

Identification of Long Noncoding RNAs Expression Profiles Between Gallstone and Gallbladder Cancer Using Next-Generation Sequencing Analysis

Qiang Wang*, Pinduan Bi*, Ding Luo, Pingli Cao, Weihong Chen, Bin Yang

Department of Hepatobiliary Surgery, the First Affiliated Hospital of Kunming Medical University, Kunming, People's Republic of China

*These authors contributed equally to this work

Correspondence: Bin Yang, Department of Hepatobiliary Surgery, the First Affiliated Hospital of Kunming Medical University, No. 295, Xichang Road, Wuhua District, Kunming, People's Republic of China, Tel +86 013987642890, Email yangbin@kmmu.edu.cn

Background: Gallstone disease (GS) is an important risk factor for Gallbladder cancer (GBC). However, the mechanisms of the progression of GS to GBC remain unclear. Long non-coding RNA (lncRNA), modulates DNA/RNA/proteins at epigenetic, pre-transcriptional, transcriptional and posttranscriptional levels, and plays a potential therapeutic role in various diseases. This study aims to identify lncRNAs that have a potential impact on GS-promoted GBC progression.

Methods and Results: Six GBC patients without GS, six normal gallbladder tissues, nine gallstones and nine GBC patients with GS were admitted to our hospital. The next-generation RNA-sequencing was performed to analyze differentially expressed (DE) lncRNA and messenger RNA (mRNA) in four groups. Then overlapping and specific molecular signatures were analyzed. We identified 29 co-DEGs and 500 co-DElncRNAs related to gallstone or GBC. The intersection and concatenation of co-DEGs and co-DElncRNA functionally involved in focal adhesion, Transcriptional misregulation in cancers, Protein digestion and absorption, and ECM-receptor interaction signaling pathways may contribute to the development of gallbladder cancer. Further exploration is necessary for early diagnosis and the potential treatment of GBC. FXYD2, MPZL1 and PAH were observed in both co-DEGs and co-DElncRNA and validated by qRT-PCR.

Conclusion: Our data identified a series of DEGs and DElncRNAs, which were involved in the progression of GBC and GS-related metabolism pathways. Compared to GBC, the GS profile was more similar to para-tumor tissues in transcriptome level and lower risk of cancer. Further exploration is necessary from GBC patients with different periods of follow-up gallstone.

Keywords: gallstone, gallbladder cancer, long non-coding RNA, next-generation sequencing analysis, computational biology

Introduction

Gallbladder cancer (GBC) is the most common malignant tumor of the biliary system (accounting for 80–95%).¹ It is the sixth most common malignant tumor of the digestive tract and occurs three to four times as frequently in women as in men.² Gallbladder cancer is relatively high in Latin America and Asia.³ GBC is not sensitive to radiotherapy and chemotherapy.^{1,4} There is currently no recognized chemotherapy regimen that can effectively treat gallbladder cancer.⁵ Once the tumor metastasizes or recurs after surgery, doctors are helpless and cannot effectively achieve effective treatment through comprehensive treatment. Gallbladder cancer patients have an extremely poor prognosis.⁶ Therefore, the patients would benefit from an early and accurate diagnosis of gallbladder cancer.

Due to the lack of specific clinical manifestations of gallbladder cancer, and no other highly specific auxiliary examination methods, most patients are already at the advanced stage of gallbladder cancer due to tumor progression when they are first diagnosed, and they lose the opportunity for radical surgery.⁷ Mostly, GBC is diagnosed incidentally in patients undergoing cholecystectomy for the treatment of cholecystitis or cholelithiasis.⁸ Epidemiological and pathological evidence confirms that gallstone (GS) patients are more susceptible to GBC than healthy individuals. Based on this, exploring the molecular signatures between GBC and GS could help to understand the molecular mechanism of the progression of GS to GBC, and

to find highly specific biomarkers and effective drug treatment targets for GBC, which is beneficial to the development of therapeutic drugs for gallbladder cancer.

The occurrence and progression of tumors is a complex process involving multiple genes and multiple factors.⁹ Genetic factors have always been considered to play an important role.¹⁰ Studies in recent years have found that the function of RNA is not only to translate into protein. Many RNAs do not have the function of encoding protein, so-called non-coding RNA (ncRNA).¹¹ Among them, those with a length > 200 nt are named long non-coding RNA (lncRNA).¹² At the earliest, this type of molecule was mistakenly regarded as a by-product of RNA polymerase II transcription, a “noise” in the genome transcription process, without biological function, and did not attract widespread attention from scholars.¹³ However, accumulating evidence shows that lncRNAs play an important role in the development of organisms and diseases,¹⁴ including the occurrence and progression of malignant tumors.^{15,16} At present, the related research of lncRNA is still in its infancy. For most lncRNAs discovered by sequencing, their specific biological function is still unclear. Therefore, lncRNA as an emerging field has broad research prospects and is worthy of in-depth research by researchers.

In this study, we used next-generation RNA-seq and bioinformatics analysis to identify differentially expressed lncRNAs (DELncRNAs) and differentially expressed genes (DEGs), which may drive the GS process to GBC. Then, the differential expression of these biomarkers was validated in patients. The objective is to provide new insights into the mechanisms of GBC onset and progression.

Materials and Methods

Ethics Statement

All aspects of the present study were approved by the Ethics Committee of The First Affiliated Hospital of Kunming Medical University and accordance with the Declaration of Helsinki. Six GBC patients without gallstones (GBC-nGS), nine GBC patients with gallstones (GBC-GS) and nine gallstone patients (GS) were admitted. The paratumor (normal tissue away from the tumor) tissues from GBC-nGS were considered NS. All patients provided written informed consent for the use of their surgical specimens in the present study in accordance with the Committee’s regulations. The clinical characteristics of the patients are presented in Table 1. Tumor and paratumor (normal tissue adjacent to the tumor) tissues from GBC patients with or without GS were collected and immediately stored in liquid nitrogen until RNA-seq (Figure 1).

Table 1 Clinical Characteristics of Cases in This Study

Clinical Characters	GS/N	NS/N	GBC-nGS/N	GBC-GS/N	P value
Total	9	6	6	9	
Gender					>0.05
Male	3	2	2	3	
Female	6	4	4	6	
Age	64.8 (60.55–69)	56.2 (42.2–70)	66.0 (59.3–72.7)	64.8 (60.55–69)	0.116
TNM stage					0.574
I–II	-	-	2	1	
IIIA	-	-	1	2	
IVB	-	-	3	6	
Gallstone					
Yes	0	6	6	0	
No	9	0	0	9	
Size of stone	-	14.4 (7.5–21.4)	16.5 (10.6–22.5)	-	0.565
Smoking					0.969
Yes	4	3	3	4	
No	5	3	3	5	
Drinking					0.747
Yes	3	3	3	3	
No	6	3	3	6	

Abbreviations: NS, normal gallbladder; GS, gallstone; GBC-nGS, gallbladder carcinoma without gallstone; GBC-GS, gallbladder carcinoma with gallstone.

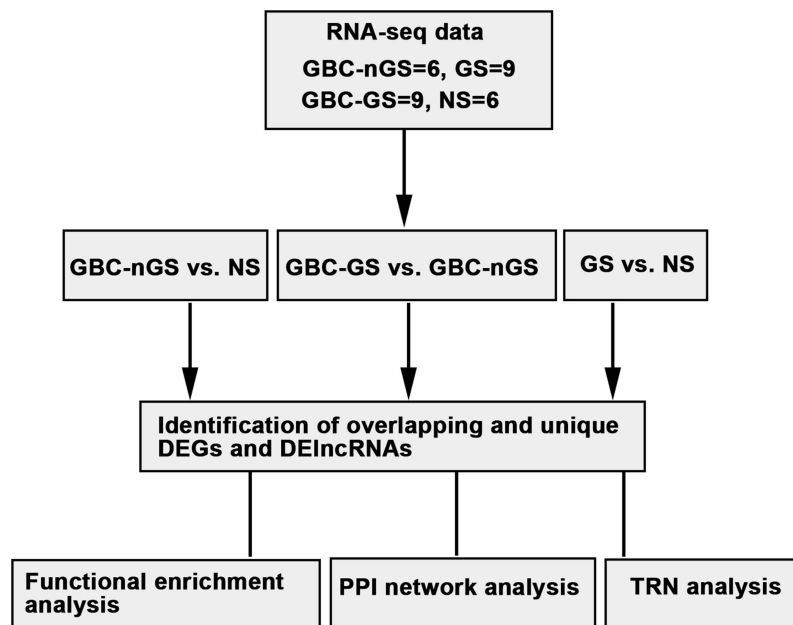


Figure 1 A Flowchart of the study process.

Abbreviations: NS, normal gallbladder; GS, gallstone; GBC-nGS, gallbladder carcinoma without gallstone; GBC-GS, gallbladder carcinoma with gallstone; DEGs, differentially expressed genes; PPI, protein-protein interaction.

RNA Sequencing

The tissue was broken into small blocks in liquid nitrogen, then total RNAs were isolated using TRIzol (Invitrogen, Carlsbad, CA). RNA was quantified and assessed for integrity using the Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA). Ribosomal RNA was removed by performing 1 to 2 washings of RNA annealed to poly-T oligo beads (Invitrogen, Life Technologies, Grand Island, NY). RNA was fragmented and reverse transcribed using random hexamers (Invitrogen). Double-stranded cDNA synthesis was performed using Pol I and RNase H. Short fragments were purified with QiaQuick PCR extraction kit (Qiagen, Hilden, Germany) and resolved with elution buffer for end reparation and poly(A) addition followed by ligation with sequencing adaptors for cluster generation and sequenced on the Illumina HiSeq 3000 (Illumina, San Diego, CA).

The Identification of Differential Expressed Genes and lncRNAs

For each sample, the obtained sequencing reads were cleaned by 5'-end quality trimming and clipping of the sequencing adaptors by SOPAnuke (<http://soap.genomics.org.cn/>). Then, we performed prealignment quality control of the cleaned sequencing reads using FastQC (v 0.11.5). Using STAR (version 2.5.0a), we aligned the RNA-Seq reads against human genome reference sequence GRCh37 under default parameters, allowing one or two base mismatches. Furthermore, read summarization of only reads spanning introns (intron spanning) was performed with HTSeq (version 0.6.0, <https://htseq.readthedocs.io/>). We included protein-coding and noncoding RNAs during read mapping, summarization, and subsequent analyses. Notably, we excluded the genes encoded on mitochondrial DNA and the Y chromosome from the analyses. In addition, the genes not expressed in 5% of the samples were excluded. Moreover, gene expression clustering analysis using k-mean clustering, principal component analysis, and subsequent statistical analyses were performed using R (version 3.0.3) and R-studio (version 0.98.1091).

RNA profile was normalized using counts per million. Using the edgeR package, we performed differential gene expression analysis for the paired samples. The DEGs and DElncRNAs were identified by using a *t*-test, first in GBC-nGS vs NS; GS vs NS; and second, GBC-GS vs GBC-nGS. DEGs and DElncRNAs (FDR filtering (fold change ≥ 2.0) and $p < 0.05$) were selected for further study. Hierarchical clustering was performed to identify DEGs and DElncRNAs expression patterns, using Cluster 3.0

(Stanford, CA, USA). Subsequently, we performed GO functional enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of DEGs using the R package ClusterProfiler.¹⁷

Functional Analysis

The gene ontology (GO; <http://www.geneontology.org>) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.kegg.jp/>) analysis were used for understanding advanced functions and biological systems. The differential GO terms and KEGG pathways were identified using Fisher's exact test with $p < 0.05$ and false discovery rate (FDR) < 0.05 .

Co-Expression Correlation Network Analysis

A co-expression correlation network was constructed for DEGs and DElncRNA across GBC-nGS and GBC-GS, to identify gene modules associated with the GS process to GBC. The Search Tool for Retrieval of Interacting Genes database (STRING version 10.5; <https://string-db.org>)¹⁸ and protein-protein interaction (PPI) network analysis (PPI score = 0.9) were used. The network was visualized using Cytoscape version 3.6.1 (<http://www.cytoscape.org/>).¹⁹

RT-qPCR Verification

For gene expression validation, cDNA was generated from total RNA using InRcute lncRNA cDNA kit (Tiangen Biotech Co., Ltd., China). cDNA (250 ng) was used for qPCR to determine the lncRNA expression levels using InRcute lncRNAqPCR kit (SYBR Green) (Tiangen Biotech Co., Ltd., China) with custom PCR primers (Table S1). Samples were run on QuantStudio 12K Flex Real-Time PCR System (Applied Biosystems; Thermo Fisher Scientific, Inc.) and were subject to 3 min at 95°C before undergoing 40 cycles of denaturing at 95°C for 3 sec followed by annealing/extension at 60°C for 20 sec. Three replicate PCRs were performed. Relative expression of each gene was quantified using the gene's $2^{-\Delta\Delta Ct}$.

Statistical Analysis

All statistical analyses were performed using SPSS software (version 22.0; SPSS, Inc., Chicago, IL, USA). For comparison, a two-tailed Student's *t*-test was performed, when appropriate. Data were presented as the mean \pm standard error. All histograms were constructed using Prism for Windows (version 6.0; GraphPad Software, Inc., La Jolla, CA, USA). $P < 0.05$ was considered to indicate a statistically significant difference.

Results

Characteristics of Patients and Sequencing Data

Clinical characteristics of GBC-nGS, GBC-GS and GS patients are presented in Table 1. The acquisition of RNA-seq data consists of an average of 8.08 G clean data per sample with 94.83% mean Q30. HISAT2 showed that 97.30% of reads were mapped into GRCh37. A total of 53,261 lncRNAs were detected, of which 839 lncRNAs were newly predicted. After quality control of sequencing data, there was no significant difference between the groups of patients. The quality control demonstrated that RNA sequencing produced high-quality data.

Identification of DEGs and DElncRNA Expression Profiles Between NS and GBC-nGS

Differential gene expression analysis using DESeq2, including 947 DEGs (517 upregulated and 432 downregulated) and 2871 DElncRNA (1559 upregulated and 1312 downregulated) (Figure 2A and D). The top 20 DEGs and DElncRNAs are shown in Table 2.

For mRNA, GO analysis suggested that 256 GO items were functionally enriched involved in complement activation, antigen binding, nuclear nucleosome, cell adhesion, angiogenesis, and extracellular matrix organization (Figure 2B). The 16 pathways ($p < 0.05$) associated with DEGs were also enriched in Systemic lupus erythematosus, alcoholism, Transcriptional misregulation in cancers, Protein digestion and absorption, and ECM-receptor interaction signaling pathway (Figure 2C).

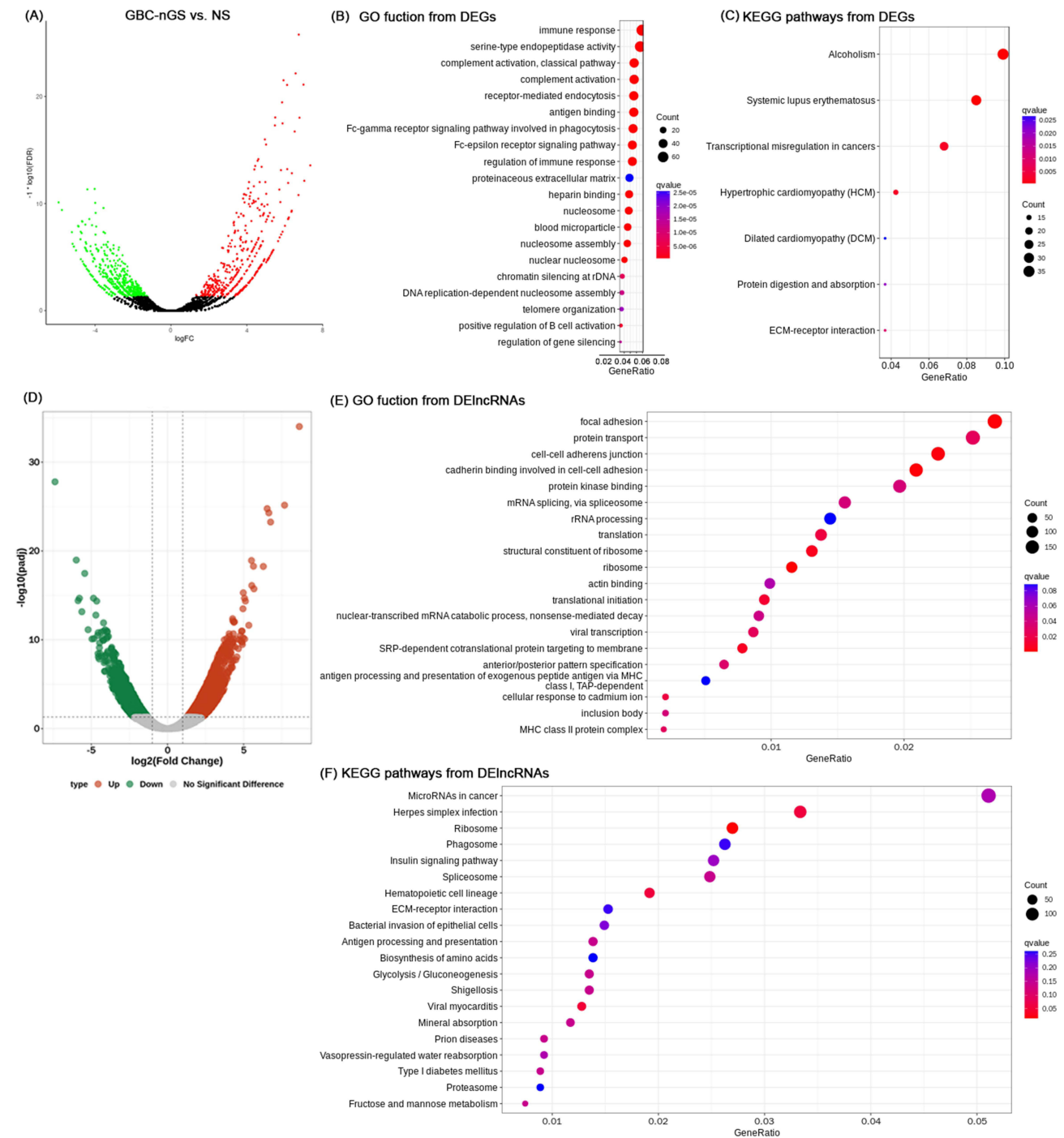


Figure 2 Differences in the lncRNA and mRNA expression profiles between GBC-nGS and NS tissues. **(A)** Volcano plot of the differential expression analysis of DEGs (fold change ≥ 2.0 and $p < 0.05$). The top 15 items of DEGs of GBC-nGS were identified by **(B)** GO function enrichment and **(C)** KEGG enrichment. **(D)** Volcano plot of the differential expression analysis of DElncRNAs (fold change ≥ 2.0 and $p < 0.05$). The top 15 items of DElncRNAs of GBC-nGS were identified by **(E)** GO function enrichment and **(F)** KEGG enrichment.

Abbreviations: GBC-nGS, gallbladder cancer tissues without gallstone; NS, normal gallbladder; DEG, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes pathway.

For lncRNAs, 1522 target genes were predicted by trans- and cis-regulation, of which 191 showed significant differences in expression. GO enrichment analyses demonstrated that 41 GO items were distinctly related to focal adhesion, protein transport, ribosome, cell-cell adherens junction, protein kinase binding et al (Figure 2E). There are 29 KEGG items

Table 2 Next-Generation Sequencing Analysis of the 20 Most Aberrantly Regulated mRNA and lncRNAs Between Gallbladder Carcinoma Without Gallstone (GBC-nGS) and Normal Gallbladder Tissue (NS)

Genes	logFC	P value	lncRNAs	logFC	P value
CHRD1	-6.240	3.92E-14	HLA-A	8.657	5.13E-39
FABP4	-6.221	3.96E-14	COPG1	-7.382	1.70E-32
C7	-5.384	7.49E-12	MXRA8	7.688	1.10E-29
COL11A1	5.379	8.80E-12	MSTRG.32747	6.546	3.58E-29
HBB	-5.257	2.18E-11	MSTRG.20776	6.653	1.34E-28
IGHV3-49	5.025	1.18E-10	CDK5RAP3	6.770	1.72E-27
PLIN4	-4.654	1.12E-09	RPS7	-5.986	3.83E-23
BCL2L2	-4.631	1.25E-09	EEF1G	5.516	4.90E-23
KCNB1	-4.658	1.37E-09	DNAH1	5.647	2.45E-22
MUC6	-4.484	2.88E-09	KCNH2	6.294	2.88E-22
GJB2	4.510	3.34E-09	OAZ1	-5.431	1.93E-21
IFI30	4.433	4.18E-09	CDKN2B	5.538	4.83E-20
IGKV1-17	4.463	4.77E-09	RP5-1085F17.3	5.676	1.22E-19
CCL14	-4.370	6.25E-09	GAS5	4.978	3.70E-19
KLF2	-4.358	7.77E-09	MSTRG.27456	-4.856	1.66E-18
CHRM2	-4.278	1.00E-08	SF1	-5.784	1.76E-18
PLCXD3	-4.333	1.17E-08	MKL2	5.0568	1.84E-18
CFD	-4.271	1.40E-08	DERL3	-5.867	3.69E-18
OGN	-4.217	1.51E-08	CD276	5.117	4.49E-18
IGHV4-61	4.201	1.98E-08	KRT8	-4.647	4.60E-18

with p -value <0.05 , including microRNA in cancer, Ribosome, ECM-receptor interaction, Glycolysis signaling pathway, and Steroid hormone biosynthesis (Figure 2F).

DEGs and DElncRNAs (nGS vs GBC-nGS) were enriched in Transcriptional misregulation in cancers, Protein digestion and absorption, ECM-receptor interaction and microRNA in cancer, which contributed to GBC progress.

Identification of DEGs and DElncRNA Expression Profiles Between NS and GS

Differential gene expression analysis using DESeq2, including 372 DEGs (150 upregulated and 252 downregulated) and 2823 DElncRNA (1323 upregulated and 1500 downregulated) (Figure 3A and D). The top 20 DEGs and DElncRNAs are shown in Table 3. Compared to DEGs (nGS vs GBC-nGS), the number of DEGs (nGS vs GS) groups is significantly reduced. We hypothesized GS tissue was more similar to para-tumor tissues in transcriptome level and lower risk of cancer.

GO analysis ($p<0.05$) was performed to DEGs enrichment in biological processes, cellular components and molecular functions. The results suggested that 239 items were functionally enriched and involved in the apical plasma membrane, serine-type endopeptidase activity, heparin binding, ion transport, extracellular matrix organization (Figure 3B). The 96 pathways ($p<0.05$) associated with DEGs (nGS vs GS) were closely associated with metabolism, including cell adhesion molecules, Protein digestion and absorption, bile secretion, mineral absorption, Drug metabolism - cytochrome P450 (Figure 3C), which contributes to the pathogenesis of gallstones (GS).

For DElncRNAs, 195 target genes were predicted by trans- and cis-regulation, and showed significant differences in expression. GO enrichment analyses demonstrated 34 GO items were distinctly related to focal adhesion, protein kinase binding, cell-cell adherens junction, ribosome et al (Figure 3E). There are 81 KEGG items with p -value <0.05 , including Tuberculosis, Purine metabolism, Ribosome, Apoptosis, Shigellosis, Glycolysis signaling pathway, VEGF signaling pathway (Figure 3F). Compared to DElncRNAs (nGS vs GBC-nGS), DElncRNAs (nGS vs GS) are less enriched in cancers, such as MicroRNAs in cancers and ECM-receptor interaction.

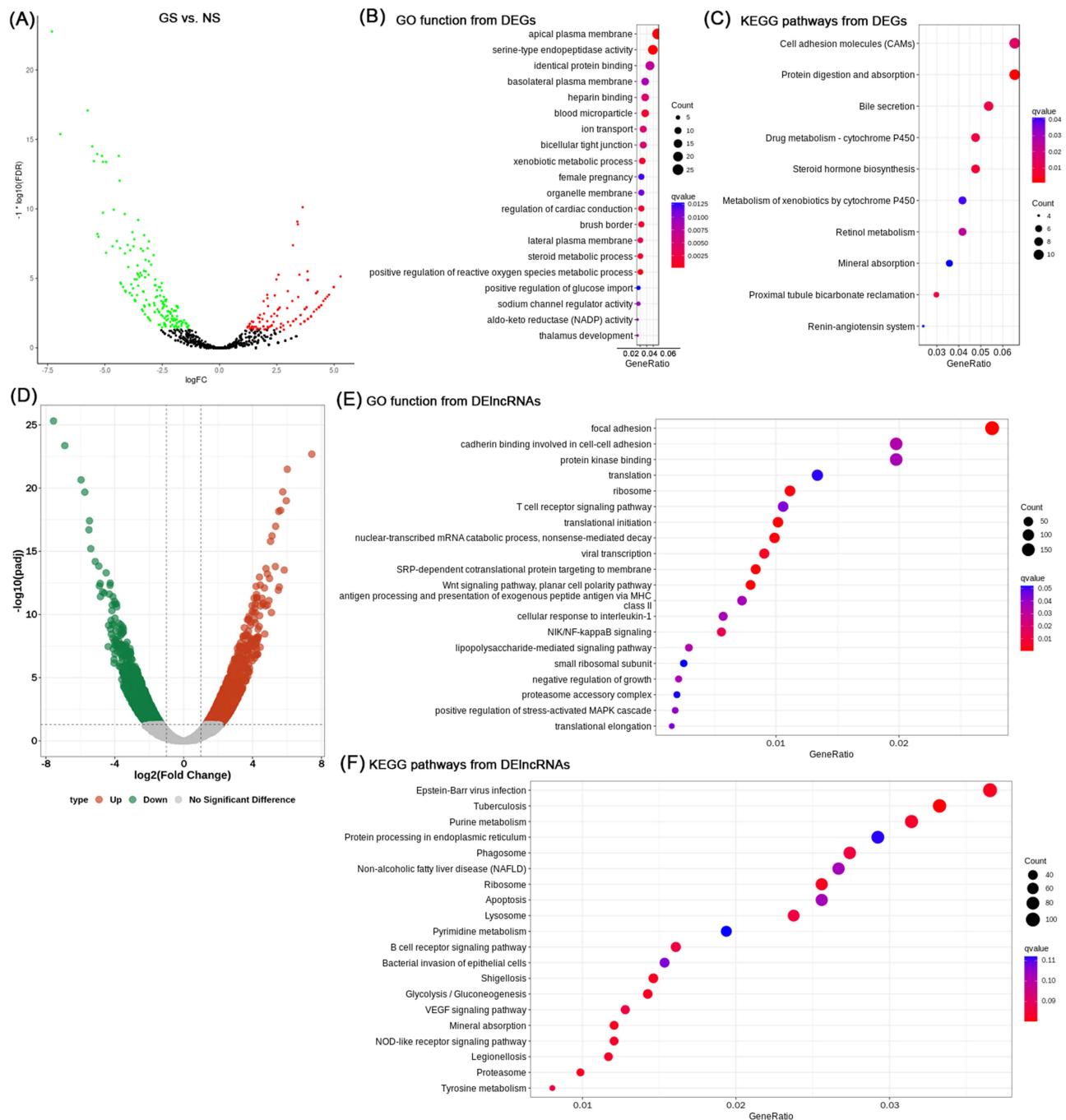


Figure 3 Differences in the lncRNA and mRNA expression profiles between GS and NS tissues. **(A)** Volcano plot of the differential expression analysis of DEGs (fold change ≥ 2.0 and $p < 0.05$). The top 15 items of DEGs of GBC-GS were identified by **(B)** GO function enrichment and **(C)** KEGG enrichment. **(D)** Volcano plot of the differential expression analysis of DElncRNAs (fold change ≥ 2.0 and $p < 0.05$). The top 15 items of DElncRNAs of GBC-nGS were identified by **(E)** GO function enrichment and **(F)** KEGG enrichment. **Abbreviations:** NS, normal gallbladder; GS, gallstone; DEG, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes pathway.

These results suggest that the expression profile of GS may be more similar to para-tumor tissues in transcriptome level and lower risk of cancer.

Identification of lncRNA and mRNA Expression Profiles Between GBC-GS and GBC-nGS

A total of 915 DEGs (408 upregulated and 507 downregulated) and 3383 DElncRNAs (1824 upregulated and 1559 downregulated) were identified between GBC-GS and GBC-nGS tissue (Figure 4A and D). The top 20 DEGs and DElncRNAs are shown in Table 4. Although DEGs were from GBC, the number of DEGs (GBC-GS vs GBC-nGS)

Table 3 Next-Generation Sequencing Analysis of the 20 Most Aberrantly Regulated mRNAs and lncRNAs Between Gallstone (GS) and Normal (NS)

Genes	GS/NS logFC	P value	lncRNAs	GS/NS logFC	P value
AKR1B10	-6.497	5.46E-15	CDK5RAP3	-7.559	2.51E-30
AMBP	-6.213	3.68E-14	DNAH1	-6.904	4.46E-28
LGALS2	-6.036	1.19E-13	TMEM86B	7.442	3.17E-27
GSTAI1	-5.887	3.25E-13	CLU	6.018	6.64E-26
SLC28A2	-5.913	3.30E-13	HLA-DQA1	-5.957	5.80E-25
SMIM24	-5.830	5.69E-13	MSTRG.5404	5.754	6.11E-24
HLA-DRB5	5.692	1.26E-12	MSTRG.20776	-5.741	7.58E-24
TNFRSF11B	5.441	7.32E-12	RPI1-30P6.6	5.956	4.14E-23
ABCC2	-5.371	1.15E-11	COPG1	5.630	2.70E-22
SLC9A3	-5.298	1.71E-11	MSTRG.13552	5.522	3.56E-22
FGF19	-5.247	2.22E-11	TGFBI	-5.469	2.17E-21
CYP4F3	-5.207	3.13E-11	OAZ1	5.341	6.55E-21
AKR7A3	-5.145	4.61E-11	KRT17	-5.501	1.34E-20
C11orf96	5.120	6.24E-11	FOSB	5.132	4.33E-20
SELE	5.040	1.03E-10	PAH	5.044	1.26E-19
VTN	-4.955	1.45E-10	IPO4	-5.393	5.11E-19
ANO3	-4.928	2.28E-10	CWC15	-5.136	5.69E-18
CA9	-4.885	2.66E-10	KMT5A	-4.928	1.38E-17
PP7080	-4.728	6.33E-10	MSTRG.24393	5.323	1.57E-17
AKR1B10	-6.497	5.46E-15	RPLP0	4.777	2.39E-17

remains relatively high, similar to DEGs (nGS vs GS). It suggests that GBC with and without gallstones may exhibit higher differences. Interestingly, several DEGs (GBC-GS vs GBC-nGS) overlap with DEGs (nGS vs GS), rather than DEGs (nGS vs GBC-nGS). This suggests that gallstones may act independently of GBC.

For DEGs, there were 333 GO items ($p < 0.05$), which were distinctly related to complement activation, antigen binding, serine-type endopeptidase activity and receptor-mediated endocytosis (Figure 4B). 37 KEGG items ($p < 0.05$) were enriched (Figure 4C). The unique DElncRNAs were associated with 33 GO items, including focal adhesion, protein transport, ribosome, cell-cell adherens junction et al (Figure 4E). 31 KEGG items ($p < 0.05$) was observed, including microRNA in cancer, Endocytosis, Spliceosome, Ribosome, Protein digestion and absorption, Glycolysis signaling pathway et al (Figure 4F).

Comparing DEGs (nGS vs GBC-nGS) and DEGs (nGS vs GS), as well as DElncRNAs, KEGG pathways (GBC-nGS vs GBC-GS) take into account the characteristics of both. Some pathways are related to cancers, like Transcriptional misregulation in cancers, Protein digestion and absorption, ECM-receptor interaction and microRNAs in cancers. Meanwhile, some pathways are related to gallstone metabolism, such as Bile secretion, Rennin secretion, and Tryptophan metabolism. This implies that there may be two aspects of differences between GBC-GS and GBC-nGS. We are more inclined to believe that the gallstone may not be associated with progression of GBC.

Co-DEGs and Co-DElncRNAs Among Three Lists

Our study aims to identify molecular signatures between GBC-GS and GBC-nGS to understand the possible mechanisms through which GS progresses to GBC. There were 29 overlapping DEGs and 552 overlapping DElncRNAs identified among three lists (Figure 5A and D).

The Gene Ontology analysis demonstrated co-DEGs were enriched biological processes like Cell cycle, DNA packaging, DNA packaging, Protein binding, and Cytoskeleton (Figure 5B). KEGG pathways analysis enriched pathways like alcoholism, systemic lupus erythematosus, and homologous recombination (Figure 5C). Co-DElncRNAs were also enriched in Negative regulation of gene expression, epigenetic, rDNA heterochromatin assembly, nucleosome and rDNA heterochromatin assembly and rDNA heterochromatin assembly, whereas the DElncRNAs were mainly linked with

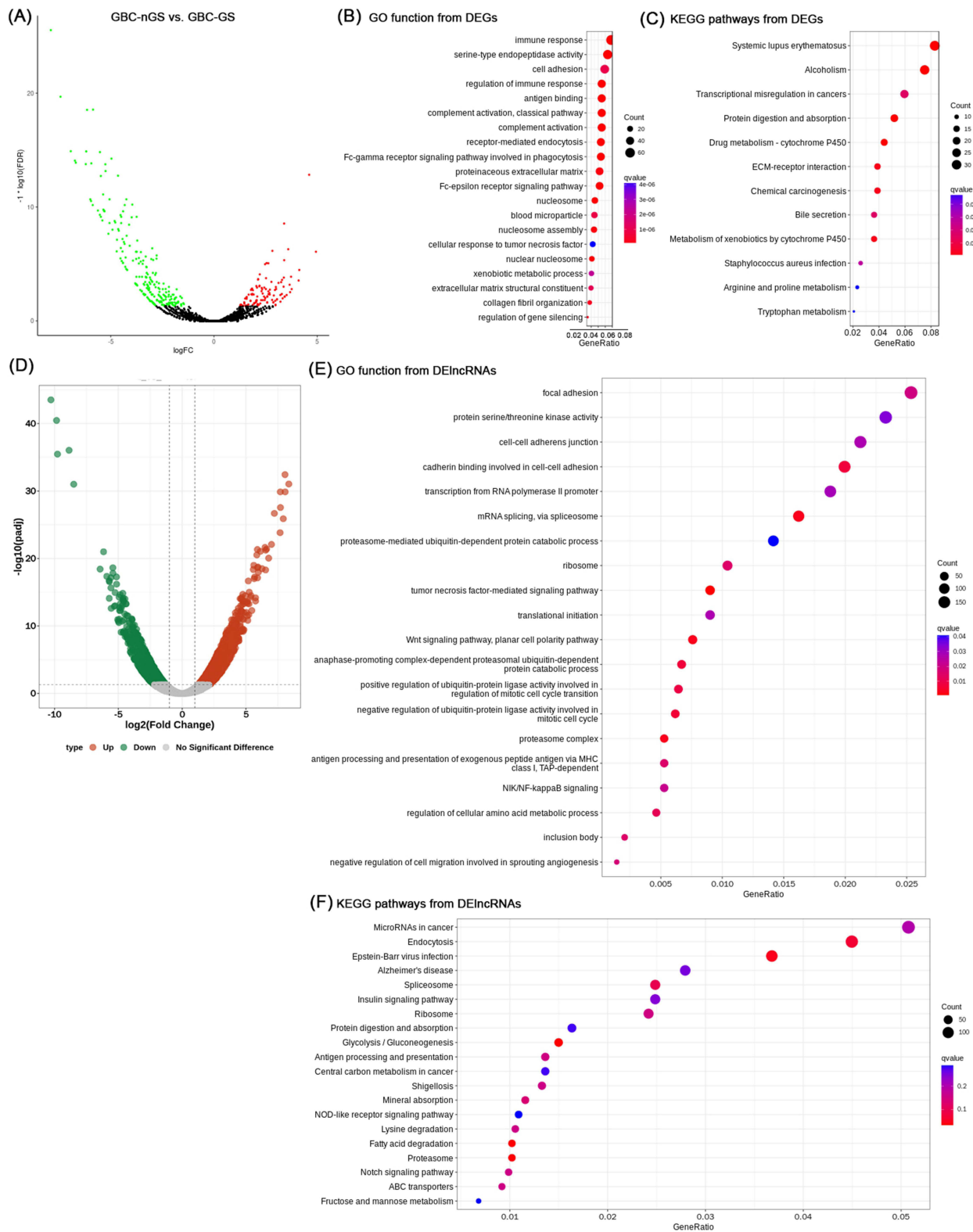


Figure 4 Differences in the lncRNA and mRNA expression profiles between GBC-nGS and GBC-GS tissues. **(A)** Volcano plot of the differential expression analysis of DEGs (fold change ≥ 2.0 and $p < 0.05$). The top 15 items of DEGs of GBC-GS were identified by **(B)** GO function enrichment and **(C)** KEGG enrichment. **(D)** Volcano plot of the differential expression analysis of DElncRNAs (fold change ≥ 2.0 and $p < 0.05$). The top 15 items of DElncRNAs of GBC-nGS were identified by **(E)** GO function enrichment and **(F)** KEGG enrichment.

Abbreviations: GBC-GS, gallbladder cancer tissues with gallstone; GBC-nGS, gallbladder carcinoma without gallstone; DEG, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes pathway.

Table 4 Next-Generation Sequencing Analysis of the 20 Most Aberrantly Regulated mRNAs and lncRNAs Between Gallbladder Carcinoma Without Gallstone (GBC-GS) and Withgallstone (GBC-nGS)

Genes	logFC	P value	lncRNAs	logFC	P value
TM4SF4	-8.231	3.14E-20	MSTRG.20766	-10.284	1.50E-48
OLFM4	8.240	3.51E-20	KRT17	-9.841	3.34E-45
PAH	-8.176	6.20E-20	HLA-DQA1	-8.867	1.33E-40
HLA-DRB5	8.159	7.59E-20	MSTRG.7518	-9.772	6.64E-40
AKR1B10	-7.865	5.28E-19	PAH	8.062	8.97E-37
GC	-7.371	1.22E-17	RP11-30P6.6	8.359	2.67E-35
COL11A1	7.277	3.21E-17	MSTRG.8672	-8.497	3.37E-35
IGHG4	7.127	8.63E-17	MSTRG.2633	7.694	5.71E-34
ALDOB	-7.057	1.10E-16	PAH	8.074	5.85E-34
SLC3A1	-6.875	4.67E-16	MSTRG.2627	7.703	1.34E-31
GSTA1	-6.743	9.63E-16	MSTRG.2627	7.228	1.10E-30
UNC80	-6.768	1.16E-15	MSTRG.30924	7.925	7.42E-30
ABCC2	-6.705	1.60E-15	PAH	7.676	9.46E-28
LGALS2	-6.601	2.75E-15	ATP7B	7.002	5.62E-26
MUCL3	-6.560	3.99E-15	AKR1C2	6.493	1.66E-25
B3GALT2	-6.533	4.37E-15	FXVD2	6.618	3.08E-25
IGHG3	6.471	5.35E-15	MSTRG.20766	5.899	3.86E-25
IGKVID-33	6.393	1.09E-14	MSTRG.14771	6.552	5.09E-25
PPP1R1B	6.398	1.25E-14	TMEM94	6.131	8.07E-25
TM4SF4	-8.231	3.14E-20	PTK7	-6.159	9.79E-25

Systemic lupus erythematosus and alcoholism pathway (Figure 5E and F). Interestingly, DEGs are enriched in immune-related pathways, while DElncRNAs are enriched in cancer-related pathways, such as pancreatic cancers, Non-small cell lung cancer, AMPK signaling pathway, etc.

Next, we mainly focused on DEGs and DElncRNAs with enlarged expression fold-change trends in GBC-GS group versus GBC-nGS group. We found three genes were co-expressed in lncRNAs and mRNAs, including FXVD2, MPZL1, PAH (Figure 5G). Then, we used qPCR to examine their levels in NS, GS, GBC-GS and GBC-nGS groups. The results showed that FXVD2 and PAH were upregulated, while MPZL1 was downregulated between GBC-GS and NS or GBC-nGS group. It is consistent with RNA-sequencing data (Figure 5H).

These results suggest that FXVD2, MPZL1, and PAH may be important mediators between immunity and gallbladder cancer. Additionally, these factors may also serve as bridges between immunity and gallstones.

Discussion

Although without the function of encoding protein, lncRNA was reported to play an important role in the progression of tumors, including gastric cancer,²⁰ hepatocellular cancer,²¹ colorectal cancer²² and pancreatic cancer.²³ However, few studies have focused on the role of lncRNAs in the progression of gallstones to gallbladder cancer. In this study, the differential expressed lncRNA and mRNA were screened in gallbladder cancer patients with gallstone and gallbladder carcinoma without gallstone. Under the premise of good quality control, we obtained a total of 29 DEGs and 500 DElncRNAs. Using the obtained DElncRNAs and DEGs, a co-expression network of DEGs was constructed. The DEGs were significantly enriched in GO function, including focal adhesion cell-cell adhesion, ribosome, spliceosome, while enriched in KEGG pathway, including Transcriptional misregulation in cancers, Protein digestion and absorption, ECM-receptor interaction, bile secretion signaling pathways.

Previous studies have used microarray chips to study the expression profiles of DEGs and DElncRNA in GBC.²⁴ Compared with matched para-cancerous tissues, the level of more than 1000 lncRNA expression changed.²⁵ Our data screened more known DElncRNAs and novel DElncRNAs. The discrepancy between the data of these studies may be

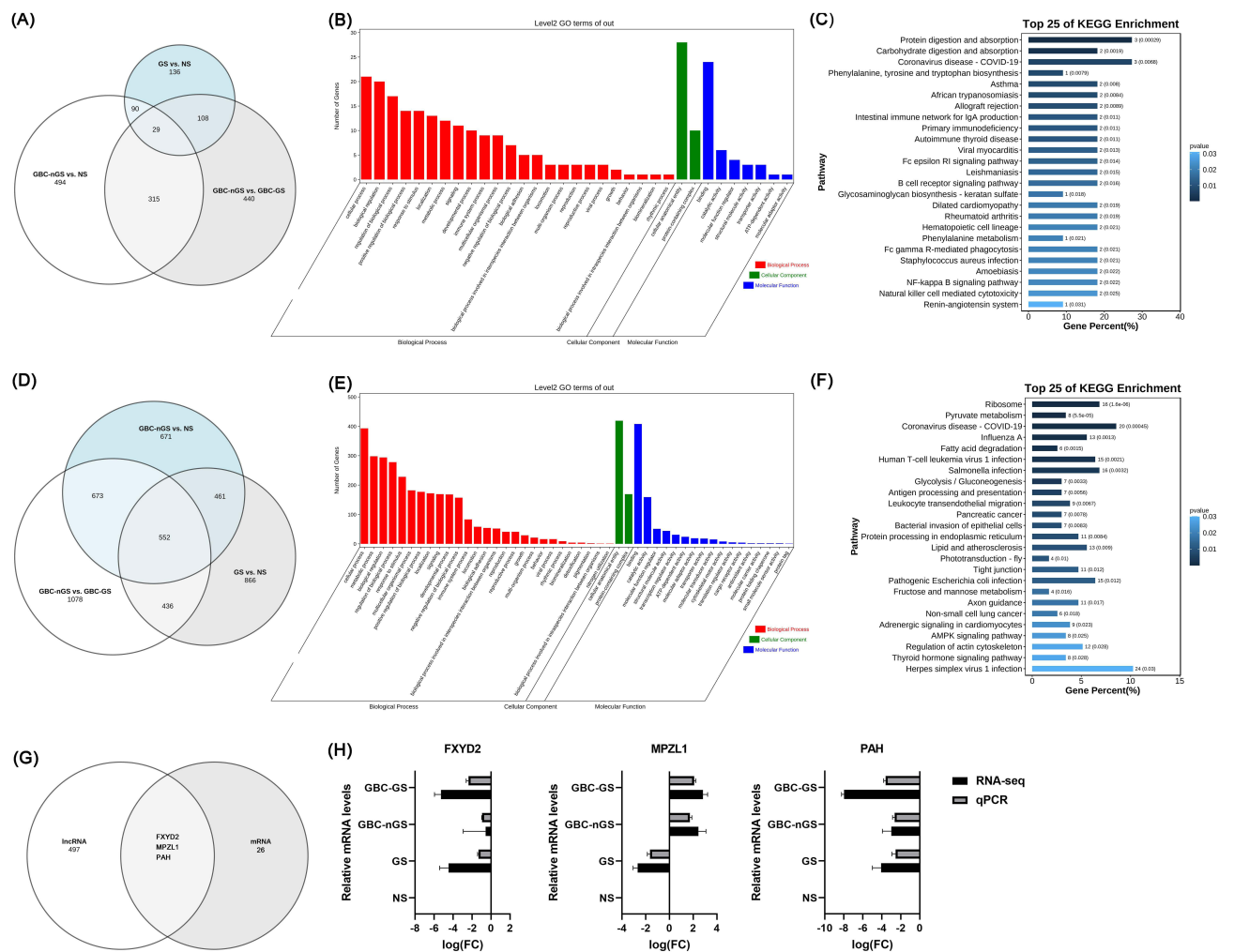


Figure 5 Differential gene and lncRNA expression profiles of overlapping and unique signatures. Venn diagram showing the number of overlapping and overlapping (A) DEGs and (B) DElncRNAs between four groups. The top 25 items of co-DEGs were identified by (C) GO function enrichment and (D) KEGG enrichment. The top 25 items of co-DElncRNAs were identified by (E) GO function enrichment and (F) KEGG enrichment. (G) Venn diagram showing the number of overlapping between co-DEGs and co-DElncRNAs predicted genes. (H) FXYD2, MPZL1, PAH were determined to be differentially expressed in four groups compared with NS group using next-generation sequencing analysis and validated using RT-qPCR in tissues.

Abbreviations: RT-qPCR, reverse transcription-polymerase chain reaction. GBC-nGS, gallbladder cancer tissues without gallstone; GBC-GS, gallbladder cancer tissues with gallstone; NS, normal gallbladder; GS, gallstone; DEG, differentially expressed genes.

due to different kinds of tumor tissues and different platforms used. The platform that was used as the Human lncRNA array V4.0 (4×180K) in Kong et al's study,²⁴ whereas the platform for the GEO data was the Affymetrix Human Gene 2.0 ST Array in Wang et al's study.²⁵ Our study is based on next-generation sequencing data. High-throughput ncRNA sequencing screened more new lncRNAs, which avoided the omission of lncRNA which plays an important role in GBC.

We identified 947 DEGs (NS vs GBC-nGS), 372 DEGs (NS vs GS), and 915 DEGs (GBC-nGS vs GBC-GS) respectively. The numbers reflect that gallstones may be more inclined towards para-tumor tissues and contribute less to gallbladder cancer. Gallstones are very common but gallbladder cancer is very rare. Most people with an inflamed gallbladder or gallstones do not get gallbladder cancer.²⁶ Our study indirectly supports the conclusions regarding gallstones not related to gallbladder cancer. Trans-regulation and cis-regulation are important regulatory mechanisms for lncRNAs.²⁷ We utilized this mechanism to predict target genes of DElncRNAs, together with DEGs followed by GO and KEGG function analysis. Our data indicates that lncRNA transcription undergoes significant changes during the process of GBC, consistent with previous studies.^{28,29} DEGs and DElncRNAs (NS vs GBC-nGS) were enriched in cancer-related pathways, while DEGs and DElncRNAs (NS vs GS) were enriched in metabolism-related pathways. Studies reported a close association between gallstones and processes such as bile secretion and transport.³⁰ It reflects the dual spectrum characteristics of DEGs and DElncRNAs (GBC-nGS vs GBC-GS),

which exhibit both tumor and stone-related profiles. This suggests that gallstones and gallbladder cancer may exist independently rather than in a promoting relationship.

Several shared signaling pathways between gallstones and gallbladder cancers are worth our attention, including ECM-receptor interaction pathways, transcription misregulation, and Protein digestion and absorption. ECM-receptor interaction pathway plays an important role in the process of tumor shedding, adhesion, degradation, movement and hyperplasia.³¹ Accumulated evidence has confirmed ECM-receptor interaction pathway is associated with several cancers, including breast cancer,³¹ kidney cancer,³² and colorectal cancer.³³ In gallbladder cancer, Kong et al showed mRNAs and lncRNAs may be involved in gallbladder cancer progression via ECM-receptor interaction pathways and the complement and coagulation cascades.²⁴ Chen et al compared the glandular epithelial between normal and tumor tissues and found that extracellular matrix (ECM)-receptor interaction plays an important role in gallbladder cancer progression.³⁴ In our study, upregulated DEGs and DElncRNAs were significantly enriched in the ECM-receptor interaction pathway, which possibly promoted the development of gallbladder cancer. Transcription misregulation arises in cancer from disease-defining genetic alterations either indirectly, via mutation of signaling factors converging on transcriptional control, or directly, via genetic alterations in gene control factors themselves.³⁵ Nepal et al' study showed that in gallbladder cancer, transcriptional levels appear to be affected by mutations, as indicated by the enrichment of mutations for the 'Transcriptional misregulation in cancer' pathway.³⁶ We agree with this viewpoint. In this study, we also found that transcription misregulation is significantly upregulated in GBC-GS. However, direct or indirect effects requires further validation.

Interestingly, the pathway "Protein digestion and absorption" is presented in both cancer and gallstones. Protein digestion and absorption have been consistently found to be involved in digestive system cancers from RNA-sequencing studies.³⁷ However, it has been identified in gallbladder cancer for the first time, including the FXVD2 and the Collagen family. Collagen is the major component of the tumor microenvironment and participates in cancer fibrosis. In gallstone, the total and insoluble collagen contents were significantly high in gallstone patients. The activity levels of collagenase and protease were significantly low.³⁸ In this study, Protein digestion and absorption participated in the progress of both gallstone and gallbladder cancer. In addition, DElncRNAs (NS vs GBC-nGS) and DElncRNAs (GBC-GS vs GBC-nGS), but not DElncRNAs (NS vs GS). were enriched in the pathway "MicroRNAs in cancers". It indicated gallbladder cancer was related to miRNAs. Previous studies have shown that lncRNAs can act as competitive endogenous RNAs (ceRNAs), affecting the function of miRNAs through sponge effects and participating in tumor progression. Our results indirectly support this view.

Finally, we analyzed DEGs and DElncRNAs related to both gallstones and tumors. We initially filtered out co-DEGs and co-DElncRNAs from all four groups. Subsequently, we overlapped the genes predicted by lncRNA cis-regulation with DEGs, and identified three genes that exhibited significant changes. For instance, FXVD2 showed a significant decrease in expression in GBC-nGS compared to NS, and the decrease was even more pronounced in GBC-GS compared to GBC-nGS. Similar patterns were observed for MPZL1 and PAH. Three genes were reported to relate to gallstone or GBC. Downregulation of FXVD2 is associated with poor prognosis in clear cell renal cell carcinoma.³⁹ Prevalence ratios (PR) of gallbladder disease are significantly higher for the Phenylalanine hydroxylase (PAH) deficiency cohort.⁴⁰ MPZL1 is highly expressed in advanced gallbladder carcinoma and promotes the aggressive behavior of human gallbladder carcinoma GBC-SD cells.⁴¹ Although the dynamic changes in these genes did not directly prove the role of gallstones in GBC, we found that these coDEGs are more closely associated with immunity. Previous studies have shown that immunity may play an important role in both GS and GBC,^{42,43} potentially serving as a bridge between the two.

This study still has some limitations. (1) Due to the nature of gallbladder disease, we were unable to complete follow-up processes from gallstones to GBC in the same patient. Therefore, the role of GS in GBC cannot be confirmed. (2) Due to the low incidence of GBC, it is also not possible to obtain sufficient GBC patients with different periods of follow-up gallstones. (3) Although we have used reliable bioinformatics tools to identify some DEGs and DElncRNAs that may be related to gallbladder cancer, these have not been confirmed by experimental evidence in this study. We will further verify the lncRNA-mRNA network in the process of GBC in vitro and in vivo studies.

Conclusions

In summary, we profiled the transcriptome landscape of GBC with/without GS using RNA sequencing and identified three genes (FXVD2, MPZL1 and PAH) and Transcriptional misregulation in cancers, Protein digestion and absorption,

ECM-receptor interaction signaling pathways, which may contribute to the development of gallbladder cancer. The differential expressed trend and diversity of DEGs and DElncRNAs indicated GS profile was more similar with para-tumor tissues in transcriptome level and lower risk of cancer. Further exploration is necessary from GBC patients with different periods of follow-up gallstone.

Data Sharing Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

This study was approved by the First Affiliated Hospital of Kunming Medical University Ethics Committee (First Affiliated Hospital of Kunming Medical University, Kunming Medical University, Kunming, China), and All patients provided written informed consent for the use of their surgical specimens in the present study in accordance with the Committee's regulations. This study was conducted in accordance with the Declaration of Helsinki.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 81272728); the Applied Basic Research Foundation of Yunnan Province, China (Grant #2018FE001(-204)); Yunnan Applied Basic Research Projects-Kunming Medical University Union Foundation (202001AY070001-040); Yunnan Provincial Fund for High-Level Reserve Talents in Health Science (H-2018068).

Author Contributions

Bin Yang and Qiang Wang conceived and designed the experiments and manuscript preparation; Pinduan Bi, Ding Luo, and Pingli Cao performed the experiments and analysis the data; Qiang Wang, Bin Yang and Weihong Chen proofread the manuscript. All authors discussed the results and contributed to the final manuscript. All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Song X, Hu Y, Li Y, Shao R, Liu F, Liu Y. Overview of current targeted therapy in gallbladder cancer. *Signal Transduc Tar Thera*. 2020;5(1):1–19. doi:10.1038/s41392-019-0089-y
2. García P, Lamarca A, Díaz J, Carrera E, Roa JC, Consortium E-LAE. Current and New Biomarkers for Early Detection, Prognostic Stratification, and Management of Gallbladder Cancer Patients. *Cancers*. 2020;12(12):3670. doi:10.3390/cancers12123670
3. Rawla P, Sunkara T, Thandra KC, Barsouk A. Epidemiology of gallbladder cancer. *Clinical Exper Hepa*. 2019;5(2):93. doi:10.5114/ceh.2019.85166
4. Jayaraman S, Jarnagin WR. Management of gallbladder cancer. *Gastroenterol Clinics*. 2010;39(2):331–342. doi:10.1016/j.gtc.2010.02.006
5. Zaidi MY, Maithel SK. Updates on gallbladder cancer management. *Current Onco Repo*. 2018;20(2):1–7. doi:10.1007/s11912-018-0664-3
6. Kim JH, Kim HN, Lee KT, et al. Gene expression profiles in gallbladder cancer: the close genetic similarity seen for early and advanced gallbladder cancers may explain the poor prognosis. *Tumor Biol*. 2008;29(1):41–49. doi:10.1159/000132570
7. Bai Y, Liu Z-S, Xiong J-P, et al. Nomogram to predict overall survival after gallbladder cancer resection in China. *World J Gastroenterol*. 2018;24(45):5167. doi:10.3748/wjg.v24.i45.5167
8. Muhammad JS, Khan MR, Ghias K. DNA methylation as an epigenetic regulator of gallbladder cancer: an overview. *Int j Surg*. 2018;53:178–183. doi:10.1016/j.ijsu.2018.03.053
9. Abdol Razak NB, Jones G, Bhandari M, Berndt MC, Metharom P. Cancer-associated thrombosis: an overview of mechanisms, risk factors, and treatment. *Cancers*. 2018;10(10):380. doi:10.3390/cancers10100380
10. Sharma A, Sharma KL, Gupta A, Yadav A, Kumar A. Gallbladder cancer epidemiology, pathogenesis and molecular genetics: recent update. *World J Gastroenterol*. 2017;23(22):3978. doi:10.3748/wjg.v23.i22.3978

11. Liu S, Yang Y, Jiang S, et al. Understanding the role of non-coding RNA (ncRNA) in stent restenosis. *Atherosclerosis*. 2018;272:153–161. doi:10.1016/j.atherosclerosis.2018.03.036
12. Gibb EA, Brown CJ, Lam WL. The functional role of long non-coding RNA in human carcinomas. *Mol Cancer*. 2011;10(1):1–17. doi:10.1186/1476-4598-10-38
13. Goodrich JA, Kugel JF. Non-coding-RNA regulators of RNA polymerase II transcription. *Nat Rev Mol Cell Biol*. 2006;7(8):612–616. doi:10.1038/nrm1946
14. Yang Y, Yujiao W, Fang W, et al. The roles of miRNA, lncRNA and circRNA in the development of osteoporosis. *Biol Res*. 2020;53(1):1–16. doi:10.1186/s40659-020-00309-z
15. Sun W, Ren S, Li R, Zhang Q, Song H. LncRNA, a novel target biomolecule, is involved in the progression of colorectal cancer. *American J Cancer Res*. 2019;9(11):2515.
16. Tu C, Yang K, Wan L, et al. The crosstalk between lncRNAs and the Hippo signalling pathway in cancer progression. *Cell Proliferation*. 2020;53(9):e12887. doi:10.1111/cpr.12887
17. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS*. 2012;16(5):284. doi:10.1089/omi.2011.0118
18. Szklarczyk D, Franceschini A, Kuhn M, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res*. 2010;39(suppl_1):D561–8. doi:10.1093/nar/gkq973
19. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: network analysis and visualization of proteomics data. *J Proteomics*. 2018;18(2):623–632. doi:10.1021/acs.jproteome.8b00702
20. Hao N-B, Y-F H, X-Q L, Wang K, Wang R-L. The role of miRNA and lncRNA in gastric cancer. *Oncotarget*. 2017;8(46):81572. doi:10.18632/oncotarget.19197
21. Yang Y, Chen L, Gu J, et al. Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat Commun*. 2017;8(1):1–13. doi:10.1038/s41467-016-0009-6
22. Yuan W, Li X, Liu L, et al. Comprehensive analysis of lncRNA-associated ceRNA network in colorectal cancer. *Biochem Biophys Res Commun*. 2019;508(2):374–379. doi:10.1016/j.bbrc.2018.11.151
23. Huang X, Zhi X, Gao Y, Ta N, Jiang H, Zheng J. LncRNAs in pancreatic cancer. *Oncotarget*. 2016;7(35):57379. doi:10.18632/oncotarget.10545
24. Kong L, Wu Q, Zhao L, Ye J, Li N, Yang H. Identification of messenger and long noncoding RNAs associated with gallbladder cancer via gene expression profile analysis. *J Cell Biochem*. 2019;120(12):19377–19387. doi:10.1002/jcb.28953
25. Wang J, Liu H, Shen X, et al. Long non-coding RNA expression profiles in gallbladder carcinoma identified using microarray analysis. *Oncol Lett*. 2017;13(5):3508–3516. doi:10.3892/ol.2017.5893
26. Roa JC, García P, Kapoor VK, Maithel SK, Javle M, Koshiol J. Gallbladder cancer. *Nature Rev Dis Prim*. 2022;8(1):69. doi:10.1038/s41572-022-00398-y
27. Mas AM, Huarte M. lncRNA–DNA hybrids regulate distant genes. *EMBO Rep*. 2020;21(3):e50107. doi:10.15252/embr.202050107
28. Yang S, Qin L, Wu P, et al. RNA sequencing revealed the multi-stage transcriptome transformations during the development of gallbladder cancer associated with chronic inflammation. *PLoS One*. 2023;18(3):56.
29. Roy N, Dhingra BR, Barah P. Integrative network-based approaches identified systems-level molecular signatures associated with gallbladder cancer pathogenesis from gallstone diseases. *J Biosci*. 2022;47(2):31. doi:10.1007/s12038-022-00267-6
30. Dosch AR, Imagawa DK, Jutric Z. Bile metabolism and lithogenesis: an update. *Surg Clin*. 2019;99(2):215–229. doi:10.1016/j.suc.2018.12.003
31. Bao Y, Wang L, Shi L, et al. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. *Cell Mol Biol Lett*. 2019;24(1):1–20. doi:10.1186/s11658-019-0162-0
32. Zhang H-J, Tao J, Sheng L, et al. Twist2 promotes kidney cancer cell proliferation and invasion by regulating ITGA6 and CD44 expression in the ECM-receptor interaction pathway. *Oncotargets Ther*. 2016;9:1801. doi:10.2147/OTT.S96535
33. Machackova T, Vychytilova-Faltejskova P, Souckova K, et al. MiR-215-5p Reduces Liver Metastasis in an Experimental Model of Colorectal Cancer through Regulation of ECM-Receptor Interactions and Focal Adhesion. *Cancers*. 2020;12(12):3518. doi:10.3390/cancers12123518
34. Chen P, Wang Y, Li J, et al. Diversity and intratumoral heterogeneity in human gallbladder cancer progression revealed by single-cell RNA sequencing. *Clinical Transl Med*. 2021;11(6):e462. doi:10.1002/ctm2.462
35. Bradner JE, Hnisz D, Young RA. Transcriptional addiction in cancer. *Cell*. 2017;168(4):629–643. doi:10.1016/j.cell.2016.12.013
36. Nepal C, Zhu B, O'Rourke CJ, et al. Integrative molecular characterisation of gallbladder cancer reveals micro-environment-associated subtypes. *J Hepatol*. 2021;74(5):1132–1144. doi:10.1016/j.jhep.2020.11.033
37. Chen Z, Shen Z, Zhang Z, Zhao D, Xu L, Zhang L. RNA-associated co-expression network identifies novel biomarkers for digestive system cancer. *Front Genetics*. 2021;12:659788. doi:10.3389/fgene.2021.659788
38. Geetha A. Biochemical and physicochemical changes in collagen isolated from the gall bladder of gall stone patients. *J Bioc Mol Bio Biop*. 2002;6(6):421–425. doi:10.1080/1025814021000036160
39. Zhang Z, Tang Y, Li L, et al. Downregulation of FXRD2 Is Associated with Poor Prognosis and Increased Regulatory T Cell Infiltration in Clear Cell Renal Cell Carcinoma. *J Immunol Res*. 2022;2022:1.
40. Burton BK, Jones KB, Cederbaum S, et al. Prevalence of comorbid conditions among adult patients diagnosed with phenylketonuria. *Mol Gene Metabol*. 2018;125(3):228–234. doi:10.1016/j.ymgme.2018.09.006
41. Liu X, Huang J, Liu L, Liu R. MPZL1 is highly expressed in advanced gallbladder carcinoma and promotes the aggressive behavior of human gallbladder carcinoma GBC-SD cells. *Molecu Med Rep*. 2019;20(3):2725–2733. doi:10.3892/mmr.2019.10506
42. Maurer KJ, Carey MC, Fox JG. Roles of infection, inflammation, and the immune system in cholesterol gallstone formation. *Gastroenterology*. 2009;136(2):425–440. doi:10.1053/j.gastro.2008.12.031
43. Espinoza JA, Bizama C, García P, et al. The inflammatory inception of gallbladder cancer. *Biochi Bio Acta Rev Cancer*. 2016;1865(2):245–254. doi:10.1016/j.bbcan.2016.03.004

International Journal of General Medicine

Dovepress

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>