Bioinformatics Analysis Identifies the Estrogen Receptor 1 (ESR1) Gene and hsa-miR-26a-5p as Potential Prognostic Biomarkers in Patients with Intrahepatic Cholangiocarcinoma

Background: Intrahepatic cholangiocarcinoma arises from the epithelial cells of the bile ducts and is associated with poor prognosis. This study aimed to use bioinformatics analysis to identify molecular biomarkers of intrahepatic cholangiocarcinoma and their potential mechanisms.

Material/Methods: MicroRNA (miRNA) and mRNA microarrays from GSE53870 and GSE32879 were downloaded from the Gene Expression Omnibus (GEO) database. Differentially expressed miRNAs (DEMs) associated with prognosis were identified using limma software and Kaplan-Meier survival analysis. Predictive target genes of the DEMs were identified using miRWalk, miRTarBase, miRDB, and TargetScan databases of miRNA-binding sites and targets. Target genes underwent Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. Hub genes were analyzed by constructing the protein-protein interaction (PPI) network using Cytoscape. DEMs validated the hub genes, followed by construction of the miRNA-gene regulatory network.

Results: Twenty-five DEMs were identified. Fifteen DEMs were upregulated, and ten were down-regulated. Kaplan-Meier survival analysis identified seven upregulated DEMs and nine down-regulated DEMs that were associated with the overall survival (OS), and 130 target genes were selected. GO analysis showed that target genes were mainly enriched for metabolism and development processes. KEGG analysis showed that target genes were mainly enriched for cancer processes and some signaling pathways. Fourteen hub genes identified from the PPI network were associated with the regulation of cell proliferation. The overlap between hub genes and DEMs identified the estrogen receptor 1 (ESR1) gene and hsa-miR-26a-5p.

Conclusions: Bioinformatics analysis identified ESR1 and hsa-miR-26a-5p as potential prognostic biomarkers for intrahepatic cholangiocarcinoma.

MeSH Keywords: Biological Markers • Cholangiocarcinoma • Gene Expression Profiling • MicroRNAs

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/921815

* Xianzheng Qin and Yuning Song contributed equally to this work

Corresponding Authors: Xianzheng Qin, e-mail: jp6303416178@qmul.ac.uk, Yuning Song, e-mail: jp6303416179@qmul.ac.uk

Source of support: Departmental sources
Background

Intrahepatic cholangiocarcinoma arises from the epithelial cells of the bile ducts and is associated with poor prognosis. Worldwide, intrahepatic cholangiocarcinoma has an increasing incidence and high mortality rate and represents about 15% of cases of primary liver cancer, with hepatocellular carcinoma (HCC) representing about 70% of cases [1–3]. The main risk factors for intrahepatic cholangiocarcinoma include sclerosing cholangitis, biliary anomalies, hepatolithiasis, hepatobiliary flukes, and liver cirrhosis [4]. Patients with intrahepatic cholangiocarcinoma often present with nonspecific symptoms or are asymptomatic. Therefore, without sensitive screening criteria, only a few cases are diagnosed at an early stage [5,6]. Also, most patients are diagnosed with late-stage intrahepatic cholangiocarcinoma with the tumor having invaded into adjacent structures or metastasized to distant sites [7–9]. Even for patients who are diagnosed at an early stage, risk factors such as cirrhosis may increase the difficulty of treatment [6]. Only about 30% of patients with intrahepatic cholangiocarcinoma can undergo surgical resection, and these patients have a high recurrence rate following surgery [5,10]. Despite clinical research on improving the management of patients with intrahepatic cholangiocarcinoma, the prognosis remains poor, with a 30% three-year survival rate and an 18% five-year survival rate [11,12]. Therefore, potential diagnostic and prognostic biomarkers for intrahepatic cholangiocarcinoma remain to be identified.

The microRNAs (miRNAs) are a family of small endogenous non-coding RNA molecules that play an important role in regulating the expression of target genes and proteins through complementary base pairs with mRNAs [13–15]. Recent studies have shown an association between miRNAs and human cancers [16]. Changes in miRNAs affect several cellular processes that include cell proliferation, cell differentiation, and signal transduction [14,17,18]. The progression of intrahepatic cholangiocarcinoma is associated with the abnormal expression of miRNAs [18–20]. Biomarkers of intrahepatic cholangiocarcinoma have included upregulated miR-31, and miR-150 and down-regulated miR-590-3p and miR-424-5p [19–23]. Wang et al. [24] found that increased expression of plasma levels of miR-150 could identify patients with intrahepatic cholangiocarcinoma with high sensitivity, specificity [22,23]. Also, miR41 directly regulates BRCA1-associated protein-1 (BAP-1), which has frequent mutations in intrahepatic cholangiocarcinoma, which is associated with reduced prognosis [20,25,26].

Epithelial-mesenchymal transition (EMT) is a biological developmental process that is considered to be the key mechanism leading to invasion and metastasis of intrahepatic cholangiocarcinoma [27,28]. In 2015, Zhang et al. showed that the expression of miR-590-3p was down-regulated in intrahepatic cholangiocarcinoma and showed that miR-590-3p influenced EMT by inhibiting the expression of the Smad interacting protein 1 (SIP1) [29]. Also, miR-424-5p has been shown to play an important role in promoting cell proliferation and metastasis in intrahepatic cholangiocarcinoma [21,30]. In 2019, Wu et al. [21] proposed that the restoration of miR-424-5p expression may be a promising approach to treat intrahepatic cholangiocarcinoma by targeting the pathway of the binding between miR-424-5p and NUAK family kinase 1 (ARK5) mRNA. Although these previous studies have resulted in the development of drug treatments, the underlying molecular mechanisms in the progression of intrahepatic cholangiocarcinoma remain to be elucidated. Therefore, new diagnostic and prognostic biomarkers in patients with intrahepatic cholangiocarcinoma may also result in new approaches to treatment.

Bioinformatics analysis of microarray data is a high-throughput technology that has been widely used to identify genetic changes in cancer. The analysis of miRNA microarrays can be used to identify potential biomarkers in intrahepatic cholangiocarcinoma [29]. This study aimed to use bioinformatics analysis to identify molecular biomarkers of intrahepatic cholangiocarcinoma and their potential mechanisms. The miRNA and mRNA expression profiles were downloaded to obtain differentially expressed miRNAs (DEMs), and differentially expressed mRNAs. The interactions between DEMs, their target genes, and differentially expressed mRNAs in intrahepatic cholangiocarcinoma were investigated through the microarray profiles of the expression of miRNAs and mRNAs. The construction of the miRNA-gene regulatory network explored the potential molecular prognostic biomarkers for intrahepatic cholangiocarcinoma, which may provide insights into future diagnosis and treatment.

Material and Methods

Microarray data

High-throughput gene expression and microarray data were obtained from the Gene Expression Omnibus (GEO) public genomics online repository (www.ncbi.nlm.nih.gov/geo) [31]. The miRNA expression dataset, GSE53870, and the mRNA expression dataset, GSE32879, were downloaded from the GEO database [29,32]. The probes were converted to the corresponding gene symbol using the annotation information in the GEO platform.

Identification of differentially expressed miRNAs (DEMs) and differentially expressed mRNAs

The R (version 1.6.2) Affy package (www.bioconductor.org/) was used for the analysis of GSE53870 and GSE32879. The median
algorithm performed the data preprocessing and normalization in R (version 3.6.1). The limma package (version 3.40.6) (http://bioconductor.org/) was used to screen the DEMs and differentially expressed mRNAs. The adjusted P-value (adj. P-value) and the Benjamini–Hochberg false discovery rate (FDR) were used in the analysis to reduce the rate of false positives. DEMs and differentially expressed mRNAs, which both satisfied the log2 (fold-change) >2 and the adj. P-value <0.05 were considered to be statistically significant and were selected for further study.

Visualization of DEMs

The Heml Heat map Illustrator (version 1.0) is an open-source bioinformatics toolkit that was used to graphically visualize multi-dimensional and numerical gene expression data as heatmaps [33]. The data of DEMs were visualized with different colors. The volcano plot was performed using the R package ggplot2 version 3.2.1 to visualize the DEMs (https://cran.r-project.org/web/packages/ggplot2/index.html).

Kaplan-Meier survival analysis

Data in GSE53870 were processed for statistical analysis to investigate the relationship between DEMs and patients with intrahepatic cholangiocarcinoma. The free R package survival package (version 3.1-7) (https://cran.r-project.org/web/packages/survival/) was used for survival analysis of the screened DEMs. The log-rank test was performed to estimate the prognosis of different DEMs. A P<0.05 was considered to be statistically significant.

Prediction and screening of the target genes of DEMs

miRWalk (version 3.0) (http://mirwalk.umm.uni-heidelberg.de/) is an open-source website used to predict the target genes [34]. TargetScan (version 7.2) (http://www.targetscan.org/vert_72/) is an online database that was used to predict the target genes by searching for the conserved sites on the paired seed region of each DEM [35]. Also, miRDB (http://mirdb.org/index.html) and miRTarBase (http://miRTarBase.mbc.nctu.edu.tw/) were used to predict the target genes [36,37]. A Venn diagram was produced by the R (venneuler) package (version 1.1-0) (https://cran.r-project.org/web/packages/venneuler/index.html) to reduce false positives of data predicted by the online databases.

Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of target genes

The GO resource (http://geneontology.org/) is an online database that provides biological information like the function of genes and gene products [38,39]. The GO resource was used to implement functional enrichment analysis of significant target genes with P<0.05 [40]. The KEGG pathway enrichment analysis was performed using KEGG Orthology Based Annotation System (KOBAS) (version 3.0) (http://kobas.cbi.pku.edu.cn/anno_iden.php). KOBAS is a web server that can be used to identify significantly enriched pathways by mapping to genes with known annotations [41–43]. P<0.05 was considered as statistically significant.

Construction of the protein-protein interaction (PPI) network of the target genes and centrality analysis

The PPI network was established using the STRING online database (version 11.0) (https://string-db.org/), which aims to collect and integrate the interactions between proteins [44]. The PPI network was constructed to analyze the relationships between the screened target genes and the interaction. A combined score >0.400 was regarded as significant PPI node pairs worth further investigation. Cytoscape (version 3.7.1), which is an open source software that integrates biomolecular interaction networks, was used to visualize the data from the STRING [45,46]. The online CentiScape plugin (version 2.2) (http://apps.cytoscape.org/apps/mcode) in Cytoscape was used to calculate the centrality parameters to identify the most significant nodes in the PPI network [47].

Hub gene selection and analysis

The hub genes were screened by Cytoscape software, the Molecular Complex Detection (MCODE) (version 1.5.1) (http://apps.cytoscape.org/apps/mcode) plugin in Cytoscape was used for detection of the PPI networks with dense connectivity [48]. The selection criteria were the degree of cutoff=2, the node score cutoff=0.2, the K-score=2, and the maximum depth=100. The network degrees >10 were identified as hub genes.

Construction of the miRNA-gene regulatory network

The miRNAs were identified by the online databases that could predict one or more target genes of the DEMs. The miRNA and gene regulatory network was constructed using Cytoscape software to identify the relationship between the target genes and miRNAs.

Results

Identification and visualization of differentially expressed miRNAs (DEMs)

After the processing of the raw data in GSE53870 (Figure 1), a total of 1104 miRNAs were identified, 436 of which were up-regulated, and 668 were down-regulated. Compared with samples of normal intrahepatic bile ducts, patients with intrahepatic
Figure 1. Normalization of the GSE53870 data from the Gene Expression Omnibus (GEO) database. The black boxes represent the microRNA (miRNA) expression values in patients with intrahepatic cholangiocarcinoma. The red boxes represent the miRNA expression values in the normal intrahepatic bile duct control samples.

Figure 2. The heatmap and volcano plot of the 25 differentially expressed miRNAs (DEMs) associated with prognosis in intrahepatic cholangiocarcinoma. (A) Heatmap of the top 25 DEMs was constructed using HemI. The level of expression is positively correlated with the size of the fluorescence value. The red color indicates high expression. The green color indicates low expression. (B) The volcano plot shows DEMs between the samples from patients with intrahepatic cholangiocarcinoma and normal samples. The red color indicates statistical significance.
cholangiocarcinoma had 25 DEMs that satisfied log^2 (fold-change) > 2 and adj. P < 0.05, consisting of ten down-regulated miRNAs and 15 upregulated miRNAs (Figure 2). The top 25 DEMs are listed in Table 1.

Kaplan-Meier survival analysis

Based on the data in GSE53870, Kaplan-Meier survival analysis identified 16 DEMs that were associated with overall survival (OS), which included seven upregulated DEMs and nine down-regulated DEMs. In patients with intrahepatic cholangiocarcinoma, high expression of hsa-miR-1308 (P=4.59E-2), hsa-miR-566 (P=4.40E-2), hsa-miR-565 (P=4.53E-2), hsa-miR-3197 (P=1.98E-3), hsa-miR-4327 (P=1.72E-2), hsa-miR-513b (P=4.45E-2), hsa-miR-513c-5p (P=2.52E-2) and low expression of hsa-miR-145-5p (P=2.94E-2), hsa-miR-143-3p (P=1.46E-2), hsa-miR-451a (P=6.69E-3), hsa-miR-27b-3p (P=3.38E-3), hsa-miR-26a-5p (P=2.67E-2), hsa-miR-194-5p (P=2.53E-2), hsa-miR-195-5p (P=8.18E-3), hsa-miR-125b-5p (P=3.53E-2) and hsa-miR-29c-3p (P=1.19E-3) were significantly associated with reduced OS. The remaining DEMs were not significant survival biomarkers.

Table 1. The top 25 differentially expressed miRNAs (DEMs) in the intrahepatic cholangiocarcinoma samples compared with the normal bile ducts samples.

| miRNA ID     | log2 FC | B       | t     | P-value   | adj. P-value | Expression  |
|--------------|---------|---------|-------|-----------|--------------|-------------|
| hsa-miR-1975 | 3.91    | 16.1352 | 8.02  | 1.33E-11  | 5.66E-10     | Upregulated |
| hsa-miR-1974 | 3.849   | 12.9719 | 7.27  | 3.38E-10  | 8.47E-09     | Upregulated |
| hsa-miR-1826 | 3.844   | 14.2929 | 7.58  | 8.75E-11  | 2.84E-09     | Upregulated |
| hsa-miR-923  | 3.842   | 23.6002 | 9.78  | 6.54E-15  | 1.73E-12     | Upregulated |
| hsa-miR-1274b| 3.386   | 22.0409 | 9.41  | 3.21E-14  | 4.43E-12     | Upregulated |
| hsa-miR-1308 | 3.335   | 18.4515 | 8.56  | 1.25E-12  | 9.22E-11     | Upregulated |
| hsa-miR-566  | 2.749   | 24.9821 | 10.11 | 1.60E-15  | 8.81E-13     | Upregulated |
| hsa-miR-565  | 2.567   | 21.3078 | 9.24  | 6.79E-14  | 7.50E-12     | Upregulated |
| hsa-miR-3197 | 2.429   | 23.2466 | 9.7   | 9.39E-15  | 2.84E-12     | Upregulated |
| hsa-miR-1274a| 2.367   | 14.0459 | 7.61  | 7.79E-11  | 2.69E-09     | Upregulated |
| hsa-miR-4327 | 2.314   | 21.4746 | 9.28  | 5.73E-14  | 7.03E-12     | Upregulated |
| hsa-miR-513b | 2.256   | 11.5083 | 6.91  | 1.51E-09  | 2.92E-08     | Upregulated |
| hsa-miR-1978 | 2.242   | 12.4543 | 7.14  | 5.73E-10  | 1.38E-08     | Upregulated |
| hsa-miR-513c-5p| 2.232  | 8.9229  | 6.27  | 2.13E-08  | 3.52E-07     | Upregulated |
| hsa-miR-1977 | 2.078   | 6.98    | 5.8   | 1.57E-07  | 1.81E-06     | Upregulated |
| hsa-miR-145-5p| –3.527 | 23.9275 | –9.86 | 4.68E-15  | 1.72E-12     | Down-regulated |
| hsa-miR-143-3p| –3.272 | 15.8895 | –7.96 | 1.71E-11  | 7.00E-10     | Down-regulated |
| hsa-miR-27a-3p| –2.715 | 6.2213  | –5.61 | 3.44E-07  | 3.68E-06     | Down-regulated |
| hsa-miR-451a | –2.526  | 26.0606 | –10.37| 5.30E-16  | 5.86E-13     | Down-regulated |
| hsa-miR-27b-3p| –2.42  | 11.6593 | –6.95 | 1.29E-09  | 2.64E-08     | Down-regulated |
| hsa-miR-26a-5p| –2.287 | 4.0601  | –5.05 | 3.21E-06  | 2.81E-05     | Down-regulated |
| hsa-miR-194-5p| –2.189 | 10.9978 | –6.79 | 2.54E-09  | 4.68E-08     | Down-regulated |
| hsa-miR-195-5p| –2.094 | 23.2455 | –9.7 | 9.40E-15  | 1.73E-12     | Down-regulated |
| hsa-miR-125b-5p| –2.093 | 8.5621  | –6.2 | 3.09E-08  | 4.45E-07     | Down-regulated |
| hsa-miR-29c-3p| –2.072 | 19.7967 | –8.88 | 3.17E-13  | 2.50E-11     | Down-regulated |

log^2 FC = log^2 (fold-change); adj. P-value = adjusted P-value; B = B-value; t = t-statistics.
analysis of all screened miRNAs associated with OS are shown in Figures 3 and 4.

**Prediction and screening of target genes of DEMs associated with patient survival**

Different databases used in this study had their own algorithms to predict the target genes. After matching the overlap of the results of miRWalk between the online databases, TargetScan, miRDB, and miRTarBase, 130 target genes were predicted from eight DEMs. TargetScan identified 990 target genes, 1183 target genes were identified in miRDB, and 392 target genes were identified in miRTarBase. The overlap of target genes between the three datasets is shown in the Venn diagram (Figure 5).

**Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of target genes**

GO and KEGG functional and pathway enrichment analysis of 130 target genes was performed to understand the screened target genes better. The results of GO biological process (BP) analysis identified target genes that were significantly enriched for the processes of substance metabolism, development, and the regulation of gene expression. GO molecular function (MF) showed that target genes were mainly enriched in protein binding, cyclic compound binding, sequence-specific DNA binding, and transcription regulator activity. GO cellular component (CC) showed that target genes were mainly involved in the nucleus, cytosol, intracellular organelles, and membrane-enclosed lumen. Also, KEGG pathway analysis showed that target genes were significantly enriched in cancer processes, including pathways in cancer, miRNA in cancer, proteoglycans in cancer, Ras, FoxO, and PI3K-Akt signaling pathways, and resistance to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors. Pathways in cancers were the most significantly enriched (P=5.42E-13). The most significant findings from GO and KEGG enrichment analysis are shown in Figure 6 and Table 2.

**Construction of the protein-protein interaction (PPI) network and centrality analysis**

Based on the information of target genes from the STRING database, the PPI network, including the combined score >0.400, with 130 nodes and 231 edges (Figure 7), was constructed by Cytoscape. CenScape was used to calculate the value of the degree of centrality, which was used in the selection of hub genes.

Figure 3. Kaplan-Meier survival analysis of upregulated differentially expressed microRNAs (miRNAs) (DEMs). The red lines show individuals with high expression of DEMs. The green lines show individuals with low expression of DEMs.
Selection and analysis of the hub genes

Hub genes were identified by the Molecular Complex Detection (MCODE) plugin in Cytoscape, and a total of 14 genes were screened from 130 target genes (Figure 8). The results were sorted by degree scores and identified the following 14 genes: KRAS, ESR1, STAT3, VEGFA, IGF1R, SMAD2, FGF2, DICER1, ACTB, CDK6, MET, FOXO1, ETS1, and HBEGF (Table 3). These 14 hub genes were used to process the GO and KEGG enrichment analysis.

Figure 4. Kaplan-Meier survival analysis of down-regulated differentially expressed microRNAs (miRNAs) (DEMs). The red lines show individuals with high expression of DEMs. The green lines show individuals with low expression of DEMs.

Figure 5. Venn diagram of the four datasets of 130 target genes. The target genes identified from the miRWalk database were screened again using the miRWalk, TargetScan, miRDB, and miRTarBase databases. The four datasets show an overlap of 130 target genes.

Selection and analysis of the hub genes

Hub genes were identified by the Molecular Complex Detection (MCODE) plugin in Cytoscape, and a total of 14 genes were screened from 130 target genes (Figure 8). The results were sorted by degree scores and identified the following 14 genes: KRAS, ESR1, STAT3, VEGFA, IGF1R, SMAD2, FGF2, DICER1, ACTB, CDK6, MET, FOXO1, ETS1, and HBEGF (Table 3). These 14 hub genes were used to process the GO and KEGG enrichment analysis. GO biologic process (BP) and KEGG enrichment analysis...
showed that hub genes were primarily enriched for the regulation of cell proliferation, anatomical structure and tube morphogenesis, and some receptor protein signaling pathways, proteoglycans, and pathways in cancer (Table 4).

Construction of the miRNA and gene regulatory network

According to the data from the results of 130 predicted target genes and eight corresponding miRNAs, Cytoscape was used to construct miRNA and gene regulatory network, to identify the regulatory association between the miRNAs and hub genes. The relationships were visualized with the miRNA and gene regulatory network (Figure 9).

Validation of the hub genes using the Gene Expression Omnibus (GEO) mRNA expression dataset

Data in the GSE32879 dataset from GEO were analyzed to validate the identity of the hub genes found in this study. A total of 766 differentially expressed mRNAs were identified, 173 of which were upregulated, and 593 were down-regulated. The overlap between hub genes from GSE53870 and differentially expressed mRNAs from GSE32879 showed that ESR1 was the only gene that occurred in both GEO datasets. Finally, hsa-miR-26a-5p and the corresponding miRNA of ESR1 were identified in the miRNA and gene regulatory network.
Table 2. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of target genes for differentially expressed miRNAs (DEMs) in intrahepatic cholangiocarcinoma.

| Pathway ID | Pathway description                                      | Count | P-value    |
|------------|----------------------------------------------------------|-------|------------|
| GO: 0031233| regulation of cellular metabolic process                 | 78    | 5.34E-12   |
| GO: 0048523| negative regulation of cellular process                 | 64    | 1.33E-10   |
| GO: 0060255| regulation of macromolecule metabolic process           | 75    | 1.01E-10   |
| GO: 0019222| regulation of metabolic process                          | 79    | 1.23E-10   |
| GO: 0032502| developmental process                                    | 73    | 8.71E-11   |
| GO: 0048519| negative regulation of biological process               | 69    | 8.51E-11   |
| GO: 0051171| regulation of nitrogen compound metabolic process       | 73    | 7.74E-11   |
| GO: 0080090| regulation of primary metabolic process                  | 73    | 4.64E-10   |
| GO: 0048856| anatomical structure development                         | 68    | 8.44E-10   |
| GO: 0056222| intracellular                                           | 123   | 1.41E-10   |
| GO: 0060255| regulation of macromolecule metabolic process           | 75    | 1.01E-10   |
| GO: 0031323| regulation of cellular metabolic process                 | 78    | 5.34E-12   |
| GO: 0048523| negative regulation of cellular process                 | 64    | 1.33E-10   |
| GO: 0060255| regulation of macromolecule metabolic process           | 75    | 1.01E-10   |
| GO: 0019222| regulation of metabolic process                          | 79    | 1.23E-10   |
| GO: 0032502| developmental process                                    | 73    | 8.71E-11   |
| GO: 0048519| negative regulation of biological process               | 69    | 8.51E-11   |
| GO: 0051171| regulation of nitrogen compound metabolic process       | 73    | 7.74E-11   |
| GO: 0080090| regulation of primary metabolic process                  | 73    | 4.64E-10   |
| GO: 0048856| anatomical structure development                         | 68    | 8.44E-10   |
| GO: 0056222| intracellular                                           | 123   | 1.41E-10   |
| GO: 0060255| regulation of macromolecule metabolic process           | 75    | 1.01E-10   |
| GO: 0019222| regulation of metabolic process                          | 79    | 1.23E-10   |
| GO: 0032502| developmental process                                    | 73    | 8.71E-11   |
| GO: 0048519| negative regulation of biological process               | 69    | 8.51E-11   |
| GO: 0051171| regulation of nitrogen compound metabolic process       | 73    | 7.74E-11   |
| GO: 0080090| regulation of primary metabolic process                  | 73    | 4.64E-10   |
| GO: 0048856| anatomical structure development                         | 68    | 8.44E-10   |
| GO: 0056222| intracellular                                           | 123   | 1.41E-10   |
| GO: 0060255| regulation of macromolecule metabolic process           | 75    | 1.01E-10   |
| GO: 0019222| regulation of metabolic process                          | 79    | 1.23E-10   |
| GO: 0032502| developmental process                                    | 73    | 8.71E-11   |
| GO: 0048519| negative regulation of biological process               | 69    | 8.51E-11   |
| GO: 0051171| regulation of nitrogen compound metabolic process       | 73    | 7.74E-11   |
| GO: 0080090| regulation of primary metabolic process                  | 73    | 4.64E-10   |
| GO: 0048856| anatomical structure development                         | 68    | 8.44E-10   |
| GO: 0056222| intracellular                                           | 123   | 1.41E-10   |
Intrahepatic cholangiocarcinoma arises from bile duct epithelial cells and has high morbidity and mortality [10,22,49]. Due to the lack of effective methods for early diagnosis, the majority of patients with intrahepatic cholangiocarcinoma do not have symptoms in the early stages, and present with late-stage disease. Despite clinical studies to improve patient management, the molecular mechanisms remain unclear, and there are no prognostic molecular biomarkers. Therefore, the identification of molecular biomarkers associated with intrahepatic cholangiocarcinoma, their biological significance, and biological functions may provide insight into the pathogenesis of intrahepatic cholangiocarcinoma at the molecular level.

In the present study, microRNA (miRNA) and mRNA microarrays from GSE53870 and GSE32879 were downloaded from the Gene Expression Omnibus (GEO) database for intrahepatic cholangiocarcinoma and were used to identify differentially expressed miRNAs (DEMs) in comparison with normal intrahepatic bile ducts. This study identified 25 DEMs from the dataset, including 15 upregulated miRNAs and ten down-regulated miRNAs. Kaplan-Meier survival analysis showed that seven upregulated miRNAs (hsa-miR-1308, hsa-miR-566, hsa-miR-565, hsa-miR-3197, hsa-miR-4327, hsa-miR-513b, and hsa-miR-513c-5p) and nine down-regulated DEMs (hsa-miR-145-5p, hsa-miR-143-3p, hsa-miR-451a, hsa-miR-27b-3p, hsa-miR-26a-5p, hsa-miR-194-5p, hsa-miR-195-5p, hsa-miR-125b-5p, and hsa-miR-29c-3p) were associated with the overall survival (OS) of patients with intrahepatic cholangiocarcinoma. The associations between some of the identified DEMs and intrahepatic cholangiocarcinoma have also been identified in previous studies. Specifically, miR-145 has been reported as a tumor suppressor, and the levels are reduced in intrahepatic cholangiocarcinoma, which affects Akt/FoxO1
### Table 3. The top 14 hub genes with the degree score.

| Gene symbol | Gene description                          | Degree |
|-------------|-------------------------------------------|--------|
| KRAS        | KRAS proto-oncogene, GTPase               | 30     |
| ESR1        | Estrogen receptor 1                       | 26     |
| STAT3       | Signal transducer and activator of tran 3 | 24     |
| VEGFA       | Vascular endothelial growth factor A      | 21     |
| IGF1R       | Insulin like growth factor 1 receptor     | 16     |
| SMAD2       | Smad family member 2                      | 15     |
| FGFR2       | Fibroblast growth factor 2                | 15     |
| ACTB        | Actin beta                                | 15     |
| CDK6        | Cyclin dependent kinase 6                 | 15     |
| MET         | MET proto-oncogene, receptor tyrosine kinase | 13     |
| FOXO1       | Forkhead box 01                           | 12     |
| FOS1        | FOS proto-oncogene 1, transcription factor| 9      |
| HBEGF       | Heparin binding EGF like growth factor    | 7      |

### Table 4. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) biologic process pathway enrichment analysis of the hub genes.

| Pathway ID  | Pathway description                                         | Count | P-value    |
|-------------|-------------------------------------------------------------|-------|------------|
| GO: 0042127 | regulation of cell population proliferation                | 12    | 3.90E-12   |
| GO: 0009653 | anatomical structure morphogenesis                          | 12    | 1.04E-10   |
| GO: 0001676 | enzyme linked receptor protein signaling pathway            | 9     | 1.20E-10   |
| GO: 007169  | transmembrane receptor protein tyrosine kinase signaling pathway | 8  | 3.74E-10   |
| GO: 0010604 | positive regulation of macromolecule metabolic process      | 13    | 5.82E-10   |
| GO: 0008284 | positive regulation of cell population proliferation        | 9     | 1.04E-09   |
| GO: 0009893 | positive regulation of metabolic process (GO: 0009893)      | 13    | 1.59E-09   |
| GO: 0010628 | positive regulation of gene expression                      | 11    | 1.77E-09   |
| GO: 0035239 | tube morphogenesis                                          | 8     | 2.51E-09   |
| GO: 0080134 | regulation of response to stress                            | 10    | 2.88E-09   |
| hsa05205   | Proteoglycans in cancer                                     | 10    | 1.69E-20   |
| hsa05200   | Pathways in cancer                                          | 10    | 1.09E-17   |
| hsa01521   | EGFR tyrosine kinase inhibitor resistance                   | 6     | 2.72E-13   |
| hsa05212   | Pancreatic cancer                                           | 5     | 3.10E-11   |
| hsa05218   | Melanoma                                                    | 5     | 4.40E-11   |
| hsa04015   | Rap1 signaling pathway                                      | 6     | 7.12E-11   |
| hsa04014   | Ras signaling pathway                                       | 6     | 1.12E-10   |
| hsa04933   | AGE-RAGE signaling pathway in diabetic complications         | 5     | 2.40E-10   |
| hsa05206   | MicroRNAs in cancer                                         | 6     | 5.52E-10   |
| hsa04068   | FoxO signaling pathway                                      | 5     | 9.46E-10   |
Increased expression of miR-145 is associated with inhibition of the growth of intrahepatic cholangiocarcinoma by inhibiting cancer cell proliferation, growth, and invasion [53,54]. Also, miR-26a was previously shown to be significantly down-regulated in cholangiocarcinoma cells in vitro, and miR-195 expression was reduced in cholangiocarcinoma cells [50–52]. A miRNA expression profile in intrahepatic cholangiocarcinoma previously reported the aberrant expression of some miRNAs, which included upregulated hsa-miR-566, while hsa-miR-29c-3p, hsa-miR-26a-5p, hsa-miR-451a, and hsa-miR-143-3p were down-regulated, which supports the findings of the DEMs identified in the present study [53,55]. However, miRNAs have different functional roles in the regulation of specific genes [56]. Therefore, target gene prediction of miRNAs is of importance. Several online databases are currently used to predict target genes of miRNAs, and each miRNA may predict a large number of target genes with the help of the algorithms from online databases. However, many gene target databases do not fully understand the relationships between miRNAs and target genes, which may result in false positives [56]. The miRWalk database predicts target genes by integrating six conventional features and seven new features [34,57]. TargetScan considers site type and searches for the conserved sites that pair the seed region of each DEM and then considers another 14 features to predict the target genes [35]. By using the support vector machine framework, miRDB may be used to predict target genes [36,58]. The online database, miRTarBase, predicts target genes by collecting and organizing the relationship between miRNAs and target genes from published studies [37]. Based on different computational methods of the online databases miRWalk, TargetScan, miRDB, and miRTarBase, the overlap of target genes in all datasets may reduce the false positives of the predicted results of miRWalk and make the identification of target genes more credible, as in the present study.

In this study, there were 130 genes selected. Gene Ontology (GO) functional enrichment analysis showed that these 130 genes were significantly enriched in the substance metabolism, development process, and regulation of gene expression. There are several previous studies have shown that regulation of cell proliferation and cellular metabolic processes are associated with cancer progression [59–61]. The findings from the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis identified target genes that were enriched for resistance to epidermal growth factor (EGFR) tyrosine kinase inhibitors, hormone resistance, and other signaling pathways, including Ras and FoxO, and cancer development pathways. Several previously published studies have shown that miRNAs are associated with cancer [62–64]. Kim et al. identified the role of signal transduction in cancer, which is consistent with the findings from the present study [65]. Tyrosine kinase inhibitors specific for EGFR1 could affect the role of EGFR and contribute to the progression of cholangiocarcinoma, as shown in a previous study by Lee et al. [66], which is consistent with the finding of EGFR tyrosine kinase inhibitor resistance identified in the present study. Rizvi et al. [22] showed that the Ras pathway was involved in malignancy. Other studies have shown that activation of the FoxO1 signaling pathway...
and the genetic variation of ESR1 might increase the risk for hepatocellular carcinoma and prostate cancer [83,84]. Given these associations, ESR1 could be regarded as a potential tumor suppressor for intrahepatic cholangiocarcinoma. The results of KEGG pathway enrichment analysis showed that ESR1 was mainly enriched in proteoglycans in cancer. Proteoglycans are components of the extracellular matrix (ECM), which has been shown in tumorigenesis of leiomyomas by VCAN by down-regulating ESR1 [85]. Also, changes in ECM are associated with the development of hepatocellular carcinoma (HCC) and liver cirrhosis [86,87]. The roles of proteoglycan and ESR1 in intrahepatic cholangiocarcinoma require further study, as liver cirrhosis is also a risk factor for intrahepatic cholangiocarcinoma.

Several previous studies have reported that hsa-miR-26a-5p acts as a tumor suppressor and in cancer [88]. The expression of hsa-miR-26a-5p was reduced in bladder cancer, colorectal cancer, and HCC [89–91]. In this study, the result also demonstrated hsa-miR-26a-5p were down-regulated. Chang et al. showed that patients with HCC who had increased expression of hsa-miR-26a-5p had an increase in overall survival (OS) rates and a reduced the risk of tumor recurrence [92]. Also, hsa-miR-26a-5p was shown to be associated with the expression of E-cadherin and vimentin, which are involved in epithelial-mesenchymal transition (EMT) [93]. EMT has been identified as an important factor in tumor metastasis in intrahepatic cholangiocarcinoma [27,28]. By targeting EMT, hsa-miR-26a-5p might interfere with tumor development to improve the prognosis of patients. Therefore, hsa-miR-26a-5p should be regarded as a potential molecular biomarker in patients with intrahepatic cholangiocarcinoma.

Conclusions

This study aimed to use bioinformatics analysis to identify molecular biomarkers of intrahepatic cholangiocarcinoma and their potential mechanisms and identified down-regulated hsa-miR-26a-5p and ESR1. The findings from the present study, combined with the findings from previous studies, support the importance of hsa-miR-145-5p, KRAS, and hsa-miR-143-3p in intrahepatic cholangiocarcinoma. Further clinical studies are required to verify the findings from this preliminary study.

Conflict of interests

None.
References:

1. Rahnenfai-Azar AA, Weisbord A, Dillhoff M et al: Intrahepatic cholangiocarcinoma: Molecular markers for diagnosis and prognosis. Surg Oncol, 2017; 26(2): 125–37

2. Altekruse SF, Devesa SS, Dickie LA et al: Histological classification of liver and intrahepatic bile duct cancers in SEER registries. J Registry Manag, 2011; 18(4): 201–5

3. Global Burden of Disease Cancer Collaboration, Fitzmaurice C, Dicker D, Pain A et al: The Global Burden of Cancer 2013. JAMA Oncol, 2015; 1(4): 505–27

4. Zhang H, Yang T, Wu M, Shen F: Intrahepatic cholangiocarcinoma: Epidemiology, risk factors, diagnosis and surgical management. Cancer Lett, 2016; 379(2): 198–205

5. Chun YS, Javle M: Systemic and adjuvant therapies for intrahepatic cholangiocarcinoma. Cancer Control, 2017; 24(3): 1073274817729241

6. Massarweh NN, El-Serag HB: Epidemiology of hepatocellular carcinoma and intrahepatic cholangiocarcinoma. Cancer Control, 2017; 24(3): 1073274817729245

7. Banales JM, Cardinale V, Carpio G et al: Expert consensus document: Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). Nat Rev Gastroenterol Hepatol, 2016; 13(5): 261–80

8. Weber SM, Ribero D, O’Reilly EM et al: Intrahepatic cholangiocarcinoma: current management and emerging therapies. Expert Rev Gastroenterol Hepatol, 2017; 11(5): 439–49

9. Bichacz B, Komuta M, Roskams T, Gores GJ: Clinical diagnosis and staging of cholangiocarcinoma. Nat Rev Gastroenterol Hepatol, 2011; 8(9): 512–22

10. Rahnenfai-Azar AA, Weisbord AB, Dillhoff M et al: Intrahepatic cholangiocarcinoma: Current management and emerging therapies. Expert Rev Gastroenterol Hepatol, 2017; 11(5): 439–49

11. Bartella I, Dufour JP: Clinical diagnosis and staging of intrahepatic cholangiocarcinoma. J Gastrointestin Liver Dis, 2015; 24(4): 481–89

12. Khan SA, Davidson BR, Goldin R et al: Guidelines for the diagnosis and treatment of cholangiocarcinoma: Consensus document. Gut, 2002; 51(3 Suppl 6): VII–9

13. Lou W, Liu J, Gao Y et al: MicroRNA regulation of liver cancer stem cells. Am J Cancer Res, 2018; 8(7): 1126–41

14. Li J, Tan S, Kooger R et al: MicroRNAs as novel biological targets for detection and regulation. Chem Soc Rev, 2014; 43(2): 506–17

15. Lou W, Ding B, Xu L, Fan W: Construction of potential glioblastoma multi-plex gene set analysis tools for better prioritization of biologically relevant sites in mammalian miRNAs. Elife, 2015; 4

16. Sticht C, De La Torre C, Parveen A, Grelt N: miRwalk: An online resource for prediction of microRNA binding sites. PLoS One, 2018; 13(10): e0206239

17. Zhang RX, Zheng Z, Li K et al: Both plasma and tumor tissue miR-146a high expression correlates with prolonged overall survival of surgical patients with intrahepatic cholangiocarcinoma. Medicine 2017; 96(44): e8267

18. Haga H, Patel T: Molecular diagnosis of intrahepatic cholangiocarcinoma. J Hepatobiliary Pancreat Sci, 2015; 22(2): 114–23

19. Yu C, Liu S, Cao W et al: miR-390-3p suppresses epithelial-mesenchymal transition in intrahepatic cholangiocarcinoma by inhibiting SIP1 expression. Oncotarget, 2017; 8(21): 34698–708

20. Funahashi M, Shigeri M et al: Prognostic role of microRNA-122 expression in animal model of hepatocellular carcinoma. J Cancer Res, 2002; 86(6): 442–54

21. Doncheva NT, Morris JH, Goodkin I, Jensen LJ: miRBase StringApp: Network analysis and visualization of microRNAs data. J Proteome Res, 2019; 18(2): 623–32

22. Szklarczyk D, Gable AL, Lyon D et al: STRING v11: Protein-protein interaction networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res, 2019; 47(D1): D419–26

23. Ashburner M, Ball CA, Blake JA et al: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet, 2000; 25(1): 25–29

24. The Gene Ontology Consortium: The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res, 2019; 47(D1): D330–38

25. Mi H, Muruganjanu A, Ebert D et al: PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res, 2019; 47(D1): D419–26

26. Dörnhaus NT, Morris JH, Goodkin I, Jensen LJ: miRBase StringApp: Network analysis and visualization of microRNAs data. J Proteome Res, 2019; 18(2): 623–32

27. Scardoni G, Tosadori G, Faizan M et al: Biological network analysis with STRING v11: Protein-protein interaction networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res, 2019; 47(D1): D607–13

28. Shannon P, Markiel A, Ozier O et al: Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res, 2003; 13(11): 2498–504

29. Szklarczyk D, Gable AL, Lyon D et al: STRING v11: Protein-protein interaction networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res, 2019; 47(D1): D607–13

30. Doncheva NT, Morris JH, Goodkin I, Jensen LJ: miRBase StringApp: Network analysis and visualization of microRNAs data. J Proteome Res, 2019; 18(2): 623–32

31. Ashburner M, Ball CA, Blake JA et al: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet, 2000; 25(1): 25–29

32. The Gene Ontology Consortium: The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res, 2019; 47(D1): D330–38

33. Scardoni G, Tosadori G, Faizan M et al: Biological network analysis with STRING v11: Protein-protein interaction networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res, 2019; 47(D1): D607–13

34. Dörnhaus NT, Morris JH, Goodkin I, Jensen LJ: miRBase StringApp: Network analysis and visualization of microRNAs data. J Proteome Res, 2019; 18(2): 623–32

35. Ashburner M, Ball CA, Blake JA et al: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet, 2000; 25(1): 25–29

36. The Gene Ontology Consortium: The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res, 2019; 47(D1): D330–38

37. Scardoni G, Tosadori G, Faizan M et al: Biological network analysis with STRING v11: Protein-protein interaction networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res, 2019; 47(D1): D607–13

38. Dörnhaus NT, Morris JH, Goodkin I, Jensen LJ: miRBase StringApp: Network analysis and visualization of microRNAs data. J Proteome Res, 2019; 18(2): 623–32

39. Ashburner M, Ball CA, Blake JA et al: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet, 2000; 25(1): 25–29

40. The Gene Ontology Consortium: The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res, 2019; 47(D1): D330–38

41. Scardoni G, Tosadori G, Faizan M et al: Biological network analysis with STRING v11: Protein-protein interaction networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res, 2019; 47(D1): D607–13

42. Dörnhaus NT, Morris JH, Goodkin I, Jensen LJ: miRBase StringApp: Network analysis and visualization of microRNAs data. J Proteome Res, 2019; 18(2): 623–32

43. Ashburner M, Ball CA, Blake JA et al: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet, 2000; 25(1): 25–29
