



## OPEN ACCESS

## EDITED BY

Justin Boddey,  
The University of Melbourne, Australia

## REVIEWED BY

Serena Cavallero,  
Sapienza University of Rome, Italy  
Fernanda Janku Cabral,  
State University of Campinas, Brazil

## \*CORRESPONDENCE

Ying Wang  
wangying@nipd.chinacdc.cn  
Jianping Cao  
caojp@chinacdc.cn

<sup>†</sup>These authors have contributed  
equally to this work

## SPECIALTY SECTION

This article was submitted to  
Parasite and Host,  
a section of the journal  
Frontiers in Cellular and  
Infection Microbiology

RECEIVED 05 April 2022

ACCEPTED 28 June 2022

PUBLISHED 11 August 2022

## CITATION

Liu L, Chen F, Jiang S, Zhong B, Li W,  
Xu K, Wang Q, Wang Y and Cao J  
(2022) Analysis of gene expression  
profile of peripheral blood in alveolar  
and cystic echinococcosis.  
*Front. Cell. Infect. Microbiol.* 12:913393.  
doi: 10.3389/fcimb.2022.913393

## COPYRIGHT

© 2022 Liu, Chen, Jiang, Zhong, Li, Xu,  
Wang, Wang and Cao. This is an open-  
access article distributed under the  
terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use,  
distribution or reproduction is  
permitted which does not comply with  
these terms.

# Analysis of gene expression profile of peripheral blood in alveolar and cystic echinococcosis

Lei Liu<sup>1,2†</sup>, Fan Chen<sup>3†</sup>, Shan Jiang<sup>4†</sup>, Bo Zhong<sup>5</sup>, Wei Li<sup>6</sup>,  
Kejun Xu<sup>6</sup>, Qi Wang<sup>5</sup>, Ying Wang<sup>2\*</sup> and Jianping Cao<sup>2,7\*</sup>

<sup>1</sup>Department of Hospital Infection Management, Sichuan Provincial Orthopedic Hospital, Chengdu, China, <sup>2</sup>National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, (Chinese Center for Tropical Diseases Research); Key Laboratory of Parasite and Vector Biology, National Health Commission of People's Republic of China; National Center for International Research on Tropical Diseases, China; World Health Organization Collaborating Center for Tropical Diseases, Shanghai, China, <sup>3</sup>Department of Microbiological Laboratory, Xindu County Center for Disease Control and Prevention, Chengdu, China, <sup>4</sup>Department of Department of Environmental and School Health, Chengdu Center for Disease Control and Prevention, Chengdu, China, <sup>5</sup>Institute of Parasitic Diseases, Sichuan Center for Disease Control and Prevention, Chengdu, China, <sup>6</sup>Department of Parasitic Diseases, Garzé Center for Disease Control and Prevention, Kangding, China, <sup>7</sup>The School of Global Health, Chinese Center for Tropical Diseases Research, Shanghai Jiao Tong University School of Medicine, Shanghai, China

RNA-sequencing (RNA-seq) is a versatile, high-throughput technology that is being widely employed for screening differentially expressed genes (DEGs) in various diseases. Echinococcosis, a globally distributed zoonosis, has been reported to impose a heavy disease burden in pastoral areas of China. Herein we aimed to explore the molecular mechanisms underlying echinococcosis. In this study, peripheral blood samples were collected from six patients with alveolar echinococcosis (AE), six patients with cystic echinococcosis (CE), and six healthy controls. RNA-Seq (mRNA) was performed to detect gene transcript and expression levels, and DEGs were subjected to bioinformatic analyses. In comparison with healthy controls, 492 DEGs (270 upregulated, 222 downregulated) were found in the AE group and 424 DEGs (170 upregulated, 254 downregulated) were found in the CE group ( $|\log_2(\text{fold change})| > 1$  and  $P < 0.05$ ). Further, 60 genes were upregulated and 39 were downregulated in both the AE and CE groups. Gene ontology enrichment analysis indicated that DEGs were mainly involved in molecular functions, including extracellular space, extracellular region, organ and system development, and anatomical structure development. Protein-protein interaction (PPI) networks were constructed to depict the complex relationship between DEGs and interacting proteins.

## KEYWORDS

echinococcosis, peripheral blood, RNA-sequencing, mRNA, DEG

## Introduction

Echinococcosis is one of the most geographically widespread parasitic zoonoses across the world. *Via* peroral infection by the eggs of *Echinococcus* spp., the parasitic metacestode (larval stage) develops as a tumor-like tissue in the liver of intermediate hosts, which include livestock, small rodents, and even humans. In humans, *E. multilocularis* and *E. granulosus* cause alveolar echinococcosis (AE) and cystic echinococcosis (CE), respectively (Kammerer and Schantz, 1993; WHO, 2011; Deplazes et al., 2017). In endemic regions, echinococcosis is a significant public health issue, imposing financial burden on patients due to high treatment and hospitalization costs (Craig et al., 2019; Wang et al., 2020). Therefore, many earlier studies on echinococcosis have focused on exploring disease pathogenesis, diagnostics, and epidemiology, aiding the development of treatment and prevention strategies (Moro and Schantz, 2006; Nunnari et al., 2012; Fu et al., 2020).

At present, echinococcosis diagnosis primarily depends on clinical symptoms, epidemiological contact history and imaging examinations. However, the sensitivity of imaging examinations to detect small lesions at the early stage is not particularly high. Serological antibody detection tests could be used to overcome this gap, but they lack species specificity due to the possibility of cross-reaction with other helminthic antigens (Schweiger et al., 2012).

Surgery is the most commonly used method to treat echinococcosis, but recurrence is an unavoidable issue. Albendazole is the preferred drug recommended by the World Health Organization, which selectively inhibits glucose uptake, leading to glycogen storage depletion in the parasite. However, the efficacy of albendazole is not yet well documented, and it can also cause side effects, such as dizziness, nausea, and leucopenia (Dehkordi et al., 2019). Immunoprophylaxis is a potential satisfactory way to prevent the prevalence of echinococcosis. Several antigens have been evaluated as vaccine candidates, but to date, they remain in the animal experiment stage (Jazouli et al., 2020; Gharibi et al., 2021).

In recent years, with the genome mapping of *Echinococcus* as well as of intermediate and terminal hosts, studies have moved their focus to the genetic level, especially research into gene expression and regulation. Rapid advancements are being made in gene-related technologies, and RNA-sequencing (RNA-seq) based on second generation high-throughput sequencing platforms is receiving a great deal of attention. Transcriptomics, which is widely employed for analyzing gene expression, splicing sites, allele-specific expression, and isoforms, has become the mainstream method for gene annotation and expression determination.

In this study, we collected peripheral blood samples from patients with AE, CE and healthy controls, and assessed gene transcripts and expression levels using RNA-Seq technology. The differential expression levels of mRNA between patients and

controls were analyzed with bioinformatics. Our findings provide new ideas and clues for echinococcosis diagnosis and treatment; further, we believe they could also support research into immunity.

## Materials and methods

### Study participants

The present study was approved by the Ethics Committee of Sichuan Centre for Disease Control and Prevention (reference no. 2018–2), and all methods were performed in accordance with the relevant guidelines and regulations of the Declaration of Helsinki. All participants were Tibetan residents living in Garzê County, Garzê Tibetan Autonomous Prefecture; their signed written informed consent was obtained before the study. Peripheral blood samples were collected from six patients with AE, six patients with CE, and six healthy controls; all the three groups were included with the same half-and-half sex ratio. None of the study participants showed history of cardiovascular and cerebrovascular diseases, metabolic diseases, tumors, or systemic autoimmune diseases.

### Blood sample collection and RNA extraction

From each participant, 4 mL venous blood were drawn into an ethylenediaminetetraacetic acid vacutainer tube and diluted with an equal volume of sterile phosphate-buffered saline. Total RNA was isolated using TRIzol; the TRIzol: blood ratio was 3:1. RNA quality was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

### RNA preparation, library construction, and illumina sequencing

A total amount of 3 µg RNA per sample was used as input material for RNA sample preparations. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA), according to manufacturer instructions, and index codes were added to attribute sequences to each sample. mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was achieved using divalent cations under elevated temperature in NEBNext® First Strand Synthesis Reaction Buffer (5×). First-strand cDNA was synthesized using random hexamer primers and M-MuLV reverse transcriptase (RNase H-), and second-strand cDNA was synthesized using DNA polymerase I and RNase H. cDNA fragments of

preferentially 250–300 bp in length were selected for PCR, and library quality was assessed on an Agilent Bioanalyzer 2100. The libraries were sequenced on an Illumina HiSeq platform and 125 bp/150 bp paired-end reads were generated. Raw data were cleaned with Cutadapt 4.0 to remove reads that were of low quality and those containing adapter and poly-N.

## Read mapping and gene expression quantification

Reference genome and gene model annotation files were downloaded from the genome website Ensembl (<http://www.ensembl.org/index.html>). An index of the reference genome was built using Bowtie v2.2.3, and paired-end clean reads were aligned to the reference genome using TopHat v2.0.12. HTSeq v0.6.1 was used to count the number of reads mapped to each gene. Subsequently, the expected number of fragments per kilobase of transcript sequence per millions base pairs (FPKM) was calculated for estimating gene expression levels (Anders et al., 2015).

## Differential expression analysis

Differential expression analysis was performed using the DESeq R package (1.18.0), which provided statistical methods for determining differential expression in digital gene expression data using a model based on negative binomial distribution. Differentially expressed genes (DEGs) were identified by comparing the expression levels of all transcripts in the AE and CE groups with those in the control group. Genes with  $|\log_2(\text{fold change})| > 1$  and an adjusted P-value  $< 0.05$  were considered to be differentially expressed (Anders and Huber, 2010; Anders and Huber, 2016).

## Enrichment analysis

Gene ontology (GO) enrichment analyses of DEGs were performed using the Goseq R package, in which gene length bias was corrected. GO terms with corrected  $P < 0.05$  were considered significantly enriched by DEGs (Young et al., 2010).

ClusterProfiler R package was used to assess the statistical enrichment of DEGs in KEGG pathways and to analyze their involvement in pertinent metabolic pathways.

## Protein–protein interaction network

PPI networks were constructed using Cytoscape v3.7.2, and Search Tool for the Retrieval of Interacting Genes (STRING) database was employed to depict the complex relationship

between DEGs and proteins encoded by them (Szkłarczyk et al., 2015).

## Verification of the expression of DEGs by quantitative RT-PCR

To verify the accuracy of our sequencing results, qRT-PCR was performed to investigate the expression levels of several DEGs. The mRNA was reverse-transcribed into cDNA using the Revert Aid First-Strand cDNA Synthesis Kit (Thermo, USA). The real-time PCR was performed using SYBR Premix Ex Taq II (Takara, Japan). GAPDH served as the internal reference gene, and primers were designed using Primer Premier 5. Each PCR was repeated in triplicate. The expression levels of genes were normalized by 18S rRNA and analyzed using the  $2^{-\Delta\Delta C_t}$  method.

## Results

### Differential expression analysis

A total of 1.02 billion raw reads were obtained, and the average number of reads per sample was 56,609,372. The error rates of 18 specimens were all less than 0.03%. Bases with  $Q_{\text{phred}} \geq 20$  accounted for 96.85% of the total number. Correlation analyses showed that the average global profiles of gene expressions between AE, CE, and control samples were highly correlated (correlation coefficient  $R^2_{\text{AE}} = 0.995$ ,  $R^2_{\text{CE}} = 0.994$ , Figure 1).

In comparison with the control group, 492 (270 upregulated, 222 downregulated) and 424 (170 upregulated, 254 downregulated) DEGs were identified in the AE and CE groups, respectively (Figure 2). Volcano plots were drawn using  $\log_2(\text{fold change})$  as the abscissa and  $-\log_{10}(t\text{-test significance P-value})$  as the ordinate. Hierarchical cluster analysis was performed to assess DEG expression patterns, which became evident *via* a heatmap (Figure 3).

The top 20 downregulated and upregulated DEGs in the AE and CE groups are shown in Figure 4. The expression levels of some genes were found to be upregulated (e.g., RAP1GAP, RPS15AP29, and AL391421.1) and downregulated (e.g., FOLR1, AC093591.1, and AARD) in patients with AE. Similarly, the expression levels of many genes were upregulated (e.g., SLC25A24P1, IHH, and AL590096.1) and downregulated (e.g., SFTPC, AL645608.2, and SFTPA2) in patients with CE.

AE and CE are pathologically different but also show some similarities. The expression of some genes was up- or downregulated in both AE and CE (Figure 5). The expression levels of 60 genes (e.g., ASMT, and TMEM244) were upregulated, and those of 39 genes (e.g., FOLR1, COL3A1, and MAGED4) were downregulated in both these groups.

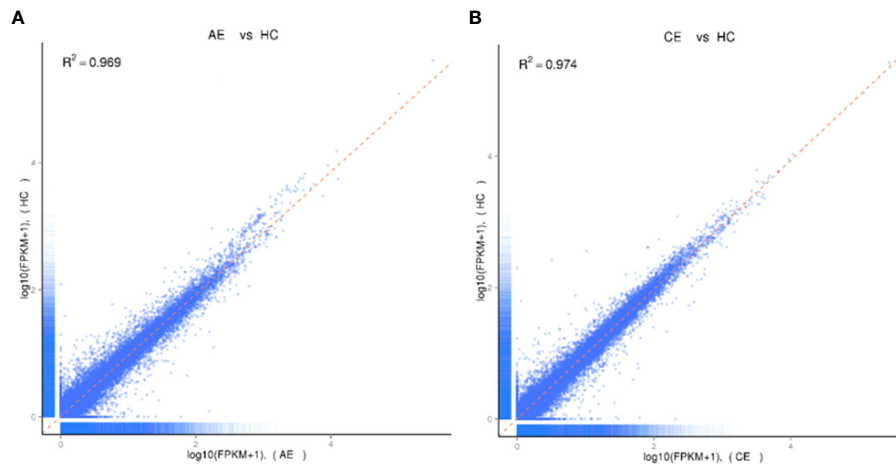


FIGURE 1

Analysis of homogeneity between AE, CE, and healthy control samples. (A) Expression analysis of AE and healthy control samples. The Pearson correlation coefficient was 0.969. (B) Expression analysis of CE and healthy control samples. The Pearson correlation coefficient was 0.974.

## GO enrichment analysis

In the AE group, 19 clusters ( $P < 0.05$ ) were significantly annotated: 7, 10, and 2 were associated with the functional categories of biological process, cellular component, and molecular function, respectively (Figure 6A). The top enrichments were closely related to extracellular space, extracellular region, organ and system development, and response to external stimulus.

In the CE group, 30 clusters ( $P < 0.05$ ) were significantly annotated: 15, 10, and 5 were associated with the functional

categories of biological process, cellular component, and molecular function, respectively (Figure 6B). The top upregulated enrichments were closely related to extracellular space, extracellular region, anatomical structure development, multicellular organismal development, and system development.

## KEGG pathway enrichment analysis

In the AE group, DEGs were mainly involved in 11 pathways ( $P < 0.05$ ), including ribosome, protein digestion and absorption,

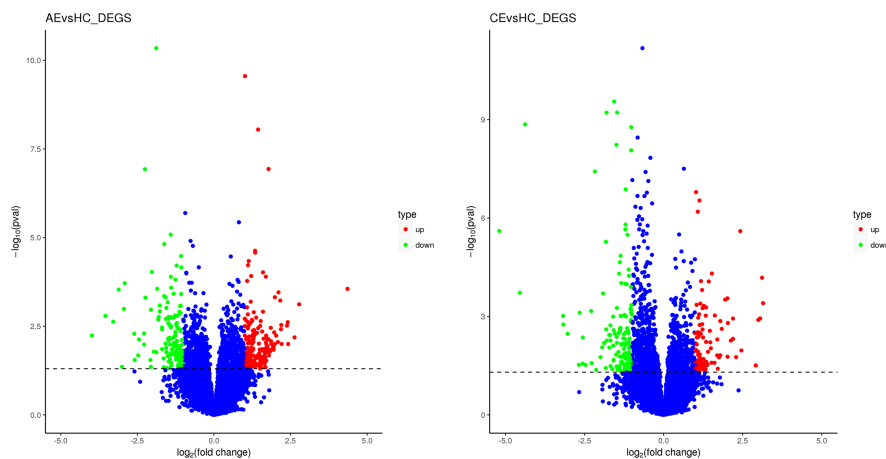


FIGURE 2

Scatterplot of DEGs. (A) The expression levels of 270 and 222 genes were up- and downregulated, respectively, in patients with AE. (B) The expression levels of 170 and 254 were up- and downregulated, respectively, in patients with CE.

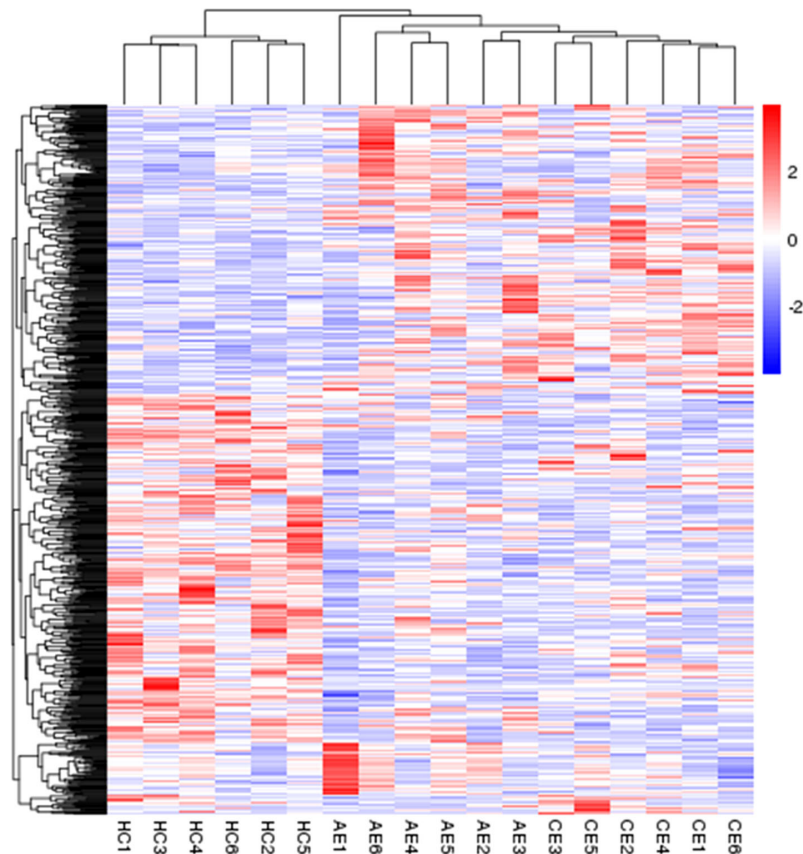


FIGURE 3

Heatmap of DEGs. The heatmap was generated using normalized FPKM values of DEGs. The expression level of each transcript is represented by a color, ranging from blue (low) to red (high).

and ECM-receptor interaction (Figure 7A). In the CE group, DEGs were mainly involved in 16 pathways ( $P < 0.05$ ), including phagosome and protein digestion and absorption (Figure 7B).

## PPI network analysis

Proteins encoded by DEGs in the AE and CE groups formed a complex PPI network (Figure 8A, B). The PPI network of protein products in the AE group was mainly linked by hub genes such as GBP1, ISG15, RSAD2, CXCL10, and VCAM-1, while that of those in the CE group was mainly linked by hub genes such as COL1A1, COL1A2, COL3A1, DCN, and FBLN1.

## Verification of the expression of DEGs by qRT-PCR

qRT-PCR was performed to validate the expression levels of four DEGs: ASMT, TMEM244, FOLR1, and COL3A1. The primer sequences were described in Table 1. The relative

expression levels of ASMT and TMEM244 were found to be upregulated, whereas those of FOLR1 and COL3A1 were downregulated in the AE and CE groups. The results of the  $2^{-\Delta\Delta C_t}$  method were consistent with those of RNA-seq (Figure 9).

## Discussion

High-throughput sequencing technology has been widely applied to identify reliable molecular markers and to elucidate mechanisms underlying different diseases. Herein we conducted a comprehensive transcriptomic study that involved healthy controls and patients with AE and CE, and detected several DEGs by RNA-Seq. In subsequent analyses, we specifically focused on DEGs that showed considerably aberrant expression levels and those that were key nodes in the PPI networks.

The expression level of *guanylate-binding protein 1* (GBP1) gene, encoding guanylate-binding protein 1 (GBP1) belonging to

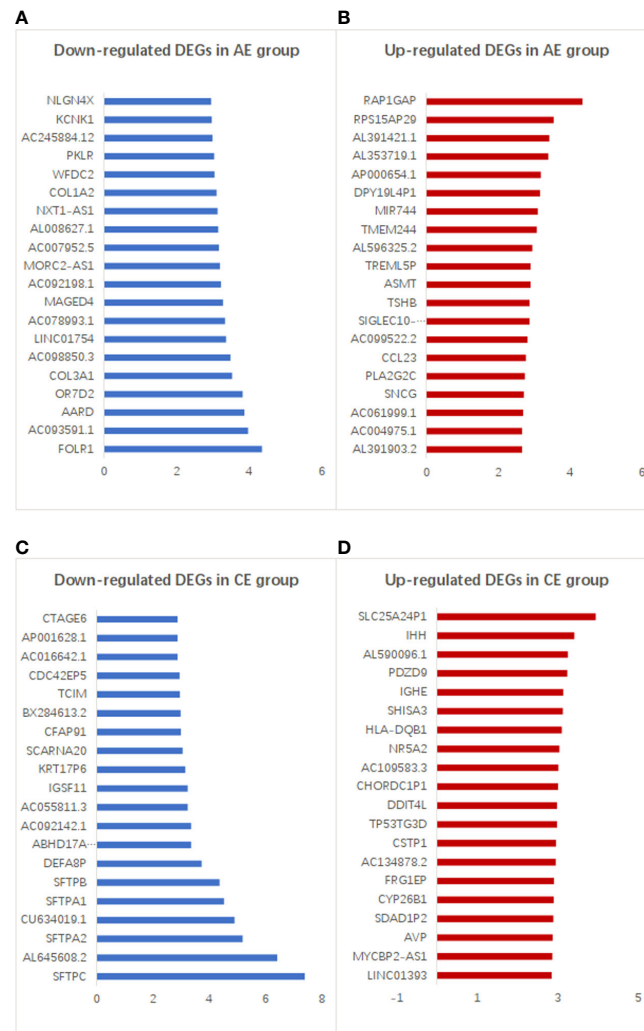


FIGURE 4

Top 20 up- and downregulated DEGs in the AE and CE groups. (A) Down- and (B) upregulated genes with higher differential expression levels in patients with AE. (C) Down- and (D) upregulated genes with higher differential expression levels in patients with CE.

the dynamin family, was upregulated in patients with AE and CE. GBP1 specifically binds to guanine nucleotides, and its expression is induced by interferon (IFN)- $\gamma$  and other inflammatory factors, such as IFN- $\alpha$ , IFN- $\beta$ , IL-1 $\alpha$ , IL-1 $\beta$ , and TNF- $\alpha$ . In addition to its association with inflammation, GBP1 participates in cell-autonomous defense against intracellular pathogens. Moreover, in recent years, GBP1 has been reported to play a pivotal role in the occurrence and development of tumors, including liver cancer. Its expression was found to be positively correlated with disease severity and tissue differentiation degree (Takagi et al., 2011). In patients with HCV, IFN induces the upregulation of the expression of GBP1, which binds to NS5B polymerase, inhibiting viral replication (Tretina et al., 2019). GBP1 was also demonstrated to be a key protective factor against *Toxoplasma gondii* infection

in human mesenchymal stem cells. Its recruitment to the parasitophorous vacuole membrane in IFN- $\gamma$ -stimulated human mesenchymal stem cells was identified to significantly inhibit *T. gondii* replication (Qin et al., 2017). Further, pathological staining of liver specimens of patients with echinococcosis showed infiltration of macrophages, neutrophils, eosinophils, and T lymphocytes around lesions (Diaz et al., 2016), causing atypical protein expression.

The expression level of *C-X-C motif chemokine 10* (CXCL10) gene, encoding C-X-C motif chemokine 10 (CXCL10), was upregulated in patients with AE and CE. CXCL10 is secreted by several types of cells, including monocytes, endothelial cells, and fibroblasts, in response to IFN- $\gamma$ . CXCL10 binds to the chemokine receptor CXCR3, which is expressed by Th1 cells, and induces lymphocytes to migrate to the infection site, causing lymphatic

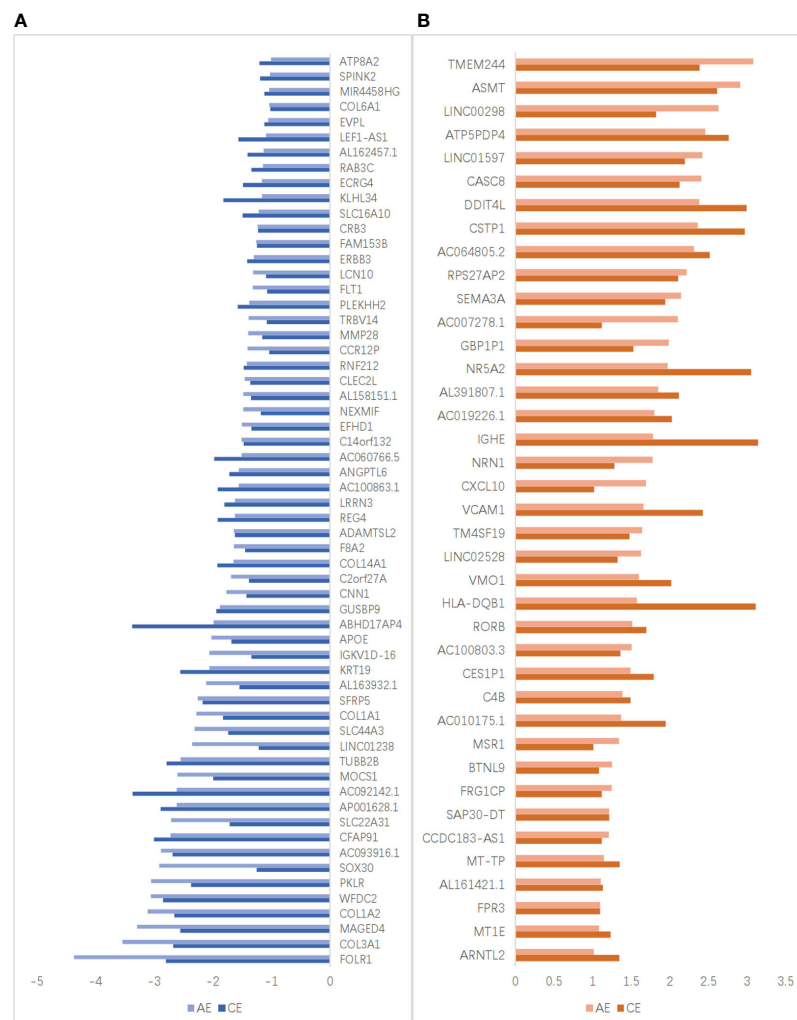
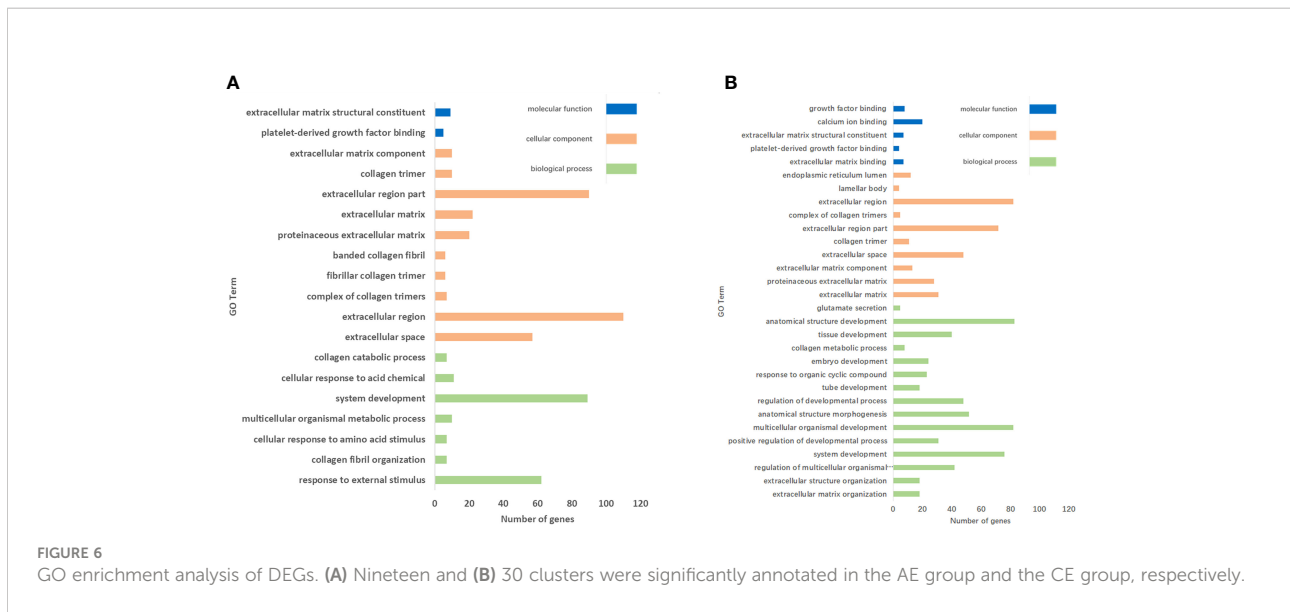


FIGURE 5

DEGs detected both in the AE and CE groups. **(A)** The expression levels of 39 genes, including FOLR1, COL3A1, and MAGED4, were downregulated both in the AE and CE groups. Light blue and dark blue lines represent the FPKM of downregulated DEGs in patients with AE and CE, respectively. **(B)** The expression levels of 60 genes, including TMEM244, ASMT, and LINC00298, were upregulated both in the AE and CE groups. Light red and dark red lines represent the FPKM of upregulated DEGs in patients with AE and CE, respectively.

infiltration and local immune response. Meanwhile, the recruited Th1 cells secrete IFN- $\gamma$ , which stimulates the local production of CXCL10 (Booth et al., 2002; Colvin et al., 2004). Serum CXCL10 levels in chronic HBV patients have been reported to be significantly higher than those in healthy controls and to be positively correlated with serum ALT and HBV DNA levels. Besides, serum CXCL10 levels in patients with HBC were also found to be associated with the severity of tissue necrotizing inflammation, with the increase in CXCL10 levels serving as an independent influencing factor for predicting the degree of fibrosis (Wang et al., 2008; You et al., 2011; Grebely et al., 2013). Higher serum CXCL10 levels may recruit more CD4<sup>+</sup> T cells for metastasis to the liver, leading to a more severe immune response; echinococcosis might have a similar mechanism.

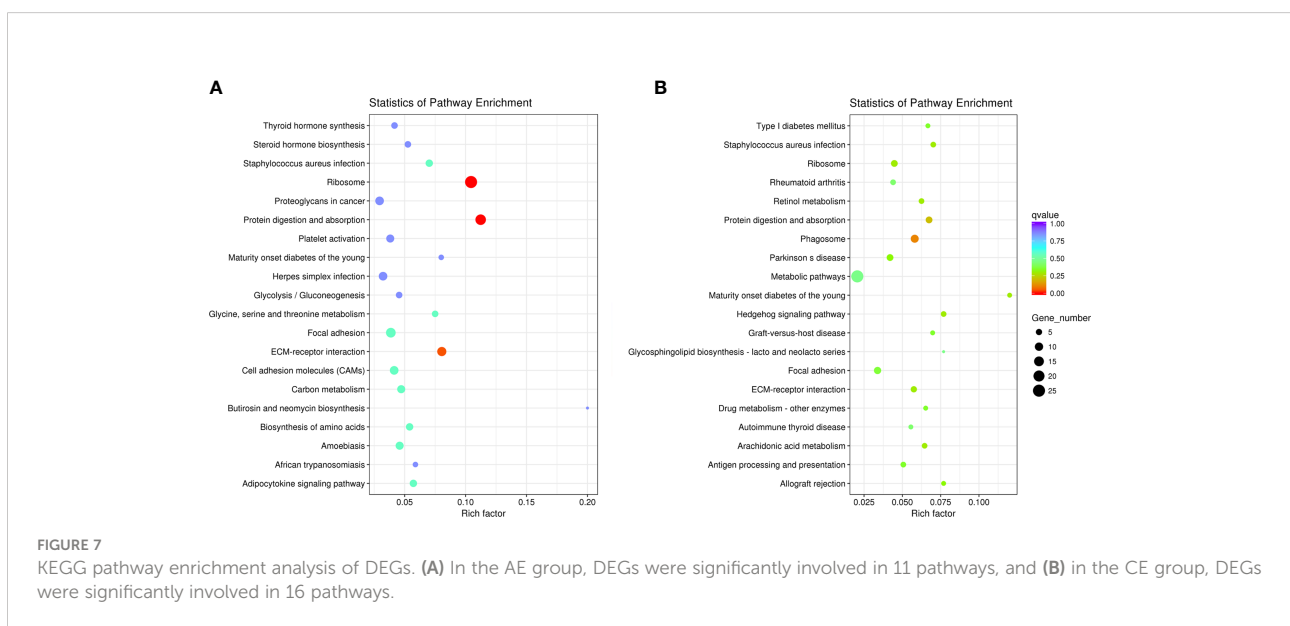
The expression level of *vascular cell adhesion molecule-1* (VCAM-1) gene, encoding vascular cell adhesion molecule-1 (VCAM-1), was upregulated in patients with AE and CE. VCAM-1 is an inducible cell surface sialoglycoprotein that was first discovered on cytokine-activated endothelium. It is expressed in large and small blood vessels after cytokine stimulation, and mediates adhesion of lymphocytes, monocytes, eosinophils, and basophils to vascular endothelium. Once endothelial cells are stimulated by cytokines, VCAM-1 is released into the bloodstream following proteolytic cleavage by vascular endothelial activation mediated by proinflammatory cytokines (Pirisi et al., 1996; Wang et al., 2014). VCAM-1 has been suggested to play a role in the pathogenesis of chronic hepatitis and cirrhosis, as its



expression levels were found to be upregulated in chronic liver disease (Antonova et al., 2013; Wadkin et al., 2017; Ali et al., 2020). Furthermore, VCAM-1 levels were observed to be significantly higher in HBsAg-positive patients and positively correlated with ALT. VCAM-1 is evidently involved in the destruction of liver cells as it participates in the killing process mediated by T and NK cells and in antibody-dependent cellular cytotoxicity mediated by neutrophils and monocytes.

*Collagen 1A1 (COL1A1)* and *COL1A2* genes encode type I collagen  $\alpha 1$  and  $\alpha 2$ , respectively; *COL3A1* gene encodes type III collagen (Exposito et al., 2010). In this study, the expression levels of these genes were downregulated in the AE and CE groups. Collagens are a family of proteins that strengthen and

support many tissues, including the cartilage, bone, tendon, skin, and sclera (Liu et al., 2016; Augusciak-Duma et al., 2018). Collagen fibers constitute the main component of the extracellular matrix (ECM). Their excessive synthesis and reduced degradation can accordingly cause abnormal proliferation of fibrous connective tissue and promote fibrosis development, which is one of the important clinical manifestations of echinococcosis in the liver. *COL1A1*, *COL1A2*, and *COL3A1* have been previously reported to be abnormally overexpressed in some tumors and parasitic diseases (Li et al., 2016; Tao et al., 2018; Ma et al., 2019). Mice infected with echinococcosis were also found to show higher *COL1A1* and *COL3A1* expression levels in the hepatic tissue (Wang et al.,





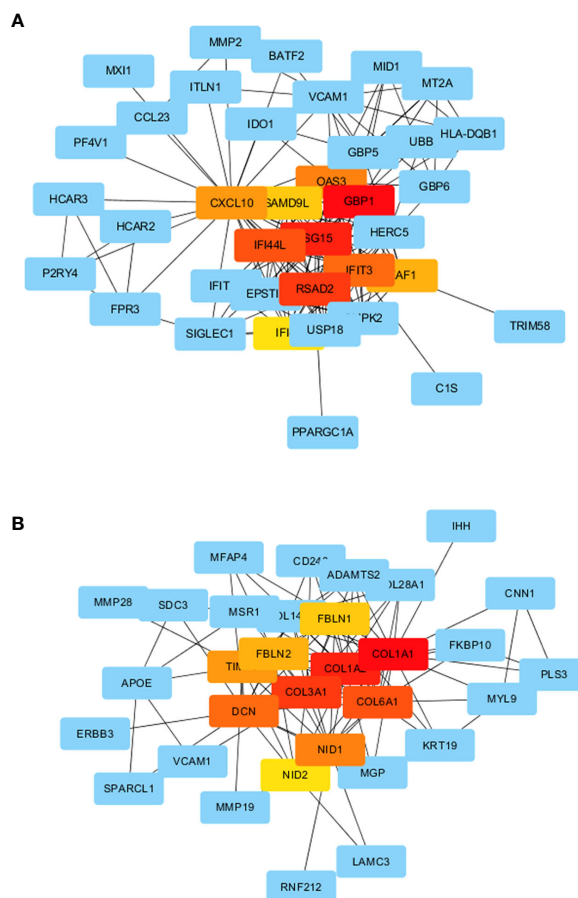


FIGURE 8 PPI networks. (A) PPI network for the AE group and (B) CE group.

2012). In this study, the expression levels of these genes in the blood samples were different. Some connection might exist between these contradictory results, and thus, further studies are warranted.

The expression level of *fibulin-1(FBLN1)* gene, which encodes fibulin-1, was significantly upregulated in patients with CE, but not in those with AE. As a member of the fibulin glycoprotein family, fibulin-1 is an ECM and blood glycoprotein,

and it functions as a bridge in the organization of ECM supramolecular structures. Fibulin-1 can also be incorporated into fibronectin-containing matrix fibers to play a role in cell adhesion and migration along protein fibers within the ECM. It is implicated in cellular transformation and tumor invasion, and it can behave both as an oncosuppressor and oncogene depending on the tissue environment (de Vega et al., 2009; Obaya et al., 2012). In a recent study, FBLN1 expression was

TABLE 1 Primer sequences used in the qRT-PCR experiment.

Gene	Forward sequence	Reverse sequence
GAPDH	CTCCTCCACCTTTGACGCT	TCTTCTCTTGTGCTCTTGCT
ASMT	GCATGTCTCTGTACCCTGGA	GGTCTTTGAAGAAATCCCCTTCC
TMEM244	GCGTGATGTTTGAGGTGCAT	CCCAAACCCATTCTTCCACAAC
FOLR1	CCCTGGAGGAAGAATGCCTG	CCAGCTCTGATCCACCTGCT
COL3A1	TTGGGTTGTCTAATATGGT	TCTCAGGATTTGTAGGGAT

GAPDH, Glyceraldehyde-3-phosphate dehydrogenase; ASMT, Acetylserotonin O-methyltransferase; TMEM244, Transmembrane protein 244; FOLR1, Folate receptor alpha; COL3A1, Collagen alpha-1(III) chain.

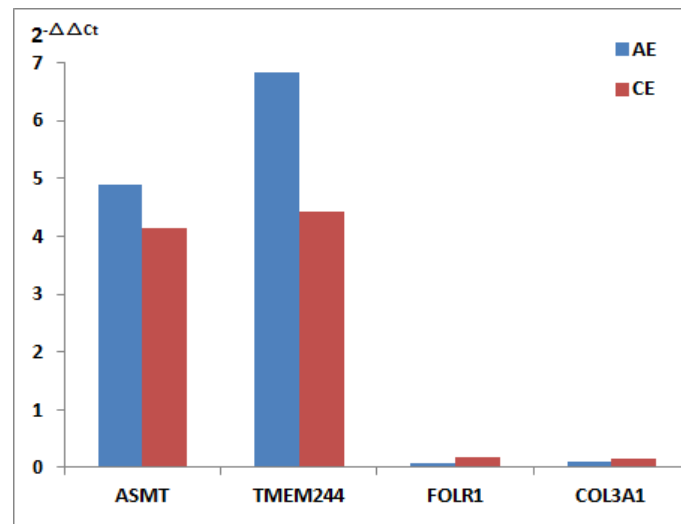


FIGURE 9  
Expression levels of DEGs by qRT-PCR. Relative expression levels ( $2^{-\Delta\Delta C_t}$ ) of ASMT, TMEM244, FOLR1, and COL3A1 in the AE and CE groups.

found to be significantly downregulated in exosomes derived from the plasma of patients with HIV, and was particularly altered in HIV<sup>+</sup> smokers (Kodidela et al., 2020). Similar results were reported for serum exosomal fibulin-1 in patients with gastric cancer (Yang et al., 2016). Some other studies have also reported that fibulin-1 expression is significantly upregulated in some cancers, validating its role in promoting tumor occurrence and development (Moll et al., 2002; Bardin et al., 2005). The changes in FBLN1 levels are related to pathological characteristics, but the underlying mechanism appears to be complex and thus warrants an in-depth investigation.

In this study, the application of RNA-Seq generated high quality, accurate data. RNA expression patterns observed in this study indicated that many immunological factors encoded by DEGs might be involved in the pathogenesis of echinococcosis. The results indicate possible target genes related to the pathogenesis of echinococcosis through statistical tests. The top DEGs identified herein could serve as potential diagnostic biomarkers and a basis for further biological and functional investigations.

This study also had a limitation. The sample size was small, which might not provide sufficient evidence to support our hypotheses. Further studies with a larger sample size are thus warranted to identify even more DEGs and to investigate their relationship with echinococcosis more comprehensively.

## Conclusions

DEGs with the most significant differences in expression levels and hub genes in PPI networks might serve as potential

diagnostic biomarkers. Further, they also can serve as a basis for biological and functional investigations of echinococcosis with larger sample sizes.

## Data availability statement

The datasets analyzed for this study can be found in the Genome Sequence Archive for Human (GSA-Human, the National Genomics Data Center), <https://bigd.big.ac.cn/gsa-human/browse/HRA001670>.

## Ethics statement

The present study was approved by the Ethics Committee of Sichuan Centre for Disease Control and Prevention (reference no. 2018-2), and all methods were performed in accordance with the relevant guidelines and regulations of the Declaration of Helsinki.

## Author contributions

Conceived and designed the experiments: LL, JC, BZ, and WL. Collected the samples: FC, QW, and KX. Performed the experiments: LL, FC, SJ, and QW. Analyzed the data: LL, FC, SJ, and YW. Wrote the paper: LL, JC, and YW. All authors have read and approved the manuscript.

## Funding

The research was supported by the Comprehensive Prevention and Control Pilot Project on Echinococcosis in Shiqu County, the China Postdoctoral Science Foundation (No. 2019M650800 to LL), the National Nature Science Foundation of China (No. 81971969 to JC), and the Three-Year Public Health Action Plan (2020-2022) of Shanghai (No. GWV-10.1-XX13 to JC).

## Acknowledgments

We would like to thank the staff at the Garzê County Center for Disease Control and Prevention (Garzê Prefecture, Sichuan Province) for their assistance with volunteer recruitment and sample collection.

## References

- Ali, R. O., Moon, M. S., Townsend, E. C., Hill, K., Zhang, G. Y., Bradshaw, A., et al. (2020). Exploring the link between platelet numbers and vascular homeostasis across early and late stages of fibrosis in hepatitis c. *Dig. Dis. Sci.* 65 (2), 524–533. doi: 10.1007/s10620-019-05760-x
- Anders, S., and Huber, W. (2010). Differential expression of RNA-seq data at the gene level—the DESeq package. *embl, Heidelberg, Germany*. Available at: <http://bioconductor.riken.jp/packages/3.4/bioc/vignettes/DESeq/inst/doc/DESeq.pdf>
- Anders, S., and Huber, W. (2016). *Differential expression of RNA-seq data at the gene level—the DESeq package* Vol. 10 (embl. Heidelberg), f1000 research.
- Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 31 (2), 166–169. doi: 10.1093/bioinformatics/btu638
- Antonova, T. V., Romanova, M. A., and Lyymar', I. (2013). Markers of endothelial dysfunction (VCAM-1 and vWF) in chronic hepatitis c. *Ter. Arkh.* 85 (12), 86–89.
- Augusiak-Duma, A., Witecka, J., Sieron, A. L., Janeczko, M., Pietrzyk, J. J., Ochman, K., et al. (2018). Mutations in the COL1A1 and COL1A2 genes associated with osteogenesis imperfecta (OI) types I or III. *Acta Biochim. Pol.* 65 (1), 79–86. doi: 10.18388/abp.2017\_1612
- Bardin, A., Moll, F., Margueron, R., Delfour, C., Chu, M. L., Maudelonde, T., et al. (2005). Transcriptional and posttranscriptional regulation of fibulin-1 by estrogens leads to differential induction of messenger ribonucleic acid variants in ovarian and breast cancer cells. *Endocrinol.* 146 (2), 760–768. doi: 10.1210/en.2004-1239
- Booth, V., Keizer, D. W., Kamphuis, M. B., Clark-Lewis, I., and Sykes, B. D. (2002). The CXCR3 binding chemokine IP-10/CXCL10: structure and receptor interactions. *Biochem.* 41 (33), 10418–10425. doi: 10.1021/bi026020q
- Colvin, R. A., Campanella, G. S., Sun, J., and Luster, A. D. (2004). Intracellular domains of CXCR3 that mediate CXCL9, CXCL10, and CXCL11 function. *J. Biol. Chem.* 279 (29), 30219–30227. doi: 10.1074/jbc.M403595200
- Craig, P. S., Giraudoux, P., Wang, Z. H., and Wang, Q. (2019). Echinococcosis transmission on the Tibetan plateau. *Adv. Parasitol.* 104, 165–246. doi: 10.1016/bs.apar.2019.03.001
- Dehkordi, A. B., Sanei, B., Yousefi, M., Sharafi, S. M., Safarnezhad, F., Jafari, R., et al. (2019). Albendazole and treatment of hydatid cyst: review of the literature. *Infect. Disord. Drug Targets.* 19 (2), 101–104. doi: 10.2174/1871526518666180629134511
- Deplazes, P., Rinaldi, L., Alvarez Rojas, C. A., Torgerson, P. R., Harandi, M. F., Romig, T., et al. (2017). Global distribution of alveolar and cystic echinococcosis. *Adv. Parasitol.* 95, 315–493. doi: 10.1016/bs.apar.2016.11.001
- de Vega, S., Iwamoto, T., and Yamada, Y. (2009). Fibulins: multiple roles in matrix structures and tissue functions. *Cell. Mol. Life. Sci.* 66 (11-12), 1890–1902. doi: 10.1007/s00018-009-8632-6
- Diaz, A., Casaravilla, C., Barrios, A. A., and Ferreira, A. M. (2016). Parasite molecules and host responses in cystic echinococcosis. *Parasite Immunol.* 38 (3), 193–205. doi: 10.1111/pim.12282
- Exposito, J. Y., Valcourt, U., Cluzel, C., and Lethias, C. (2010). The fibrillar collagen family. *Int. J. Mol. Sci.* 11 (2), 407–426. doi: 10.3390/ijms11020407
- Fu, M., Han, S., Xue, C., Wang, X., Liu, B., Wang, Y., et al. (2020). Contribution to the echinococcosis control programme in China by NIPD-CTDR. *Adv. Parasitol.* 110, 107–144. doi: 10.1016/bs.apar.2020.04.010
- Gharibi, Z., Rahdar, M., Pirestani, M., Tavalla, M., and Tabandeh, M. R. (2021). The immunization of protoscolices P29 DNA vaccine on experimental cystic echinococcosis in balb/c mice. *Acta Parasitol.* 66 (4), 1114–1121. doi: 10.1007/s11686-021-00367-7
- Grebel, J., Feld, J. J., Applegate, T., Matthews, G. V., Hellard, M., Sherker, A., et al. (2013). Plasma interferon-gamma-inducible protein-10 (IP-10) levels during acute hepatitis c virus infection. *Hepatology*. 57 (6), 2124–2134. doi: 10.1002/hep.26263
- Jazouli, M., Lightowers, M. W., Bamouh, Z., Gauci, C. G., Tadaoui, K., Ennaji, M. M., et al. (2020). Immunological responses and potency of the EG95NC-recombinant sheep vaccine against cystic echinococcosis. *Parasitol. Int.* 78, 102149. doi: 10.1016/j.parint.2020.102149
- Kammerer, W. S., and Schantz, P. M. (1993). Echinococcal disease. *Infect. Dis. Clin. North Am.* 7 (3), 605–618. doi: 10.1016/S0891-5520(20)30545-6
- Kodidela, S., Wang, Y., Patters, B. J., Gong, Y., Sinha, N., Ranjit, S., et al. (2020). Proteomic profiling of exosomes derived from plasma of HIV-infected alcohol drinkers and cigarette smokers. *J. Neuroimmune Pharmacol.* 15 (3), 501–519. doi: 10.1007/s11481-019-09853-2
- Li, J., Ding, Y., and Li, A. (2016). Identification of COL1A1 and COL1A2 as candidate prognostic factors in gastric cancer. *World J. Surg. Oncol.* 14 (1), 297. doi: 10.1186/s12957-016-1056-5
- Liu, J., Luo, C., Yin, Z., Li, P., Wang, S., Chen, J., et al. (2016). Downregulation of let-7b promotes COL1A1 and COL1A2 expression in dermis and skin fibroblasts during heat wound repair. *Mol. Med. Rep.* 13 (3), 2683–2688. doi: 10.3892/mmr.2016.4877
- Ma, H. P., Chang, H. L., Bamodu, O. A., Yadav, V. K., Huang, T. Y., Wu, A., et al. (2019). Collagen 1A1 (COL1A1) is a reliable biomarker and putative therapeutic target for hepatocellular carcinogenesis and metastasis. *Cancers.* 11 (6), 786. doi: 10.3390/cancers11060786
- Moll, F., Katsaros, D., Lazennec, G., Hellio, N., Roger, P., Giacalone, P. L., et al. (2002). Estrogen induction and overexpression of fibulin-1C mRNA in ovarian cancer cells. *Oncogene.* 21 (7), 1097–1107. doi: 10.1038/sj.onc.1205171
- Moro, P. L., and Schantz, P. M. (2006). Echinococcosis: historical landmarks and progress in research and control. *Ann. Trop. Med. Parasitol.* 100 (8), 703–714. doi: 10.1179/136485906X112257
- Nunnari, G., Pinzone, M. R., Gruttadauria, S., Celesia, B. M., Madeddu, G., Malaguarnera, G., et al. (2012). Hepatic echinococcosis: clinical and therapeutic aspects. *World J. Gastroenterol.* 18 (13), 1448–1458. doi: 10.3748/wjg.v18.i13.1448
- Obaya, A. J., Rua, S., Moncada-Pazos, A., and Cal, S. (2012). The dual role of fibulins in tumorigenesis. *Cancer Lett.* 325 (2), 132–138. doi: 10.1016/j.canlet.2012.06.019
- Pirisi, M., Fabris, C., Falletti, E., Soardo, G., Toniutto, P., Vitulli, D., et al. (1996). Serum soluble vascular-cell adhesion molecule-1 (VCAM-1) in patients with acute and chronic liver diseases. *Dis. Markers.* 13 (1), 11–17. doi: 10.1155/1996/129325

## 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.

- Qin, A., Lai, D. H., Liu, Q., Huang, W., Wu, Y. P., Chen, X., et al. (2017). Guanylate-binding protein 1 (GBP1) contributes to the immunity of human mesenchymal stromal cells against *Toxoplasma gondii*. *Proc. Natl. Acad. Sci. U. S. A.* 114 (6), 1365–1370. doi: 10.1073/pnas.1619665114
- Schweiger, A., Grimm, F., Tanner, I., Müllhaupt, B., Bertogg, K., Müller, N., et al. (2012). Serological diagnosis of echinococcosis: the diagnostic potential of native antigens. *Infect.* 40 (2), 139–152. doi: 10.1007/s15010-011-0205-6
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic. Acids Res.* 43 (Database issue), D447–D452. doi: 10.1093/nar/gku1003
- Takagi, K., Takayama, T., Nagase, H., Moriguchi, M., Wang, X., Hirayanagi, K., et al. (2011). High TSC22D3 and low GBP1 expression in the liver is a risk factor for early recurrence of hepatocellular carcinoma. *Exp. Ther. Med.* 2 (3), 425–431. doi: 10.3892/etm.2011.236
- Tao, R., Fan, X. X., Yu, H. J., Ai, G., Zhang, H. Y., Kong, H. Y., et al. (2018). MicroRNA-29b-3p prevents *Schistosoma japonicum*-induced liver fibrosis by targeting COL1A1 and COL3A1. *J. Cell Biochem.* 119 (4), 3199–3209. doi: 10.1002/jcb.26475
- Tretina, K., Park, E. S., Maminska, A., and MacMicking, J. D. (2019). Interferon-induced guanylate-binding proteins: Guardians of host defense in health and disease. *J. Exp. Med.* 216 (3), 482–500. doi: 10.1084/jem.20182031
- Wadkin, J., Patten, D. A., Kamarajah, S. K., Shepherd, E. L., Novitskaya, V., Berditchevski, F., et al. (2017). CD151 supports VCAM-1-mediated lymphocyte adhesion to liver endothelium and is upregulated in chronic liver disease and hepatocellular carcinoma. *Am. J. Physiol. Gastrointest. Liver Physiol.* 313 (2), G138–G149. doi: 10.1152/ajpgi.00411.2016
- Wang, J. H., Wei, X. F., Zhang, C. S., Lv, G. D., Jiang, T., Lu, X. M., et al. (2012). The expression and significance of Col1a1, Col3a1, and Col4a1 in hepatic tissue of BALB/c mice infected with *Echinococcus multilocularis* in the middle stages of infection. *J. Pathog. Biol.* 7 (09), 653–656. doi: 10.13350/j.cjpb.2012.09.021
- Wang, P. C., Weng, C. C., Hou, Y. S., Jian, S. F., Fang, K. T., Hou, M. F., et al. (2014). Activation of VCAM-1 and its associated molecule CD44 leads to increased malignant potential of breast cancer cells. *Int. J. Mol. Sci.* 15 (3), 3560–3579. doi: 10.3390/ijms15033560
- Wang, Q., Yang, L., Wang, Y., Zhang, G. J., Zhong, B., Wu, W. P., et al. (2020). Disease burden of echinococcosis in Tibetan communities—a significant public health issue in an underdeveloped region of western China. *Acta Trop.* 203, 105283. doi: 10.1016/j.actatropica.2019.105283
- Wang, J., Zhao, J. H., Wang, P. P., and Xiang, G. J. (2008). Expression of CXC chemokine IP-10 in patients with chronic hepatitis b. *Hepatobiliary Pancreat. Dis. Int.* 7 (1), 45–50.
- World Health Organization (2011). *Working to overcome the global impact of neglected tropical diseases: first WHO report on neglected tropical diseases* <https://www.who.int/publications/i/item/9789241564090> [Accessed February 11, 2010]
- Yang, H., Fu, H. L., MAO, J. H., Li, X., Wang, M., Qian, H., et al. (2016). Detection and diagnostic application of serum exosomal fibulin-1 in the patients with gastric cancer. *Clin. Lab. Sci.* 34 (2), 107–109. doi: 10.13602/j.cnki.jcls.2016.02.07
- Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11 (2), R14. doi: 10.1186/gb-2010-11-2-r14
- You, C. R., Park, S. H., Jeong, S. W., Woo, H. Y., Bae, S. H., Choi, J. Y., et al. (2011). Serum IP-10 levels correlate with the severity of liver histopathology in patients infected with genotype-1 HCV. *Gut Liver.* 5 (4), 506–512. doi: 10.5009/gnl.2011.5.4.506