and TET3) and reader (MeCP2, MBD1 and MBD2) proteins were compared between groups. We only focused on proteins with significant group-wise difference as well as alternation trend conforming to their biological characteristics. As a result, two eraser proteins TET1 (Suppl. Fig. 12C) and TET3 (Suppl. Fig. 12D) met the selection criteria and were strongly correlated with the methylation of probe cg02086742 and expression of PTBP1 in three patient groups. These observations provided a possible novel strategy for targeted therapy in HCC.

Patients stratified by TP53 mutational status exhibited differences in drug sensitivity

With the aim of facilitating systematic therapy development, analyses on immunotherapy, chemotherapy and targeted therapy were conducted above. Next, we investigated the drug response diversity in HCC TP53 MT and WT groups. By utilizing the drug sensitivity data in GDSC, the half-maximal inhibitory concentration (IC50) value of patients in the two groups were predicted. Among all drugs tested, 7 drugs demonstrated a group-wise statistical difference (p-value<0.05) (Figure 7A). More specifically, drugs AZD8055, OSI.027 and RO.3306 showed higher effectiveness in TP53 MT patients (Fig. 7B-D), among which AZD8055 got the lowest IC50 values (around 1). Besides, the group-wise ID50 value difference exacerbated for the OSI.027 compound (Fig. 7C). As for drugs JQ1, NU7441, SB216763 and ZM447439, their IC50 values were significantly lower in TP53 WT group (Suppl. Fig. 13A-D). Compound SB216763 presented relatively higher IC50 values, while JQ1 got the lowest IC50 values in TP53 WT HCC, which warranted its therapeutic strength.

Multi-omics validation of PTBP1 in public databases

To verify the universality of our discoveries and eliminate possible biological biases, expressional data from ICGC and CTRP databases were collected. For ICGC database, the RNA expression level of PTBP1 in TP53 MT samples was significantly higher than the normal control group (Fig. 8A). When examining the
protein expressional levels in CTRP database, though not reaching statistical significance, a certain reduction in \textit{TP53} WT samples was observed (Fig. 8B). These verifications shared high consistency with our discoveries in ceRNA network analyses and underpinned the connection between \textit{TP53} mutations and \textit{PTBP1} alternations.

In addition, the immunohistochemistry of \textit{PTBP1} in liver cancer was retrieved from the HPA database to provide an experimental point of view on protein expression. As illustrated in Figure 8C, the prevalence of \textit{PTBP1} expression in HCC was confirmed, again indicated its potential in HCC patient stratification and treatment.

**Discussion**

Liver hepatocellular carcinoma (HCC), one of the most common and mortal cancer type, has attracted pullulating attention to mechanism exploration and therapy development. Owing to the perturbative effect and usualness of \textit{TP53} mutational abnormality in HCC, a comprehensive investigation on the altered microenvironment and its connection with systemic therapy could assist the development of novel medicinal strategies. In this study, HCC tumor samples from TCGA database were primarily stratified by \textit{TP53} mutational status. Based on the DEMs, DELs and DEmis identified from groups, \textit{TP53} mutant-specific ceRNA network was constructed. After confirming the prognostic value of key network node genes, therapeutic correlation analyses on node expression and immune cell infiltrations, cancer-immunity cycle activity, chemotherapy resistance were conducted. Group-wise drug sensitivity as well as the DNA methylome of node genes was also scrutinized. Finally, generality of the identified gene target was tested by public databases. To our knowledge, this is the first study focusing on the ceRNA network mediated by \textit{TP53} mutations and its therapeutic associations in HCC, which could undoubtedly complement the gap in precision medicine.

Basing on the premise that patients stratified by \textit{TP53} mutational status exhibited significantly different survival and the functional diversity of the gene, we utilized ceRNA network to unveil the genetic alternations as well as pathways provoked by \textit{TP53} abnormality. Most of the previous applications of ceRNA network focused on the discrepancy regulators between tumor and normal samples\textsuperscript{44,45}. However, the intrinsic heterogeneity of tumor often engenders homogeneous cancer subtypes possessing distinct molecular features and therapeutic responses, which could invalidate most of the biomarkers from previous studies. Emerging studies have incorporated tumor heterogeneity in ceRNA network analyses\textsuperscript{46,47} but works considering key driver mutations were scarce. Interestingly, the DEGs identified in patients separated by \textit{TP53} mutational status were enriched in similar pathways, denoting that the effect conferred by \textit{TP53} anomaly should be investigated in a more microcosmic way, e.g. through the construction of regulatory networks. Noticing our ceRNA network only existed in \textit{TP53} mutated patients, the specificity, veracity and efficacy of the biomarkers in network was emphasized.

Multiple gene targets including hsa-miR-133b, AC017104.6, RP5-1092A11.5, RP11-365O16.6 and \textit{PTBP1} were prioritized in our ceRNA network construction procedure, some of which were reported to be involved
in cancer-specific ceRNA networks. For example, hsa-miR-133b was identified as a prognostic marker in esophageal cancer\textsuperscript{48} and cervical cancer\textsuperscript{49}. Our results italicized its role in \textit{TP53}-mutant HCC homeostasis maintenance. As for the three lncRNA nodes, we uniquely reported their prognostic value as well as the associations with immune infiltration and chemotherapy resistance in HCC, again supplemented the therapeutic gene target set. Functioning as the only mRNA in our network, \textit{PTBP1} demonstrated the strongest correlation with immune cell infiltrations and chemoresistance. With the ability to control almost all steps of messenger RNA metabolism and processing\textsuperscript{50}, the elevation of \textit{PTBP1} caused an adverse effect in glioblastoma\textsuperscript{51} and colorectal cancer\textsuperscript{52}. In HCC, relationships between \textit{PTBP1} up-regulation and tumor growth\textsuperscript{50}, tumor invasion as well as metastasis\textsuperscript{53} were uncovered. Our work for the first time proved the detriment of \textit{PTBP1} on \textit{TP53}-mutant patient survival. Considering the relationship between expressions of \textit{PTBP1} and \textit{TP53} in glioma\textsuperscript{54}, we speculate that the \textit{TP53} mutational deactivation in HCC triggered the higher expression of \textit{PTBP1}, which quantity was maintained by our identified ceRNA network and acted as the accomplice of mutated \textit{TP53} in tumor development.

In the immune-related analyses, infiltration level of common lymphoid progenitor cells was found closely associated with expressions of ceRNA network nodes but the cell prevalence disparity between \textit{TP53 MT} and \textit{TP53 WT} group did not reach statistical significance. Since common lymphoid progenitor cells give rise to multiple types of lymphocytes, we suspect the key pathways’ activity as well as the infiltration of some differentiated lymphocyte in the lymphoid lineage could be different between \textit{TP53 MT} and WT group. Indeed, the infiltrations of T cell CD4 + Th2, T cell CD4 + memory and mast cell were statistically different in the two groups. For the T cell CD4 + Th2 and CD4 + memory T cells, they demonstrated higher infiltration levels in \textit{TP53 MT} patients. A recent study reported that Th2 cells initiate an anti-tumor response through the MHC II complex pathway\textsuperscript{55} but the elimination of tumors is associated with CD8 + T cells. Regarding the observation on impaired PD-1 related pathway in HCC, adoptive CD8 + T cell therapy (ACT) could possibly benefit \textit{TP53 MT} HCC patients. Besides, CD4 + T memory cells could instigate secondary immune responses by influencing CD8 + T cell response early events\textsuperscript{56}, again demonstrated the potential of applying ACT in specific HCC groups. Moreover, the two-sidedness of mast cells in tumor development was previously reported\textsuperscript{57}. An elevated infiltration of mast cells was found in our \textit{TP53 WT} HCC samples while higher mast cell prevalence correlated with worse prognosis in \textit{TP53 MT} group. We suspect \textit{TP53} mutations triggered such cellular functional switch and approaches targeting mast cells could be a new therapeutic option for HCC.

In our multi-faceted analyses on systematic therapy, chemotherapy resistance mechanisms were found closely associated with ceRNA network nodes, among which the key mRNA \textit{PTBP1} demonstrated the strongest correlation. Aside from prioritizing it as a novel inhibition therapy target, we also proposed possible directions of epigenetic therapeutics. Moreover, drugs including AZD8055, OSI.027 and RO.3306 exhibited higher efficacy in \textit{TP53}-mutated HCC patients. All of them hamper the cell proliferation process. Considering the unrestricted cell cycle granted by \textit{TP53} mutational abnormality, these drugs could function as therapeutic surrogates in the way of conquering liver cancer harboring \textit{TP53} mutations.
However, our study still has some limitations. Firstly, only data from TCGA database was used in our TP53 mutant-specific ceRNA network construction procedure, which may cause possible racial bias on the results. With the advent of multi-omics sequencing data in multiracial liver cancer, these potential partialities could be eliminated. Secondly, relatively rigorous RNA inclusion and exclusion criteria were applied in our analyses, which could nurture our network’s specificity but limit the number of gene targets identified. Attempts pursuing a perfect balance between network breadth and depth should be made in the future studies. Thirdly, we confirmed the prognosis relativity of the key ceRNA network nodes in TCGA data. Further investigations incorporating multi-source samples with outcome information could better consolidate our discoveries. Moreover, our conclusions on immunotherapy and chemotherapy were mainly derived from correlation analyses. With the availability of future data measuring the transcriptomic dynamics conferred by therapies, the underlying association between our ceRNA network and precision medicine could be further elucidated in HCC. Last but not least, most of our conclusions were sorely data-based. Though multiple databases were involved in the validation of the network nodes, using experiments including quantitative polymerase chain reaction (qPCR), western blot as well as immunofluorescence (IF) to validate the expressions of key network members in cell lines or curated clinical samples could undoubtedly solidify our findings. Besides, techniques including luciferase assay and cross-linked immuno-precipitation could also aid the unveiling of biological mechanisms amongst the network nodes. By applying experiments in the subsequent studies, the role of our ceRNA network in TME modulation could be further illustrated.

Conclusions

In conclusion, the present study identified a ceRNA network mediated by TP53 mutations and focused on the therapeutic alternations associated with the network, providing unprecedented prognostic and diagnostic markers for HCC. Additionally, our comprehensive analyses complemented the understanding of systemic therapy in HCC, which could undoubtedly promote the development of precision medicine in HCC.

Abbreviations

HCC: hepatocellular carcinoma

ceRNA: competing endogenous RNA

OS: overall survival

miRNA: microRNA

IncRNA: long non-coding RNA

mRNA: messenger RNA
TCGA: The Cancer Genome Atlas

DEM: differentially-expressed mRNA

DEL: differentially-expressed lncRNA

DEmi: differentially-expressed miRNA

DEG: differentially-expressed gene

MT: mutant type

WT: wild type

GO: gene oncology

PCC: Pearson Correlation Coefficient

DSS: disease-specific survival

K-M: Kaplan-Meier

ssGSEA: single sample Gene Set Enrichment Analysis

TME: tumor microenvironment

SCC: Spearman's rank correlation coefficients

DFS: disease-free survival

HR: Hazard Ratio

Declarations

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest

All the authors declare that they have no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.
Informed consent

Not applicable.

Availability of data and materials

The datasets analyzed during the current study are available in the TCGA GDC Data Portal (https://portal.gdc.cancer.gov/) and cBioPortal database (https://www.cbioportal.org/).

Authors’ Contributions

JY contributed to the study design. ZX, HG, RL, JJ, LG, ZY and JY analyzed the data. ZX drafted the manuscript. JY and JP revised the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgements

Not applicable.

References


Figures
Figure 1

Workflow of the TP53 mutation-related ceRNA network construction in liver hepatocellular carcinoma and the successive comprehensive analysis procedure.
Figure 2

Enrichment analyses on DEMs identified from TP53 MT HCC samples. (A) GO Biological Process enrichment analysis results. The enriched terms for up-regulated and down-regulated DEMs were marked with red and cyan colors. (B) Similar with (A), but using KEGG database. (C) Similar with (A), but using Reactome database.
The constructed ceRNA network possessed a TP53 mutant-specificity. (A) Visualization of the constructed ceRNA network. (B) Heatmap illustrating the expression of ceRNA network nodes in three groups. (C-F) Expressional comparisons on ceRNA network nodes in TP53 MT and WT group, including (C) PTBP1, (D) AC017104.6, (E) RP5-1092A11.5, and (F) RP11-365O16.6. Values were compared using Wilcoxon rank sum test. ****: p-value<=0.0001.
Figure 4

Prognostic analysis of ceRNA network nodes and survival comparisons on groups stratified by PTBP1 Expression. (A-B) Forest plots of univariate cox regression analysis on ceRNA network nodes using overall (A) survival and (B) disease-specific survival information. (C-D) Kaplan–Meier survival curves showed the (C) overall survival and (D) disease-specific survival of patients stratified by high/low-PTBP1 expression.
Figure 5

Relationship between immune cell infiltrations and ceRNA network nodes. (A) Heatmap illustrating the correlations between immune-infiltrating cells and ceRNA network nodes in aided the imputation of drug response values TP53 MT patients. Clusters with high correlation values were marked with red rectangle. (B-E) Correlation analyses between PTBP1 expression and four kinds of immune infiltrating cells in TP53
MT group, including (B) Common lymphoid progenitor, (C) T cell CD4+ T helper (Th) 2, (D) T cell CD4+ memory and (E) Mast cell. Correlations were quantified by Spearman's rank correlation coefficients.

**Figure 6**

Inspection on the chemoresistance gene clusters closely associated with ceRNA network nodes. (A-C) Heatmap illustrating the multi-category gene clusters associated with the expression of key ceRNA
network components, including (A) PTBP1, (B) RP5-1092A11.5 and (C) AC017104.6.

Figure 7

Drugs with significant response difference in TP53 MT and WT patients were identified. (A) Drugs with different sensitivity in TP53 MT and WT group, shown by chord diagram, colored by group-specificity. The size of the arc represented the response difference significance. (B-D) Comparisons on IC50 values of
drugs with higher sensitivity in TP53 MT patients, including (B) AZD8055, (C) OSI027 and (D) RO3306. Values were compared using Wilcoxon rank sum test. *: p-value<=0.05, ****: p-value<=0.0001.

Figure 8

Multi-omics validation of PTBP1 expression in public databases. (A) Comparison on the RNA expression levels of PTBP1 using TP53 MT and normal samples from ICGC database. (B) Comparison on the protein
expression levels of PTBP1 using TP53 MT and TP53 WT samples from CTRP database. (C) Immunohistochemistry-derived protein expression of PTBP1 from normal and tumor samples in the HPA database. Expressions were compared using Wilcoxon rank sum test. ns: not significant, ****: p-value<=0.0001.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigureLegend.docx](#)
- [SupplementaryFigurePPT.pptx](#)