

# LncRNA expression profile, functional analysis and potential biomarkers in children with pancreaticobiliary maljunction

**Lian Zhao**

Soochow University Affiliated Children's Hospital

**San-li Shi**

Xi'an Jiaotong University Second Affiliated Hospital

**Wan-liang Guo** (✉ [gwsuzhou@163.com](mailto:gwsuzhou@163.com))

Xi'an Jiaotong University Second Affiliated Hospital

---

## Research article

**Keywords:** pancreaticobiliary maljunction, long non-coding RNAs, gene ontology, biomarker

**Posted Date:** May 27th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-558683/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

## Background

The present study was aimed to investigate the expression profile of long non-coding RNAs (lncRNAs) and its potential as biomarker for pancreaticobiliary maljunction (PBM).

## Methods

The differential expression of lncRNA and messenger RNA (mRNA) from 15 pediatric patients with PBM and 15 control subjects were analyzed with microArray and validated with quantitative reverse-transcription polymerase chain reaction (qRT-PCR). Gene Ontology enrichment and Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis were used to investigate the biological functions of these genes. The area under curve (AUC) of receiver operating characteristic (ROC) curve was used to predict the biomarker of lncRNA in PBM.

## Results

There were 2915 mRNAs and 173 lncRNAs upregulated, 2121 mRNAs and 316 lncRNAs downregulated in PBM cases, respectively. The enriched GO associated with differentially expressed mRNA were extracellular matrix, extracellular region and kinetochore. The most enriched pathway of Protein digestion and absorption was associated with cancer and PI3K-Akt signaling. Cis- and Trans-target gene analysis indicated that mRNAs were predicted to be regulated by one lncRNA and one mRNA corresponded to several lncRNAs. The results showed that the expression trends of *NR\_110876*, *NR\_132344*, *XR\_946886* and *XR\_002956345* were consistent with the microarray results, and the difference was statistically significant *NR\_132344*, *XR\_946886* and *XR\_002956345* ( $P < 0.05$ ). The AUC of ROC curve was significant only for *XR\_946886* (0.837,  $P < 0.001$ ).

## Conclusion

The present study indicated that lncRNAs were involved the pathogenesis of common bile duct in PBM and *XR\_946886* would be biomarkers for PBM.

## Background

Pancreaticobiliary maljunction (PBM) is a rare congenital anomaly observed in clinical practice, in which the junction of pancreatic and the biliary ducts is located outside the duodenal wall [1, 2, 3]. Due to two-way reflux of bile and pancreatic juice in PBM, bile retention with cholangitis, pancreatitis and malignancies are commonly associated complications in PBM [4, 5, 6].

Cyst excision and Roux-en-Y hepaticojejunostomy is the first choice for PBM treatment in clinic [7, 8, 9]. However, the postoperative problems in bile duct, such as cholangitis, hepatolithiasis and carcinogenesis at the residual bile ducts (8,10,11,12,13) are worthy of attention. Thus, it is urgent to investigate the biomarkers that link PBM to the related complications of bile duct, in order to early detect the complications in PBM and develop effective diagnostic and therapeutic methods.

Long noncoding RNAs (lncRNAs), a group of non-coding RNAs longer than 200 nt, play an important role in regulating the occurrence and progression of many diseases, such as infectious and malignant tumors [14, 15, 16], through modulating biological processes, such as cellular proliferation, motility, immune and inflammation response to diseases [17, 18, 19]. In addition, some lncRNAs has been identified as biomarkers for diagnosis and prognosis of diseases [20, 21]. However, the genome-wide expression and functional roles of lncRNAs in PBM is unclear.

To the best of our knowledge, there has been no publication focused on lncRNAs in common bile duct of PBM and PBM-linked biologic markers. Here, we investigated the expression profiles, functional analysis and potential biomarkers of lncRNAs in the common bile duct of patients with PBM. Our findings would provide basis for developing novel diagnostic and therapeutic targets for PBM.

## Material And Methods

### 1. Study subjects

The present study was approved by the institutional review board of the Children's Hospital of Soochow University. Written informed consent was signed by the guardian of the subjects before surgery. A total of 15 subjects with PBM and 15 normal subjects as control were enrolled in

the present study. The expression of lncRNA and mRNA was analyzed with microarrays and verified with qRT-PCR. The clinical and pathological characteristics of the 15 patients were summarized in Table 1.

## 2. RNA Microarray analysis

Microarray experiments to investigate the differentially expressed genes between PBM and control group were performed by Western SCI Biotech Company, Chong qing, China (<http://www.westernsci.cn>). The Agilent Human lncRNA Microarray 2018Version (4\*180k, Design ID: 085630) was used in this experiment.

The tissue of common bile duct of PBM cases were collected and stored at  $-80^{\circ}\text{C}$  before RNA extraction. Total RNA was isolated using TRIZOL (Invitrogen, Carls-bad, CA, USA) according to the manufacturer's protocol and quantified by the NanoDrop ND-2000 (Thermo Scientific). The RNA integrity was assessed using Agilent Bioanalyzer 2100 (Agilent Technologies). The sample labeling, microarray hybridization and washing were performed based on the manufacturer's standard protocols. Briefly, total RNA was transcribed to double strand cDNA, then synthesized into cRNA and labeled with Cyanine-3-CTP. The labeled cRNAs were hybridized onto the microarray. After washing, the arrays were scanned by the Agilent Scanner G2505C (Agilent Technologies).

For data analysis, the Feature Extraction software (version 10.7.1.1, Agilent Technologies) was used to analyze the array images to get raw data. Genespring (version 13.1, Agilent Technologies) were employed to finish the basic analysis with the raw data. To begin with, the raw data was normalized with the quantile algorithm. The probes that at least 1 out of 2 conditions have flags in "P" were chosen for further data analysis. Differentially expressed genes or lncRNAs were then identified through fold change. The threshold set for up- and down-regulated genes was a fold change  $\geq 2.0$ . Afterwards, gene ontology (GO) analysis and KEGG analysis were applied to determine the roles of these differentially expressed mRNAs and lncRNAs. Finally, Hierarchical Clustering was performed to display the expression pattern of distinguishable genes among different samples.

## 3. LncRNA function prediction

For function prediction of lncRNAs, we adopted the method from previous study [22]. In brief, the co-expressed mRNAs for each differentiated lncRNAs were calculated and subjected to functional enrichment analysis. The enriched functional terms were used as the predicted functional term of a given lncRNA. The co-expressed mRNAs of lncRNAs were identified by calculating the Pearson correlation with correlation P-value  $< 0.05$ . The Hypergeometric cumulative distribution function was used to calculate the enrichment of functional terms in annotation of co-expressed mRNAs. The relationship between lncRNA and function prediction terms", the top 200 and top 500 prediction relationships with the highest prediction reliability were selected. The frequency of each functional prediction term was counted, and the GO or pathway term with more functional annotation was counted to reflect the overall situation of the functional distribution of different lncRNA. The top 20 significant lncRNAs were selected to draw a bar chart.

## 4. Fine mapping of genome coexpression of lncRNAs and adjacent coding genes

The Cis-regulatory regions were identified by the following procedures. For each lncRNAs, the mRNAs were identified as "Cis-regulated mRNAs" when: (1) the mRNAs loci were within 300k windows up- and down-stream of the given lncRNA, (2) the Pearson correlation of lncRNA-mRNA expression was significant (P-value of correlation  $\leq 0.05$ ).

For each differentially expressed lncRNAs, the co-expressed genes and the enrichment significance of differentially expressed gene in each transcription factor (TF) entry was calculated by hypergeometric distribution test. The calculated results will return a p value of enrichment significance, and a small P value indicates that the gene is enriched in the TF entry.

The intersection between the coding gene set co-expressed by lncRNAs and the target gene set of TF/chromatin regulatory complex was calculated, and the enrichment degree of the intersection was calculated by using hypergeometric distribution test. This calculation obtained the TF significantly related to lncRNAs and identified the TF/chromatin regulatory factors that may play a regulatory role with lncRNAs. And the network diagram visualized using the analysis results of hypergeometric distribution.

According to hypergeometric distribution calculation, multiple lncRNA-TF pairs were obtained from each lncRNA. Each lncRNA-TF pair was the result of multiple gene enrichment.

## 5. Quantitative Reverse Transcription-Polymerase Chain Reaction (qRT-PCR) Assay

In order to verify the microarray results, we randomly selected a total of 6 differently expressed lncRNAs that associate with carcinogenesis or chronic inflammation in common bile duct in PBM cases and measured the lncRNA expression levels of these genes by qRT-PCR. qRT-PCR was

performed using SYBR Green PCR Kit (Applied BI) according to the manufacturer's instructions. Gene expression was normalized to  $\beta$ -actin lncRNA. The relative expression of gene transcript was calculated using the  $2^{-\Delta\Delta Ct}$  method.

## 6. Statistical Analysis

SPSS 20.0 software was used for statistical analysis. The results of measurement data were expressed as mean  $\pm$  SD. The comparison between the groups was performed using the rank sum test. A fold change of 2.0 ( $P < 0.05$ ) was used as a cutoff for differential expression of lncRNAs and mRNAs screened by microarray. Receiver operator characteristic (ROC) curve analysis was conducted and area under the ROC curve (AUC) was calculated to evaluate the diagnostic accuracy of selected genes in PBM patients.  $P < 0.05$  was considered statistically significant.

# Results

## 1. General information

A total of 15 pediatric subjects with PBM and 15 control subjects were included in the study. Demographic and clinical data were summarized in Table 1. There was no significant difference between 2 groups in body weight, age, gender or other general information.

## 2. Differential gene expression between subjects with PBM and control subjects

From the microarray data, there were 2915 mRNAs upregulated, 2121 mRNAs downregulated, 173 lncRNAs upregulated and 316 lncRNAs downregulated in subjects with PBM (cut-off = 2.0). Principal Component

Analysis (PCA) results of differentiated lncRNA and heatmap were presented in figure 1. The top 20 upregulated and downregulated lncRNAs were presented in Tables 2. The most upregulated lncRNA was LOC105378608 with fold change of 14.39.

## 3. Functional analysis of differentially expressed genes

To avoid omission of any PBM-related genes, we performed functional analyses for all the differentially expressed mRNAs. It was found that the most enriched GOs associated with the differentially expressed transcripts were extracellular matrix ( $p=2.9E-12$  up-regulated), extracellular region ( $p=0.000427$ , up-regulated), and kinetochore ( $p=2.29E-8$ , up-regulated) (Figure 2). The most enriched pathway was Protein digestion and absorption which could be related to 250 differentially expressed genes. Some pathways were associated with cancer, such as pathways in cancer (associated with 11 genes), PI3K-Akt signaling pathway (associated with 15 genes, Figure 3).

## 4. Co-expression analysis and target prediction

The functions of lncRNAs were performed by inter-acting with their targets. In the present study, the potential cis- and trans-acting were predicted. The mRNAs 100 kb upstream and downstream of the lncRNAs were searched for cis analysis, and we found that 5 lncRNAs were related with cis target genes (Figure 4). The trans analysis of lncRNAs was performed by constructing co-expression networks (Fig. 5) of dysregulated mRNAs and lncRNAs based on the expression correlation coefficient (Pearson correlation  $P$ -value  $< 0.05$ ). The data showed that most lncRNAs acted in a trans-acting manner. The top 200 prediction relationships with the highest prediction reliability were selected to

count the frequency of each TF, and the TF with more functional annotation was counted to reflect the overall functional distribution of different lncRNAs (Figure 6). With the detected 964 lncRNAs that corresponded to 9358 mRNAs, more than one mRNAs were predicted to be regulated by one lncRNA and one mRNA corresponded to several lncRNAs.

## 5. qRT-PCR Validation

The up-regulated genes (*NR\_110876*, *NR\_132344*, *XR\_946886*, *XR\_002956345*, *NR\_135295* and *XR\_002957935*) that were associated with carcinogenesis or chronic inflammation in common bile duct were verified by qRT-PCR. The results showed that the expression trends of *NR\_110876*, *NR\_132344*, *XR\_946886* and *XR\_002956345* were consistent with the microarray results, and the difference was statistically significant *NR\_132344*, *XR\_946886* and *XR\_002956345* ( $P < 0.05$ ). However, the expression trends of *NR\_135295* and *XR\_002957935* were not consistent with the microarray results (Figure 6). To further assess the accuracy of *NR\_110876*, *NR\_132344*, *XR\_946886*, *XR\_002956345* as biomarker in PBM, analysis of the ROC curve was performed with the AUC. The results indicated that only the AUC of *XR\_946886* was significantly different (0.837,  $p < 0.001$ ) (Figure 8).

# Discussion

In PBM cases, the reflux of pancreatic juice into the bile duct could damage the bile duct wall and lead to bile duct dilatation, cholestasis, stone formation, epithelial hyperplasia and malignant transformation in the biliary tract [4,5,6,7]. The biological effects of reflux of pancreatic juice and its relationship with chronic inflammation and carcinogenesis have attracted great attention in recent years. Previous studies regarding the transcriptomes of congenital dilation of common bile duct (CDC) have discovered numerous novel transcriptomic alterations. Kaneko et al. identified several downregulated genes in the gallbladder of children with PBM which may contribute to the pathophysiology and found some upregulated noncoding RNAs that may be important for biliary carcinogenesis [23]. In our previous study [24], we found 876 differentially expressed genes in children with PBM, of which 530 genes were up-regulated and 346 genes were down-regulated. Wong et al. identified 21 damaging de novo variants (DNV) in 31 cases with CDD, 6 DNV-carrying genes associated with human developmental disorders and 4 DNV-carrying genes linked with cholangio- and hepatocellular carcinomas [25]. However, there is rare report about lncRNA in PBM patients.

In the present study, we identified 489 lncRNA and 5036 mRNA transcripts that were found differentially expressed in the bile duct tissues of patients with PBM. These differentially expressed lncRNAs may affect many pathways, including those involved in Protein digestion and absorption, ECM-receptor interaction, Focal adhesion, PI3K-Akt signaling, carcinogenesis and inflammation, which were consistent with previous studies [21-23]. These findings, together with our qPCR data, suggest that lncRNAs may contribute to PBM associated complication, such as carcinogenesis and chronic inflammation in common bile duct. Among these differentially expressed lncRNAs, the upregulation of NR\_110876, NR\_132344, XR\_946886 and XR\_002956345 was confirmed by qPCR; however, only XR\_946886 showed significant difference in the AUC of ROC. The findings suggested that XR\_946886 might be involved in the development of complication of PBM and could become new biomarkers for carcinogenesis or chronic inflammation in common bile duct in PBM.

The potential functions of lncRNAs were commonly predicted by their target genes. In the present study, the relationships between lncRNAs and their target genes collocated (within 100 kb) were analyzed. We found that 5 cis target genes were related with differentially expressed lncRNA and that most lncRNAs acted in a trans manner. For example, we found that E2F4 was closely associated with lncRNA and target genes, which is consistent with previous reports that E2F4 is associated with the development and progression of cancers [26,27].

There were some limitations in the present study. The sample size was relatively small and only six related lncRNAs identified in microarray experiments were validated via qRT-PCR. Future studies should be conducted with larger sample sizes and in-depth verification of more candidate genes. In addition, the potential mechanisms of XR\_946886 and its clinical values in PBM still require further studies.

In summary, the present study provided an overall analysis of lncRNAs in common bile duct in PBM. By further exploring potential functions and pathways in which the lncRNAs were involved, we anticipate that lncRNAs such as XR\_946886 will be biomarkers or even potential therapeutic targets for PBM treatment.

## Abbreviations

long non-coding RNAs=lncRNAs; pancreaticobiliary maljunction=PBM; messenger RNA=mRNA; qRT-PCR= quantitative reverse-transcription polymerase chain reaction; ROC= receiver operator characteristic; GO= Gene Ontology; KEGG= Kyoto Encyclopedia of Genes and Genomes;

## Declarations

### Ethics approval and consent to participate

Informed written consent was obtained from all participants' parents. The study protocol was approved by the Ethics Committee of the Children's Hospital of Soochow University, Suzhou, China (No. 20170506016).

### Consent for publication

Written informed consent for publication of the case was obtained for each participant's parents.

### Availability of data and materials

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

### Competing interests

The authors declare that they have no competing interests.

### Funding

This work was supported by National Natural Science Foundation (No.81971685). The funding bodies only provided the financial means to buy reagents for experiments, and the funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

### Authors' contributions

Study conception and design: WL G. Acquisition of data: LZ. Analysis and interpretation of data: SL S. Drafting of manuscript: LZ. Critical revision of manuscript: WL G. All authors have read and approved the manuscript.

### Acknowledgements

Not applicable.

### Authors' information

Department of Radiology, Children's Hospital of Soochow University, Suzhou, China, 215025

## References

1. Le Roy B, Gagnière J, Filaire L, Fontarensky M, Hordonneau C, Buc E (2016) Pancreaticobiliary maljunction and choledochal cysts: from embryogenesis to therapeutics aspects. *Surgical and radiologic anatomy* : SRA 38 (9):1053-1060. doi:10.1007/s00276-016-1669-y.
2. Kamisawa T, Kuruma S, Chiba K (2016) [Congenital biliary dilatation and pancreaticobiliary maljunction]. *Nihon Shokakibyo Gakkai Zasshi* 113 (12):1991-1997. doi:10.11405/nisshoshi.113.1991.
3. Urushihara N, Hamada Y, Kamisawa T, Fujii H, Koshinaga T, Morotomi Y, Saito T, Itoi T, Kaneko K, Fukuzawa H, Ando H (2017) Classification of pancreaticobiliary maljunction and clinical features in children. *J Hepatobiliary Pancreat Sci* 24 (8):449-455. doi:10.1002/jhbp.485.
4. Kamisawa T, Kuruma S, Tabata T, Chiba K, Iwasaki S, Koizumi S, Kurata M, Honda G, Itoi T (2015) Pancreaticobiliary maljunction and biliary cancer. *J Gastroenterol* 50 (3):273-279. doi:10.1007/s00535-014-1015-2.
5. Kamisawa T, Kaneko K, Itoi T, Ando H (2017) Pancreaticobiliary maljunction and congenital biliary dilatation. *Lancet Gastroenterol Hepatol* 2 (8):610-618. doi:10.1016/s2468-1253(17)30002-x.
6. Eda K, Mizuochi T, Takaki Y, Higashidate N, Hashizume N, Fukahori S, Asagiri K, Yamashita Y, Yagi M (2018) Adenomyomatosis of the Gallbladder With Pancreaticobiliary Maljunction in a Child. *J Pediatr Gastroenterol Nutr* 67 (4):e82. doi:10.1097/mpg.0000000000001367.
7. Ono S, Fumino S, Iwai N (2011) Diagnosis and treatment of pancreaticobiliary maljunction in children. *Surg Today* 41 (5):601-605. doi:10.1007/s00595-010-4492-9.
8. Kamisawa T, Ando H, Suyama M, Shimada M, Morine Y, Shimada H (2012) Japanese clinical practice guidelines for pancreaticobiliary maljunction. *J Gastroenterol* 47 (7):731-759. doi:10.1007/s00535-012-0611-2.
9. Guo WL, Zhan Y, Fang F, Huang SG, Deng YB, Zhao JG, Wang J (2018) Factors affecting the operating time for complete cyst excision and Roux-en-Y hepaticojejunostomy in paediatric cases of congenital choledochal malformation: a retrospective case study in Southeast China. *BMJ Open* 8 (5):e022162. doi:10.1136/bmjopen-2018-022162
10. Tsuchida A, Kasuya K, Endo M, Saito H, Inoue K, Nagae I, Aoki T, Koyanagi Y (2003) High risk of bile duct carcinogenesis after primary resection of a congenital biliary dilatation. *Oncol Rep* 10 (5):1183-1187.
11. Kobayashi S, Asano T, Yamasaki M, Kenmochi T, Nakagohri T, Ochiai T (1999) Risk of bile duct carcinogenesis after excision of extrahepatic bile ducts in pancreaticobiliary maljunction. *Surgery* 126 (5):939-944. doi:10.1016/s0039-6060(99)70036-x.
12. Sheng Q, Lv Z, Xu W, Xiao X, Liu J, Wu Y (2017) Reoperation After Cyst Excision with Hepaticojejunostomy for Choledochal Cysts: Our Experience in 18 Cases. *Med Sci Monit* 23:1371-1377. doi:10.12659/msm.900313.
13. Sugiura T, Endo T, Ito K, Goto K, Sato Y, Kondo S, Suzuki T, Hashimoto T (2013) Recurrent Cholangitis with Congenital Hepatic Fibrosis and Pancreaticobiliary Maljunction after Roux-en-Y Reconstruction. *European J Pediatr Surg Rep* 1 (1):43-45. doi:10.1055/s-0033-1341420.
14. Song W, Miao DL, Chen L (2018) Comprehensive analysis of long noncoding RNA-associated competing endogenous RNA network in cholangiocarcinoma. *Biochem Biophys Res Commun* 506 (4):1004-1012. doi:10.1016/j.bbrc.2018.10.186.
15. Jiang Y, Huang Y, Cai S, Song Y, Boyer JL, Zhang K, Gao L, Zhao J, Huang W, Liang G, Liangpunsakul S, Wang L (2018) H19 Is Expressed in Hybrid Hepatocyte Nuclear Factor 4a(+) Periportal Hepatocytes but Not Cytokeratin 19(+) Cholangiocytes in Cholestatic Livers. *Hepatol Commun* 2 (11):1356-1368. doi:10.1002/hep4.1252.
16. Ren GJ, Fan XC, Liu TL, Wang SS, Zhao GH (2018) Genome-wide analysis of differentially expressed profiles of mRNAs, lncRNAs and circRNAs during *Cryptosporidium baileyi* infection. *BMC Genomics* 19 (1):356. doi:10.1186/s12864-018-4754-2.

17. Long X, Song K, Hu H, Tian Q, Wang W, Dong Q, Yin X, Di W (2019) Long non-coding RNA GAS5 inhibits DDP-resistance and tumor progression of epithelial ovarian cancer via GAS5-E2F4-PARP1-MAPK axis. *J Exp Clin Cancer Res* 38 (1):345. doi:10.1186/s13046-019-1329-2.
18. Tang Y, Zhou T, Yu X, Xue Z, Shen N (2017) The role of long non-coding RNAs in rheumatic diseases. *Nat Rev Rheumatol* 13 (11):657-669. doi:10.1038/nrrheum.2017.162.
19. Sallam T, Sandhu J, Tontonoz P (2018) Long Noncoding RNA Discovery in Cardiovascular Disease: Decoding Form to Function. *Circ Res* 122 (1):155-166. doi:10.1161/circresaha.117.311802.
20. Zhou M, Zhao H, Xu W, Bao S, Cheng L, Sun J (2017) Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. *Mol Cancer* 16 (1):16. doi:10.1186/s12943-017-0580-4.
21. Song P, Jiang B, Liu Z, Ding J, Liu S, Guan W (2017) A three-lncRNA expression signature associated with the prognosis of gastric cancer patients. *Cancer Med* 6 (6):1154-1164. doi:10.1002/cam4.1047.
22. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458 (7235):223-227. doi:10.1038/nature07672.
23. Kaneko K, Ito Y, Ono Y, Tainaka T, Tsuchiya H, Shimoyama Y, Ando H (2011) Gene expression profiling reveals upregulated UCA1 and BMF in gallbladder epithelia of children with pancreaticobiliary maljunction. *J Pediatr Gastroenterol Nutr* 52 (6):744-750. doi:10.1097/MPG.0b013e318214bd30.
24. Guo WL, Geng J, Zhao JG, Fang F, Huang SG, Wang J (2019) Gene expression profiling reveals upregulated FUT1 and MYBPC1 in children with pancreaticobiliary maljunction. *Braz J Med Biol Res* 52 (8):e8522. doi:10.1590/1414-431x20198522.
25. Wong JK, Campbell D, Ngo ND, Yeung F, Cheng G, Tang CS, Chung PH, Tran NS, So MT, Cherny SS, Sham PC, Tam PK, Garcia-Barcelo MM (2016) Genetic study of congenital bile-duct dilatation identifies de novo and inherited variants in functionally related genes. *BMC Med Genomics* 9 (1):75. doi:10.1186/s12920-016-0236-z.
26. Bhawe K, Roy D (2018) Interplay between NRF1, E2F4 and MYC transcription factors regulating common target genes contributes to cancer development and progression. *Cell Oncol (Dordr)* 41 (5):465-484. doi:10.1007/s13402-018-0395-3.
27. Long X, Song K, Hu H, Tian Q, Wang W, Dong Q, Yin X, Di W (2019) Long non-coding RNA GAS5 inhibits DDP-resistance and tumor progression of epithelial ovarian cancer via GAS5-E2F4-PARP1-MAPK axis. *J Exp Clin Cancer Res* 38 (1):345. doi:10.1186/s13046-019-1329-2.

## Tables

**Table 1.** The clinical and pathological characteristics of PBM

Variables	Todani types I (n=7)	Todani types IV (n=8)
Abdominal pain	6	4
Jaundice	1	4
Mass	0	0
Fever	1	1
Vomiting	5	3
Gender (male)	0	3
Age (months)	9-71	4.5-50
Pathological findings		
Cyst wall hyperplasia	7	8
Gallbladder wall congestion	7	8

**Table 2:** The top 20 differentiated expressed lncRNA in patient samples when compared to normal samples

Gene Symbol	Gene ID	Accession	Description	Length	Regulation ([A] Vs [B])	FC (abs) ([A] Vs [B])	p ([A] Vs [B])	Type
LOC105371291	105371291	XR_933627.3	uncharacterized LOC105371291, transcript variant X3	1138	down	12.056273	0.01451713	ncRNA
LOC105374803	105374803	XR_940242.2	uncharacterized LOC105374803	1305	down	9.0651	0.00494713	ncRNA
LINC00261	140828	NR_001558.3	long intergenic non- protein coding RNA 261	4912	down	7.9187694	0.01998967	ncRNA
PSORS1C3	100130889	NR_152831.1	psoriasis susceptibility 1 candidate 3 (non- protein coding), transcript variant 5	624	down	6.993259	0.01583394	ncRNA
LOC112267956	112267956	XR_002956345.1	uncharacterized LOC112267956	2005	down	6.9718432	0.01816426	ncRNA
LOC158434	158434	NR_132344.1	uncharacterized LOC158434	2662	down	6.0362997	0.00271874	ncRNA
LOC105372653	105372653	NR_134564.1	uncharacterized LOC105372653	573	down	5.8000107	2.83E-04	ncRNA
LINC01320	104355288	NR_126404.1	long intergenic non- protein coding RNA 1320	1131	down	5.6668267	0.01277497	ncRNA
LOC105371825	105371825	XR_001752940.1	uncharacterized LOC105371825	3133	down	5.4514194	0.02630284	ncRNA
PP7080	25845	NR_024158.1	uncharacterized LOC25845	1743	down	5.4081116	0.01141873	ncRNA
LOC105369381	105369381	XR_001748496.1	uncharacterized LOC105369381, transcript variant X2	5254	down	5.392252	0.00472638	ncRNA
EMX2OS	196047	NR_144378.1	EMX2 opposite strand/antisense RNA, transcript variant 2	7201	down	5.325151	0.00589902	ncRNA
SNORA25B	109623459	NR_145801.1	small nucleolar RNA, H/ACA box 25B	127	down	5.319602	0.00461494	ncRNA
LOC101928875	101928875	XR_001748392.2	uncharacterized LOC101928875	728	down	5.1403937	0.006804	ncRNA
GBA3	57733	NR_102356.1	glucosylceramidase beta 3 (gene/pseudogene), transcript variant 2, non-coding	1243	down	5.103703	0.00271457	ncRNA
LOC101927269	101927269	NR_110075.1	uncharacterized LOC101927269	507	down	4.978601	0.03097457	ncRNA
LOC105377924	105377924	NR_134603.1	uncharacterized LOC105377924	415	down	4.93889	0.0150742	ncRNA
LOC105374510	105374510	XR_001741602.1	uncharacterized LOC105374510, transcript variant X3	4802	down	4.9087567	0.03875371	ncRNA
LOC102724834	102724834	XR_429171.4	uncharacterized LOC102724834	3131	down	4.588613	0.0016174	ncRNA
LOC107985242	107985242	XR_001738352.1	uncharacterized LOC107985242,	1899	down	4.5234885	0.00128315	ncRNA

			transcript variant X2						
LOC105378608	105378608	XR_946886.2	uncharacterized LOC105378608	3022	up	14.387957	0.0011432	ncRNA	
LOC105376384	105376384	XR_930616.3	uncharacterized LOC105376384, transcript variant X3	1188	up	6.8661094	9.25E-04	ncRNA	
HECW1-IT1	100127950	NR_135295.1	HECW1 intronic transcript 1	1839	up	6.84244	0.03381596	ncRNA	
LOC105371401	105371401	XR_002957935.1	uncharacterized LOC105371401	3094	up	6.693866	0.0253292	ncRNA	
CASC23	103581031	NR_125366.1	cancer susceptibility 23 (non-protein coding)	1583	up	6.3594065	0.00331883	ncRNA	
LOC105374641	105374641	XR_925740.2	uncharacterized LOC105374641	294	up	5.8032284	0.00589144	ncRNA	
LOC105379057	105379057	XR_001742782.1	uncharacterized LOC105379057, transcript variant X3	8261	up	5.3560147	4.11E-04	ncRNA	
LOC107986570	107986570	XR_001743976.1	uncharacterized LOC107986570, transcript variant X2	770	up	5.3314576	0.01600899	ncRNA	
LINC00578	100505566	NR_047568.1	long intergenic non-protein coding RNA 578	1222	up	5.320564	0.03225355	ncRNA	
MIR4435-2HG	541471	NR_136164.1	MIR4435-2 host gene, transcript variant 6	2046	up	5.3023853	0.00741963	ncRNA	
LOC107984270	107984270	XR_001747591.1	uncharacterized LOC107984270, transcript variant X1	3786	up	5.1539755	0.01410949	ncRNA	
MIR503HG	84848	NR_024607.1	MIR503 host gene	786	up	5.049058	0.00200934	ncRNA	
ADAM5	255926	NR_001448.1	ADAM metallopeptidase domain 5 (pseudogene)	1718	up	5.048116	9.81E-04	ncRNA	
MIR8061	102466251	NR_107028.1	microRNA 8061	75	up	4.8443885	5.84E-04	ncRNA	
LOC105375855	105375855	XR_928920.2	uncharacterized LOC105375855	7758	up	4.6685147	0.01672733	ncRNA	
SH3RF3-AS1	100287216	NR_029193.1	SH3RF3 antisense RNA 1	2793	up	4.144925	0.01781914	ncRNA	
LOC100506688	100506688	NR_104615.1	uncharacterized LOC100506688, transcript variant 2	3451	up	4.080634	0.00154016	ncRNA	
LOC107986185	107986185	XR_001741392.1	uncharacterized LOC107986185	513	up	3.9532552	4.36E-05	ncRNA	
LOC107986821	107986821	XR_001745275.1	uncharacterized LOC107986821, transcript variant X1	3257	up	3.8777435	0.0174942	ncRNA	
LOC105373592	105373592	XR_001739685.1	uncharacterized LOC105373592, transcript variant X2	877	up	3.8124175	0.00616176	ncRNA	

# Figures

Figure 1A Principal Component Analysis results of differentiated mRNA

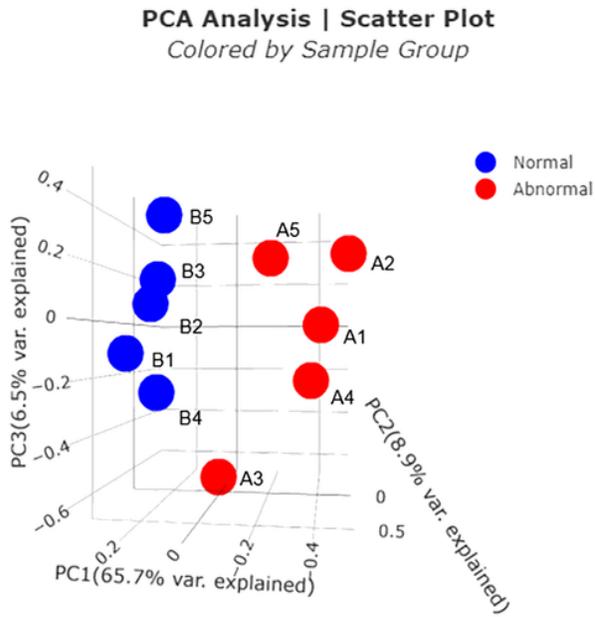


Figure 1B Clustergrammer analysis of differentiated lncRNA

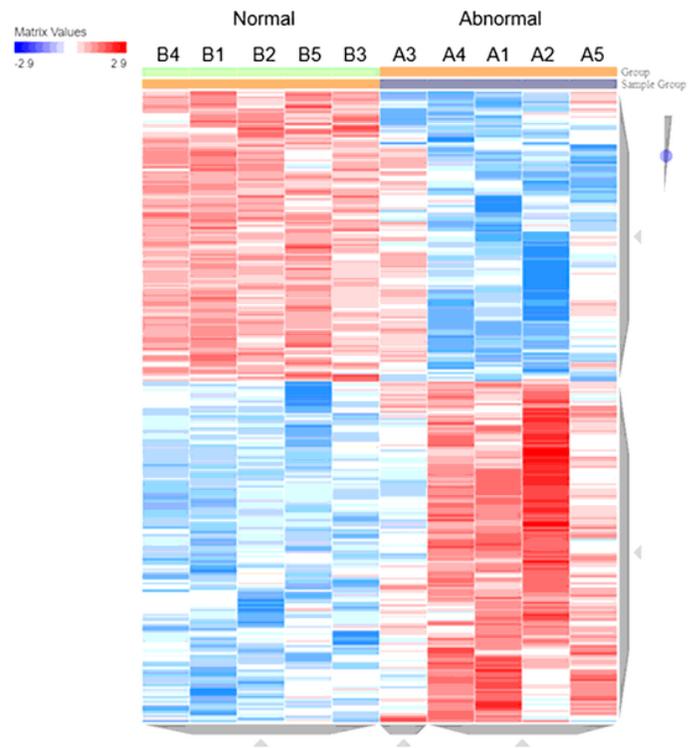


Figure 1

Principal component analysis (PCA,1A) results of differentiated lncRNA and heatmap(1B).

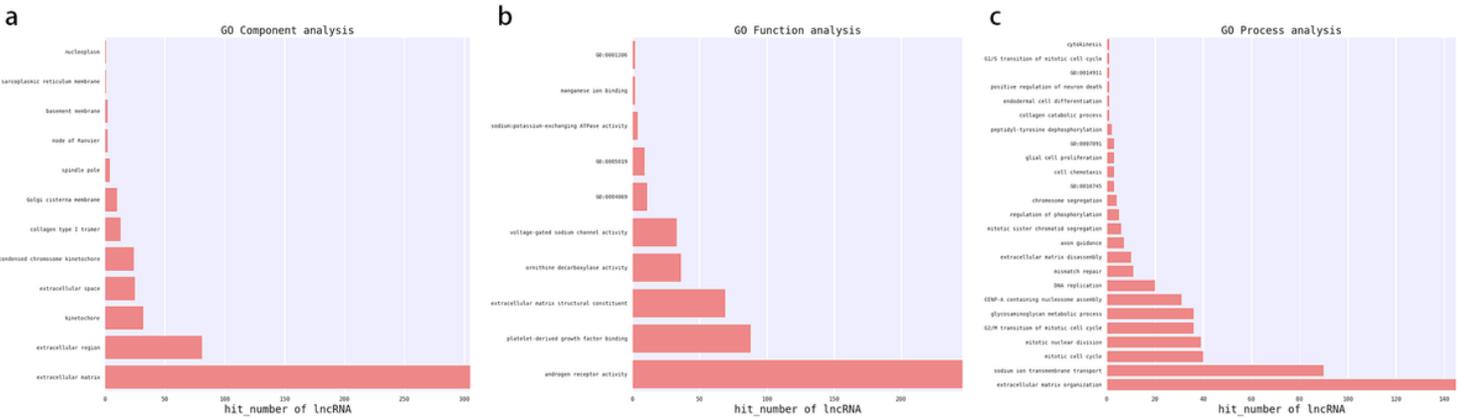
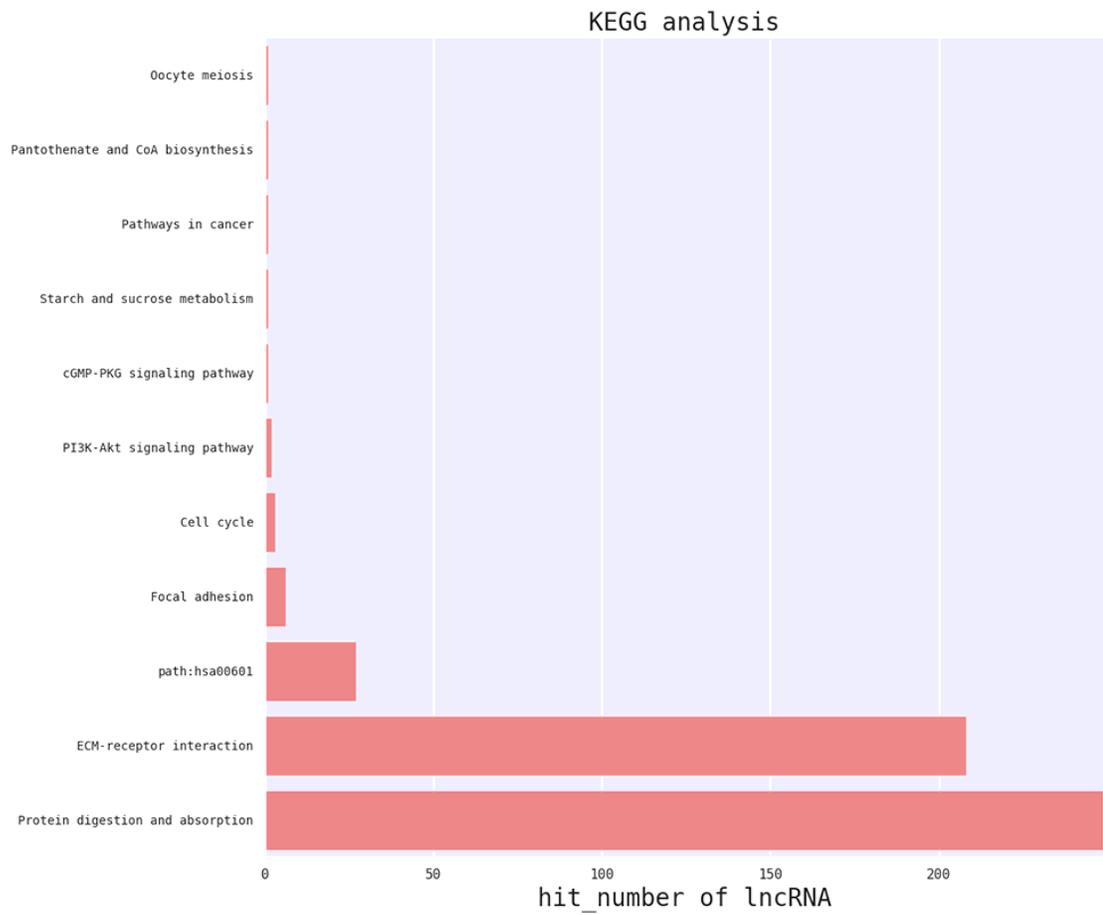


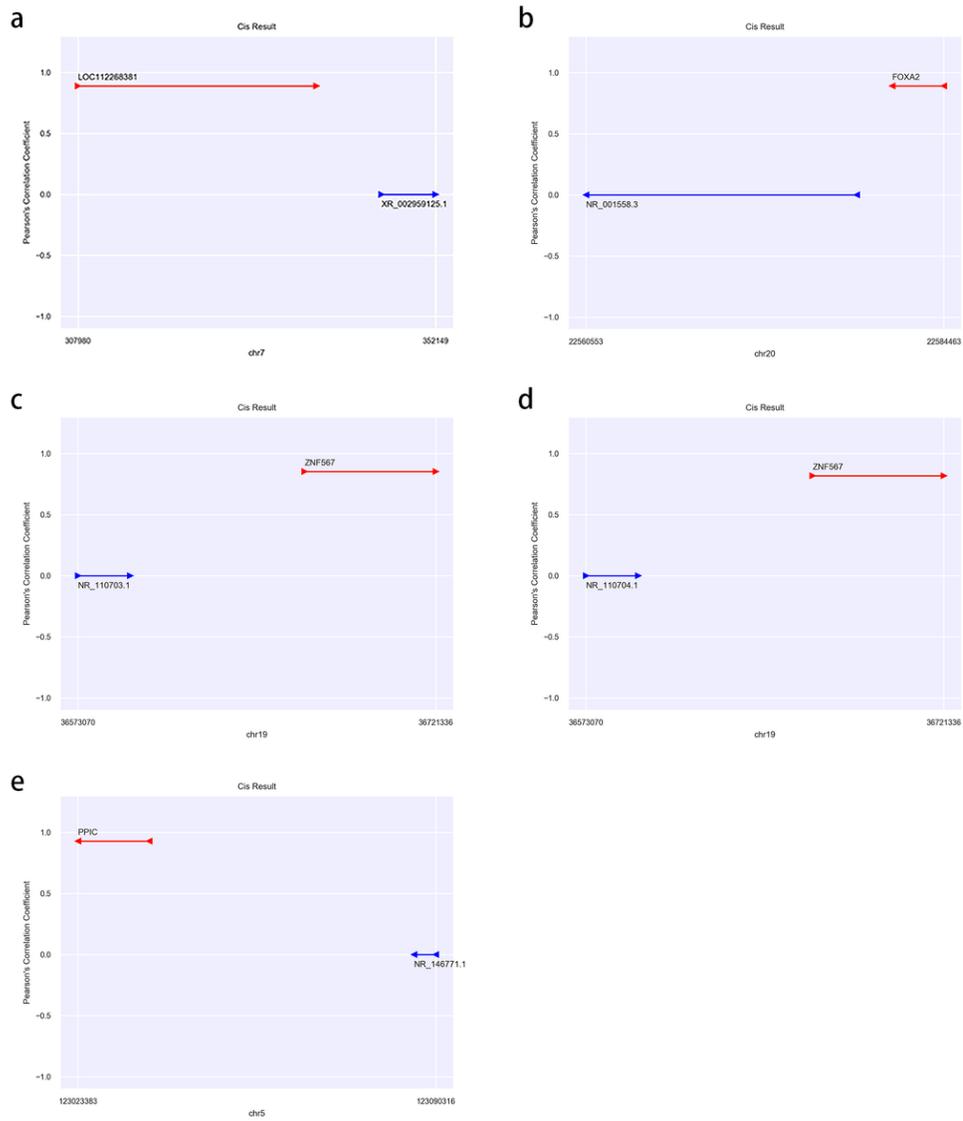
Figure 2

Functional analysis of differentially expressed genes with GO enrichment analysis.



**Figure 3**

Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis with the differentially expressed lncRNAs.



**Figure 4**

5 lncRNAs were related with cis target genes in PBM.

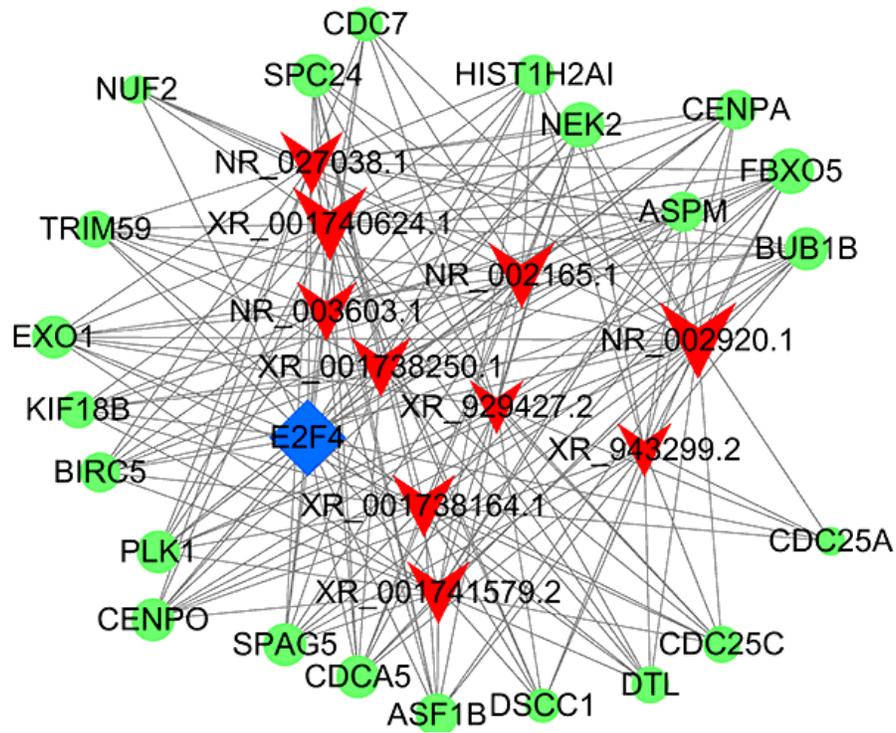
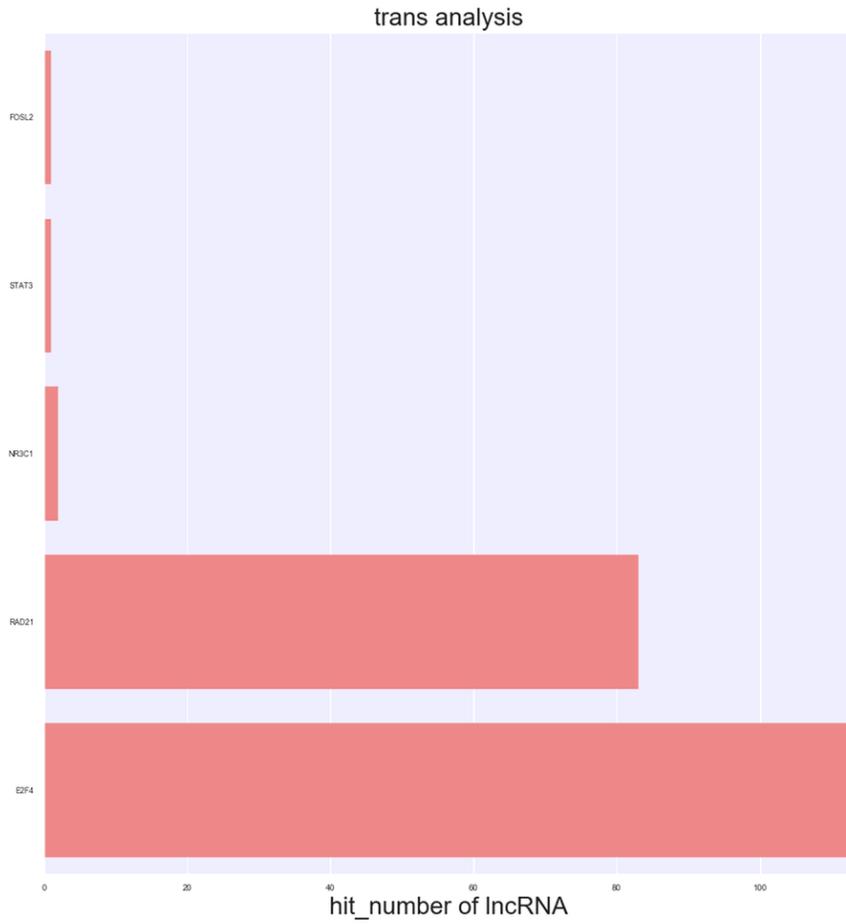


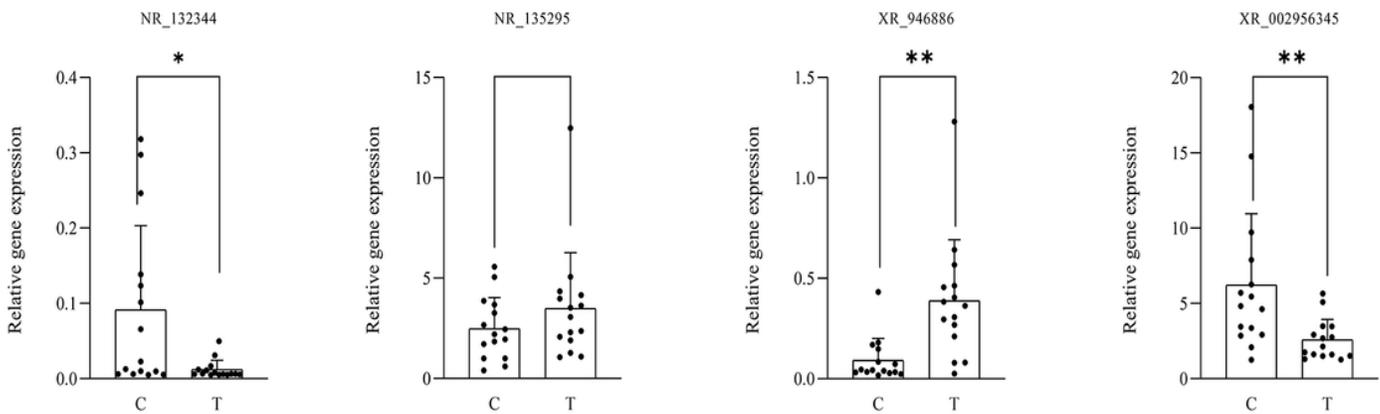
Figure 5

The trans-target prediction analysis of differentially expressed lncRNAs in PBM.



**Figure 6**

The functional distribution of differentially expressed lncRNAs with TF analysis.



**Figure 7**

Validation of the microarray test by qPCR.

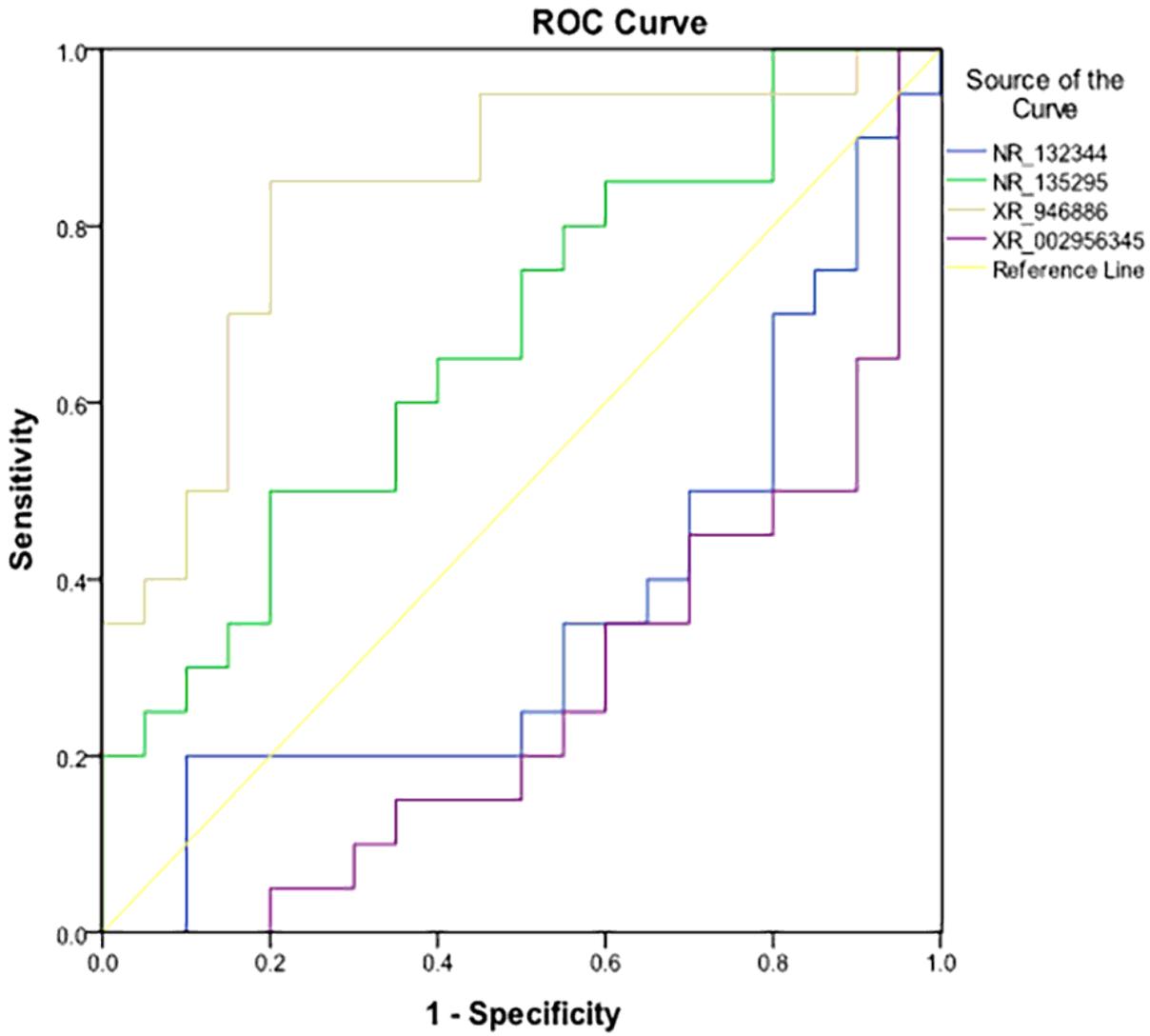


Figure 8

ROC analysis of genes NR\_110876, NR\_132344, XR\_946886, XR\_002956345 with AUC.