Chapter V

Analysing noncoding RNA linked to GBC pathogenesis and constructing a ceRNA regulatory network to identify potential noncoding regulatory signatures in GBC.

5.1 Introduction

The human genome project and next-generation sequencing technologies have revealed that only about 1.2% of the mammalian genome is responsible for encoding proteins, while the remaining percentage of the genome constitutes the ncRNAs [1-2]. Among these ncRNAs, lncRNAs show the most diversity and are expressed throughout development in a time- and tissue-specific manner [3-4]. LncRNAs have been found to play critical roles in cancer progression, development, and metastasis. Therefore, they are the subject of widespread research due to their significant involvement in physiological and pathological processes [5]. Several mechanisms have been implicated in lncRNA-mediated gene regulation, owing to their ability to interact with DNA, RNA, or protein. LncRNAs, for example, can act as signals to promote transcription, decoys to repress transcription, epigenetic regulators, or scaffolds to interact with various protein partners to form ribonucleoprotein complexes [6]. Rather than acting as coding RNAs, lncRNAs regulate gene expression through transcriptional, post-transcriptional, or epigenetic mechanisms. LncRNAs have a significant role in various cell processes, such as chromosomal imprinting, differentiation, proliferation, and DNA damage response [7-8].

LncRNAs exhibit diverse roles in the process of carcinogenesis, positioning them as potential targets for effective targeted therapy. Some function as oncogenes that drive angiogenesis, invasion, and metastasis, and others are tumor-suppressive lncRNAs that promote chemosensitivity and cellular apoptosis [9]. Many lncRNAs have been discovered over the previous few decades of RNA biology research, including *Xist* [10] and *H19* [11], which are considered milestones in lncRNA biology. Recently, a large number of lncRNAs including *MEG3* [12], *H19* [13], *MALAT1* [14], *HOTAIR* [15], *LET* [16], *CCAT1* [17], and *ANRIL* [12] have been identified to be associated with several cancer types including GBC. Studies have shown that lncRNAs play a role as either promoters or suppressors of tumors in the development and progression of GBC. The abnormal expression of these lncRNAs is linked to unfavourable survival outcomes among GBC patients [18]. While there's been notable advancement in comprehending the involvement of lncRNAs in GBC advancement, the comparative expression patterns and their roles in the pathogenesis of GBC related to gallstones and independent of gallstones have not been documented yet.

This chapter aims to identify and examine the differential expression of lncRNAs and understand the lncRNA-mediated regulation and reprogramming in GBC pathogenesis and

development. This study includes the identification of novel and annotated lncRNA from in-house generated transcriptomic datasets in two distinct GBC groups- GBC without gallstones (GBC) and GBC with gallstones (GBC+GS) through a network biology-based approach. The detailed workflow of the methodology implemented in this study is presented in **Figure 5.1.**

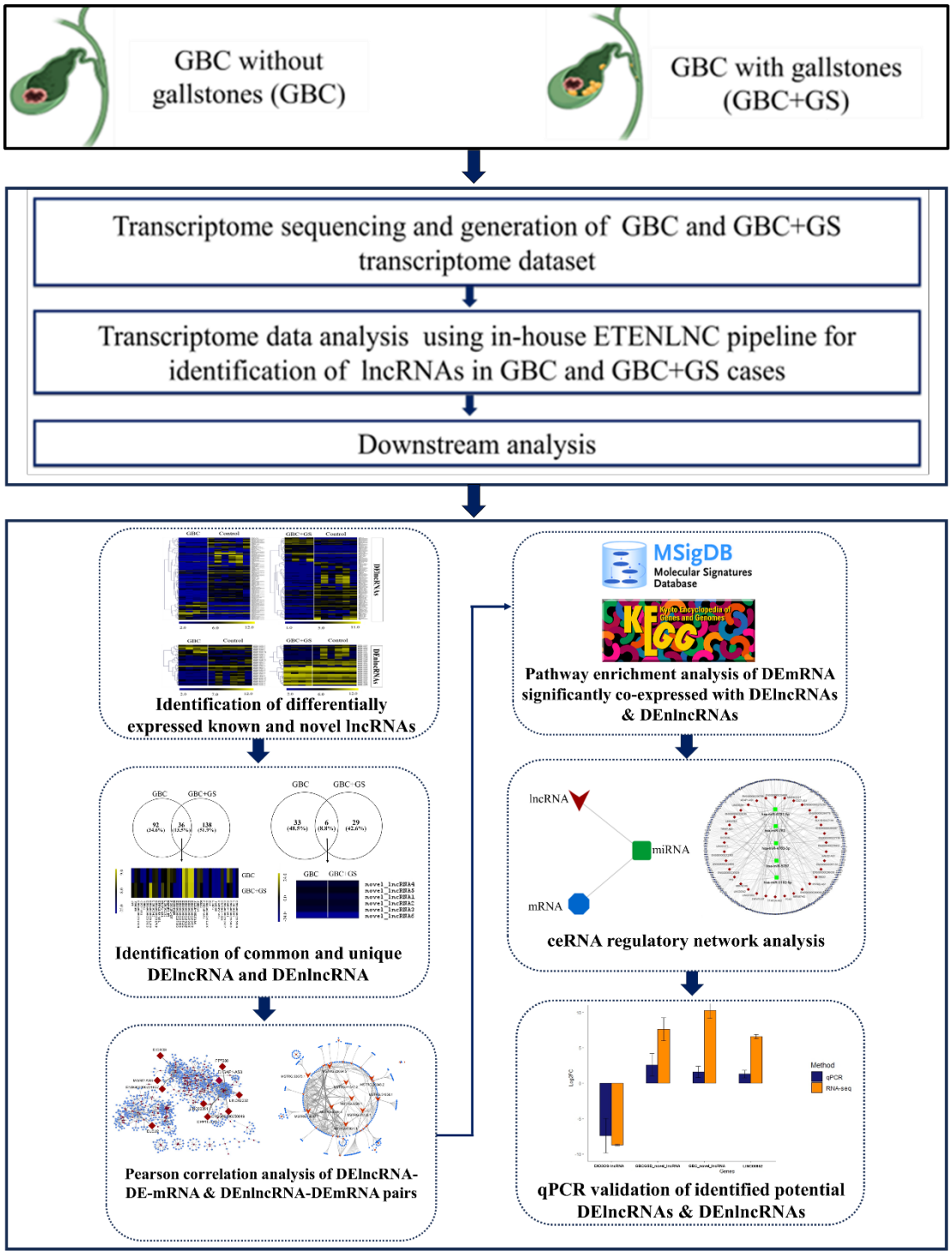
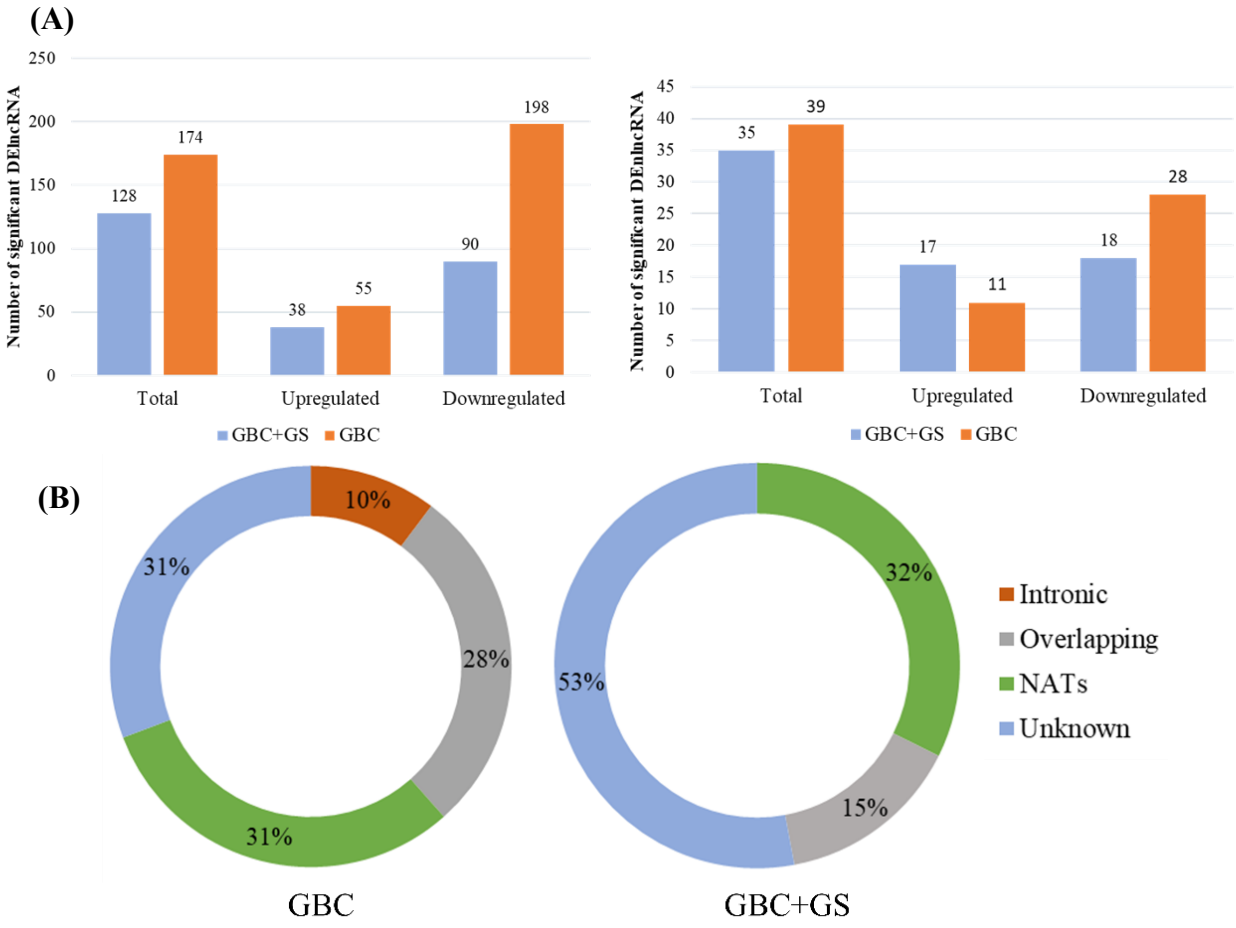


Figure 5.1: Overall schematic workflow for identification of crucial lncRNAs involved in GBC pathogenesis.

5.2 Results

5.2.1 Identification of differential lncRNA profile in GBC and GBC+GS cases

The analysis of differential gene expression revealed that in cases of GBC, 145 lncRNAs exhibited significant differential expression, while in GBC+GS cases, 175 lncRNAs displayed significant differential expression. Additionally, using the ETELNC pipeline, a total of 39 and 35 DEnovel lncRNAs (nlncRNAs) were respectively identified in GBC and GBC+GS cases [**Figure 5.2 A**]. Notably, both groups of GBC exhibited a higher count of downregulated DElncRNAs and DEnlncRNAs compared to upregulated ones. These DEnlncRNAs were classified based on their genomic position, with a significant portion falling under 'u' (unknown) class codes, denoting transcripts that are not present in the reference annotation utilized by the GFFCompare tool. These mainly encompass novel unannotated transcripts and intergenic transcripts. In both GBC categories, the second most prevalent class of lncRNAs identified as natural antisense transcripts (NATs), originating from overlaps in exons from the opposite strand [**Figure 5.2 B**].



The complete linkage hierarchical clustering analysis with DElncRNA and DEnlncRNA identified distinct clusters in GBC and GBC+GS cases compared to control samples. These distinctions in co-expressed groups of lncRNAs indicate a unique set of lncRNA signatures for each GBC case and the pathogenesis of GBC and GSD-related GBC results from differential patterns of lncRNA expression. The DElncRNA and DEnlncRNA expression profiles for the GBC and GBC+GS cases as compared to the control samples are represented in **Figure 5.2 C**.

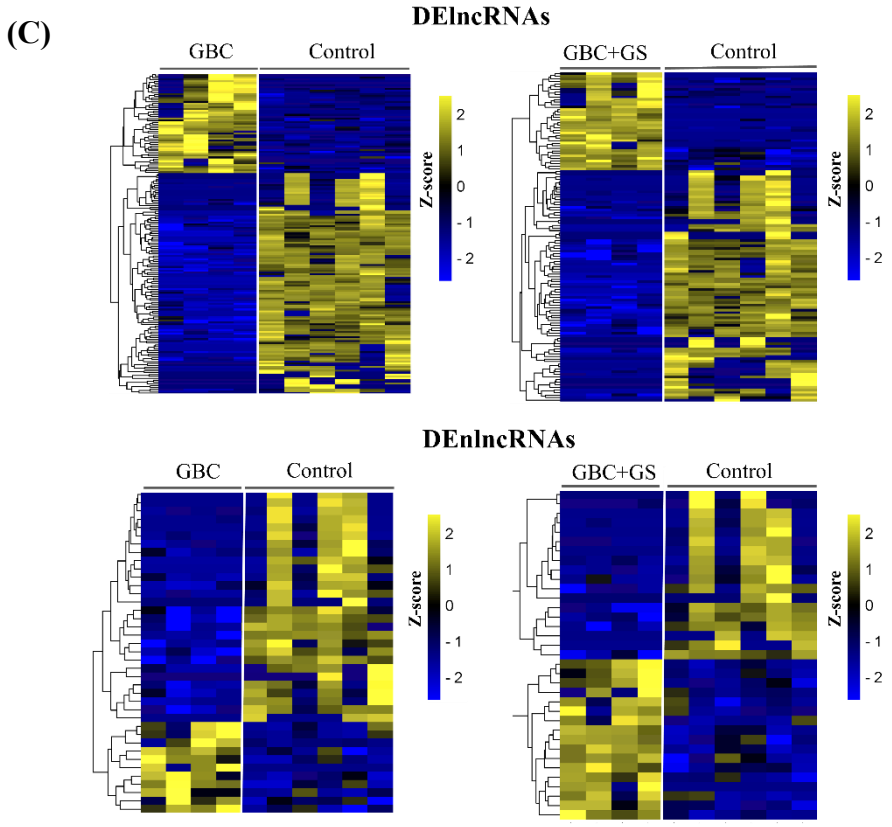


Figure 5.2: Identification of differential expression profile of lncRNA in GBC and GBC+GS group. (A) Bar graph showing the total, up and down-regulated DElncRNA and DEnlncRNA identified in GBC and GBC+GS groups. **(B)** Pie chart representing the percentage of novel lncRNAs classified under various categories based on their genomic positions. **(C)** Hierarchical clustering pattern of DElncRNA and DEnlncRNA expression profile in GBC and GBC+GS cases compared to control.

5.2.2 Identification of shared and specific lncRNA signatures in GBC and GBC+GS groups.

From **Fig 5.2 C**, it was observed that the differential expression of the DElncRNAs and DEnlncRNAs varies significantly between the two GBC groups (GBC and GBC+GS).

Therefore, to identify the shared and specific lncRNA signatures, the DElncRNAs and DEnlncRNAs identified in each group compared to controls were overlapped using venny 2.1. A total of 36 shared DElncRNAs have been identified between the two GBC groups. The majority of the shared DElncRNA signatures were found to be significantly downregulated. Gene-wise hierarchical clustering of the shared DElncRNAs showed similar expression of DElncRNAs in the two groups, except for 6 DElncRNAs which includes *HAND2-AS1*, *NR2F1-AS1*, *ENSG00000291178, MIR99AHG*, *CD27-AS1* and *JPX* [**Figure 5.3 A-B**]. Furthermore; the experimental validated overlapped DElncRNAs are found to be involved with key cancer hallmarks including proliferation, metastasis, invasion, and apoptosis [**Figure 5.3 C**]**.** Only 6 common DEnlncRNAs were identified between GBC+GS and GBC and no distinct DEnlncRNA expression level has been observed between the two GBC groups. Unlike shared DElncRNAs, the shared DEnlncRNAs are found to be highly upregulated. From this analysis, it was revealed that only a small percentage of shared DElncRNAs (13.5%) and DEnlncRNAs (8.8%) were identified among GBC and GBC+GS cases, which suggests that a distinct group of lncRNA signatures are associated with the pathogenesis and development of GBC and GBC+GS cases.

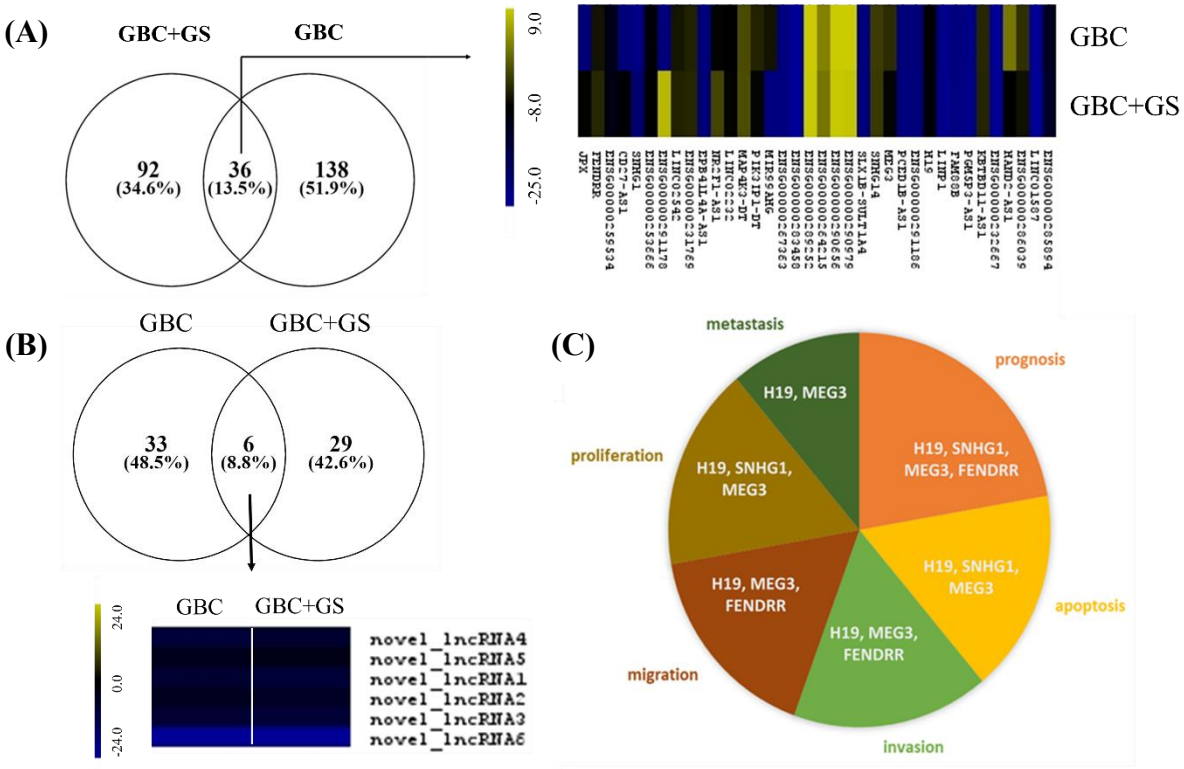


Figure 5.3: Identification of overlapping DElncRNAs and DEnlncRNA between GBC & GBC+GS group. Heatmap representing hierarchical clustering of 36 common DElncRNAs **(A)** and 6 common DEnlncRNAs **(B)** between two GBC groups. **(C)** Pie chart representing the

experimentally validated common DElncRNAs and their associated cancer functional state in GBC.

5.2.3 Screening of hub DElncRNAs and DEnlncRNAs through the construction of mRNA-lncRNA and mRNA-nlncRNA correlation networks

Several studies have demonstrated the association of lncRNAs in the regulation of mRNA at the transcriptional, post-transcriptional, and epigenetic levels. Therefore, to understand the correlation or association of the identified DElncRNAs and DEnlncRNAs with the DEmRNAs identified in case study 4 (referred to in 4.2.4), the expression correlation network has been constructed by applying a stringent cut-off correlation coefficient $|r| \geq 0.90$ and p-values of no less than 0.01 to identify highly co-expressing DElncRNA-DEmRNA and DEnlncRNA-DEmRNA pairs [**Table 5.1**]. A total of 548 and 252 DEmRNAs are significantly co-expressed with the expression of DElncRNAs and DEnlncRNA in the GBC+GS group respectively. Similarly, 777 and 423 DEmRNAs were coexpressed with DElncRNAs and DEnlncRNAs in the GBC group respectively.

Table 5.1: The statistics of correlation networks constructed with DEmRNA-DElncRNA and DEmRNA-DEnlncRNA pairs.

GBC groups	Coexpression pairs	Number of nodes	Number of edges	Path length	Network diameter
GBC	(DE) lncRNA-mRNA	931	3844	4.666	14
	(DE) nlncRNA-mRNA	462	2476	5.372	10
GBC+GS	(DE) lncRNA-mRNA	676	6248	4.951	11
	(DE) nlncRNA-mRNA	286	1291	6.014	12

Based on degree centrality, The top 10 significant DElncRNAs and DEnlncRNAs in each GBC group were identified through Pearson correlation networks. The top 10 hub DElncRNAs and DEnlncRNAs identified in each GBC group are clustered with their coexpressed DEmRNA respectively [**Figure 5.4**].

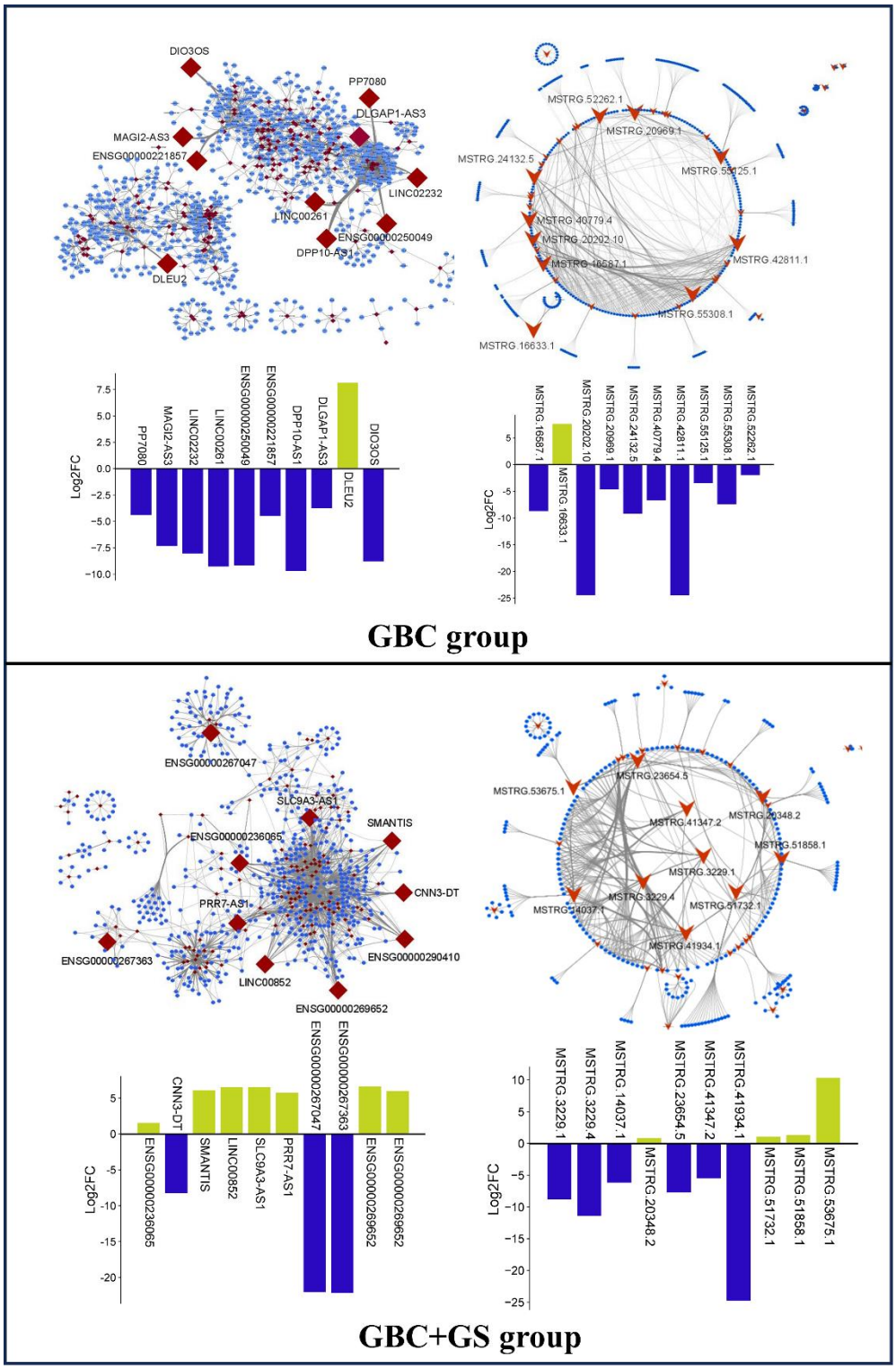


Figure 5.4: lncRNA-mRNA expression correlation network. Construction of DElncRNA-DEmRNA and DEnlncRNA-DEmRNA expression correlation networks in GBC and GBC+GS groups respectively. The blue and red nodes indicate DEmRNAs and DElncRNAs respectively. The bar plots represent the top ten hub DElncRNAs and DEnlncRNAs determined through correlation networks based on degree centrality in GBC and GBC+GS groups respectively. The blue bars represent downregulation and the yellow bars indicate upregulation.

5.2.4 Pathway enrichment analysis of the DEmRNAs coexpressed with hub DElncRNA and DEnlncRNA identified in GBC and GBC+GS cases

The pathway enrichment analysis was carried out by taking the DEmRNAs significantly co-expressed with the hub DElncRNAs and DEnlncRNAs to identify the signature pathways associated with GBC and GBC+GS pathogenesis. The top five significantly enriched (p-value<0.05) pathways were identified in the GBC and GBC+GS groups from two different pathway databases- the KEGG database and MsigDB hallmark. The KEGG pathways linked with the DEmRNAs in the GBC+GS group are involved in important cancer-related signaling pathways such as TNF-alpha signaling via NF-kB, mTORC1 signaling, and Wnt-beta catenin signaling. Whereas, metabolic processes such as bile secretion, retinol metabolism, and steroid hormone are significantly enriched hallmark pathways identified from the MsigDB hallmark database. Interestingly, it was found that the xenobiotic metabolism is found to be significantly enriched in both pathway databases. Xenobiotics are an important player in the cancer development. Altered expression of xenobiotics-related genes can activate pre-carcinogen and eventually lead to carcinogenesis [48-50]. In contrast, the enriched pathways identified in the GBC group are mainly associated with mitotic spindle, glycolysis, and transcriptional misregulation in cancer. DNA base excision repair, pancreatic cancer, and T-cell leukemia virus infection has been identified to be the common pathways in both MsigDB and KEGG databases.

The most significant pathway associated with the DEmRNAs significantly coexpressed with the hub DEnlncRNAs in the GBC group is bile secretion. Normal bile secretion is one of the pivotal functions of the gallbladder. Aberration of gene expression linked to bile metabolism and secretion pathway might initiate carcinogenic processes. whereas; the DEmRNAs coexpressed with DEnlncRNA in GBC+GS cases are associated with diverse pathological pathways including fatty acid metabolism, viral carcinogenesis, protein absorption and digestion, apical junction, and estrogen response early. Overall, the functional pathway enrichment revealed that the DEmRNAs coexpressed with the hub DElncRNAs and DEnlncRNAs identified in GBC and GBC+GS are associated with distinct biological pathways, which suggests that GBC induced by gallstones and GBC without gallstones progresses through independent pathophysiological pathways and mechanisms. The significantly enriched pathways associated with DEmRNAs coexpressed with hub DElncRNA and DEnlncRNA have been listed in **Table 5.2** and **Table 5.3**, respectively.

Table 5.2 List of top five enriched pathways linked with DEmRNAs coexpressed with hub DElncRNA identified in GBC and GBC+GS group. The pathways in bold font represent common enriched pathways identified between the KEGG and MsigDB databases.

Databases	Term	P-adjusted values	Combined Score
Enriched pathways associated with DElncRNA targets in GBC			
KEGG pathways	Bile secretion	2.80E-07	214.03
	Drug metabolism	1.32E-05	112.30
	Metabolism of xenobiotics by cytochrome P450	1.81E-05	134.17
	Retinol metabolism	1.31E-04	101.06
	Steroid hormone biosynthesis	9.78E-04	68.77
MsigDB hallmarks	TNF-alpha Signaling via NF-kB	1.65E-06	92.21
	mTORC1 Signaling	0.002	26.18
	UV Response Up	0.002	26.88
	Wnt-beta Catenin Signaling	0.005	45.03
	Xenobiotic Metabolism	0.008	16.96
Enriched pathways associated with DElncRNA targets in GBC+GS			
KEGG pathways	**Base excision repair**	3.99E-05	245.45
	Proteasome	1.50E-04	147.22
	Transcriptional misregulation in cancer	9.37E-04	39.91
	Pancreatic cancer	0.001	66.97
	Human T-cell leukemia virus 1 infection	0.001	31.45
MsigDB hallmarks	Mitotic Spindle	1.71E-04	56.6
	Glycolysis	0.001	37.09
	Pancreatic cancer	0.001	66.97
	Human T-cell leukemia virus 1 infection	0.001	31.45
	DNA Base Excision Repair	0.011	21.27

Table 5.3 List of top five enriched pathways linked with DEmRNAs coexpressed with hub DEnlncRNA identified in GBC and GBC+GS group. The pathways in bold font represent common enriched pathways identified between the KEGG and MsigDB databases.

	Term	P-adjusted values	Combined Score
Enriched pathways associated with DEnlncRNA targets in GBC			
KEGG pathways	**Bile secretion**	1.47E-05	109.65
	Folate biosynthesis	7.45E-05	198.96
	Metabolism of xenobiotics by cytochrome P450	5.56E-04	60.96
	ABC transporters	6.52E-04	82.31
	Steroid hormone biosynthesis	0.016	24.21
MsigDB hallmarks	**Bile Acid Metabolism**	0.008	24.73
	Estrogen Response Late	0.014	15.23
	Peroxisome	0.039	13.46
Enriched pathways associated with D-nlncRNA targets in GBC+GS			
KEGG pathways	Protein digestion and absorption	0.015	25.31
	Viral carcinogenesis	0.019	16.05
	Fatty acid degradation	0.019	38.31
	Herpes simplex virus 1 infection	0.04	7.98
	UV Response Down	0.006	29.64
MsigDB hallmarks	**Fatty Acid Metabolism**	0.008	25.23
	Estrogen Response Early	0.018	16.50
	Apical Junction	0.018676	16.50
	Xenobiotic Metabolism	0.018	16.50

5.2.5 Analysis of lncRNA-mediated interaction and construction of ceRNA regulatory networks.

Several studies have documented that lncRNAs can act as potential miRNA targets which regulate miRNA at the transcriptional and post-transcriptional level and influence the expression of target mRNA. Hence, it is important to identify the GBC-related lncRNAs that can effectively act as miRNA sponges or decoy and influence the mRNA expression. The highly interacted DElncRNA and DEnlncRNA and the top 5 interacting miRNAs based on the degree centrality in each GBC group were considered for the construction of the ceRNA networks. In the GBC group, the hub lncRNAs- *DIO3OS* and *LINC00261* were found to act as

ceRNA by sponging the miRNA targets [**Figure 5.5 A**]. In the GBC+GS group, the hub DElncRNA, *LINC00852* identified through coexpression analysis was found to be acting as regulatory lncRNA in the ceRNA network [**Figure 5.5 B]**. Apart from that, a novel ceRNA regulatory network was also constructed by taking the DEnlncRNA identified in each GBC group. The DEnlncRNA- *MSTRG.16633.1* and *MSTRG.53675.1* are overexpressed in the GBC and GBC+GS group as compared to the control and were found to interact with multiple miRNAs.

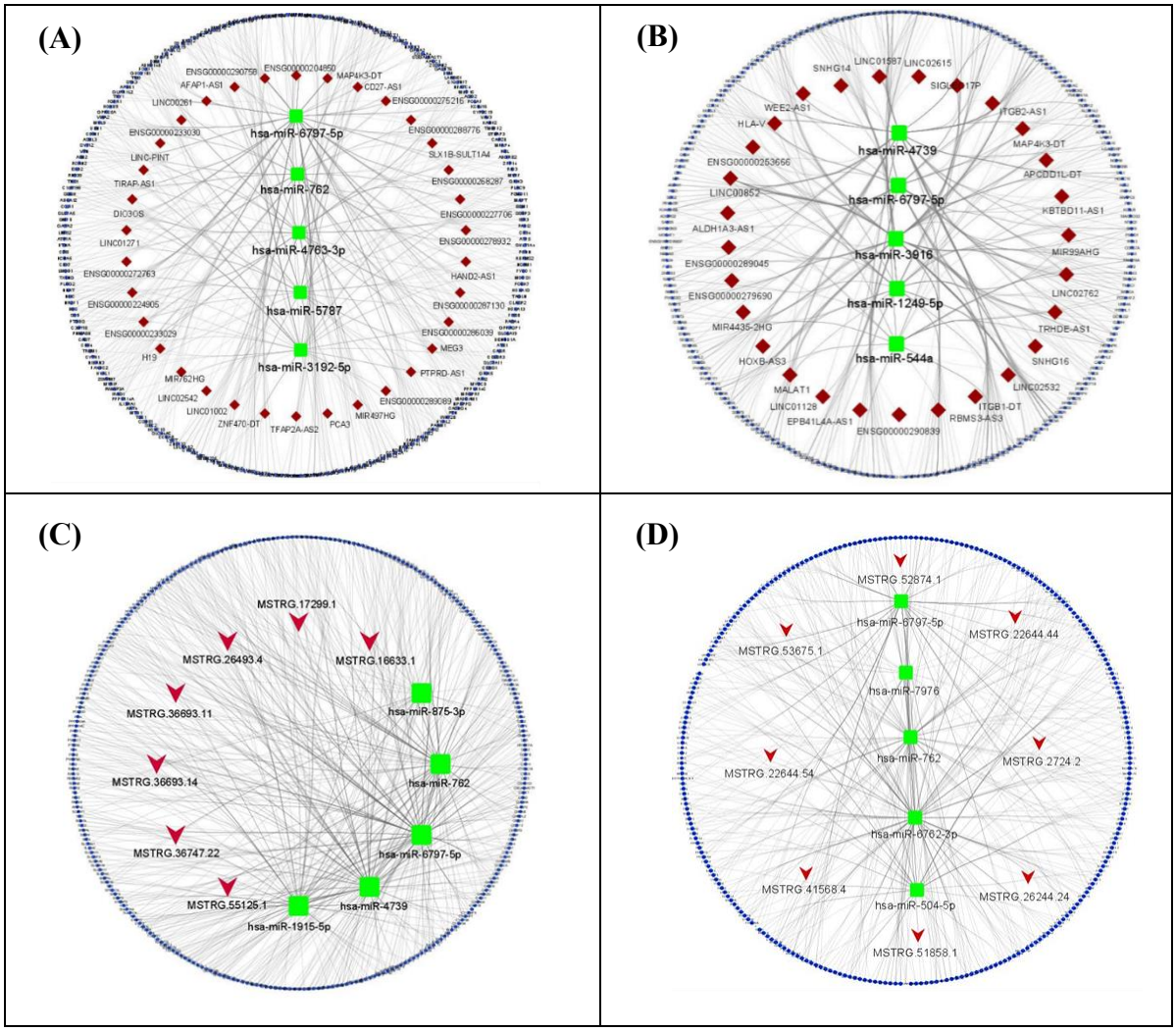


Figure 5.5: Construction of CeRNA networks. (A-B) CeRNA regulatory networks showing the highly connected miRNAs and their target DElncRNAs and DEmRNAs in GBC and GBC+GS groups respectively. (**C-D**) CeRNA regulatory networks showing the highly connected miRNAs and their target DEnlncRNAs and DEmRNAs in GBC and GBC+GS groups respectively DEmRNAs, DElncRNAs, and miRNAs are represented by blue octagon shapes, red diamonds, and green rectangles respectively.

The parent gene and genomic location of the hub DEnlncRNA identified using integrative coexpression and ceRNA network analysis has been provided in Table 5.4

Table 5.4: List of hub DEnlncRNAs and their parental gene information and genomic location.

Study groups	Hub DEnlncRNA	Parent gene	Class codes	Chromosome
GBC	MSTRG.16633	ENSG00000274798	Overlapping	15
	MSTRG.55125	RAI2	NATs	X
GBC+GS	MSTRG.53675	Unknown	Unknown	CHR_HSCHR6_MHC_SSTO_CTG1
	MSTRG.51858	Unknown	Unknown	CHR_HSCHR17_7_CTG4

The pathway enrichment analysis of the interacted DEmRNAs in the ceRNA networks constructed for the GBC and GBC+GS groups revealed that the interacting DEmRNAs in each GBC group are associated with distinct and diverse biological pathways. The interacting DEmRNAs in the GBC group are significantly enriched in metabolic pathways involving glycosaminoglycan degradation, proteoglycans, and aldosterone regulation pathways [**Figure 5.6 A-B**]. Conversely, in the GBC+GS group, the interacting DEmRNAs in the ceRNA networks are significantly associated with multiple cancer-related pathways including adherens junctions and oncogenic signaling transduction pathways such as Ras signaling, NOD-like receptor signaling, and Epstein-Barr virus infection pathways [**Figure 5.6 C-D**].

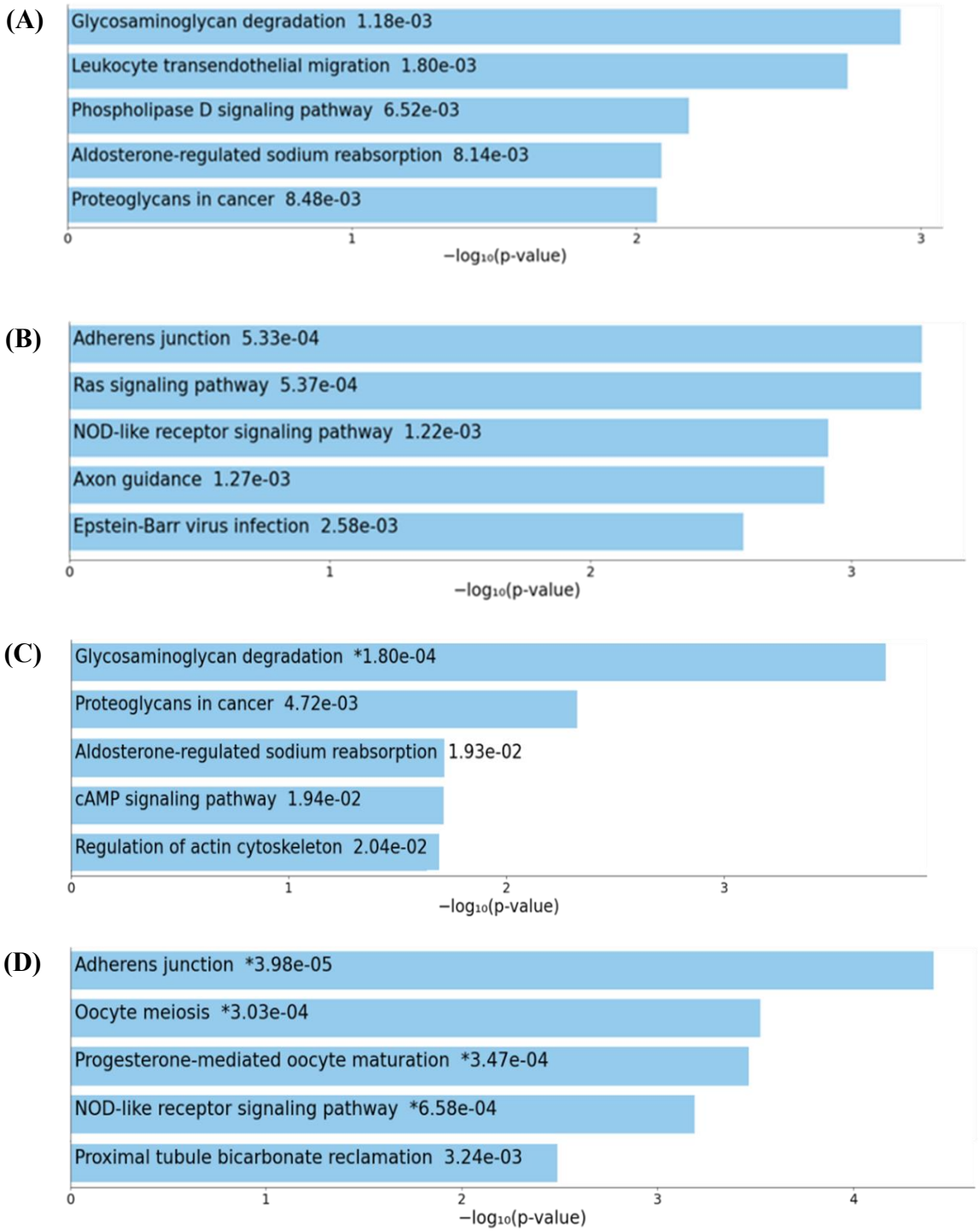


Figure 5.6: Pathway analysis of the interacting mRNAs in the ceRNA networks. (A-B) The bar plot represents the top five significantly enriched pathways linked to interacting DEmRNAs in respective GBC and GBC+GS ceRNA networks constructed with DElncRNAs. **(C-D)** The bar plot represents the top five significantly enriched pathways linked to interacting DEmRNAs in respective GBC and GBC+GS ceRNA networks constructed with DEnlncRNAs.

The highly connected miRNAs in the ceRNA networks constructed for the GBC and GBC+GS groups have been selected based on degree centrality. The miRNA- *hsa-miR-6797-5p* and *hsa-miR-762* are the most conserved miRNAs in all the ceRNA networks. The top 5 hub miRNAs identified in each GBC group are listed in **Table 5.5**.

Table 5.5: List of top five miRNAs identified in GBC and GBC+GS groups through construction of ceRNA networks. The asterisk symbol represents the miRNAs conserved in all the ceRNA networks.

Hub miRNAs identified from DElncRNA and DEnlncRNA mediated ceRNA networks in the GBC group			
miRNA	**Degree**	**miRNA**	**Degree**
hsa-miR-3192-5p	406	*hsa-miR-4739*	619
hsa-miR-5787	407	*hsa-miR-762*	532
hsa-miR-762	576	*hsa-miR-1915-5p*	154
hsa-miR-6797-5p	666	*hsa-miR-6797-5p*	374
Hub miRNAs identified from DElncRNA and DEnlncRNA mediated ceRNA networks in the GBC+GS group			
hsa-miR-4739	594	*hsa-miR-7976*	61
hsa-miR-1249-5p	350	*hsa-miR-6762-3p*	267
hsa-miR-3916	375	*hsa-miR-6797-5p*	216
hsa-miR-6797-5p	391	*hsa-miR-762*	519
hsa-miR-544a	235	*hsa-miR-504-5p*	61

In the previous chapter, under case study 4, the transcriptome sequencing and data analysis identified a few potential DEmRNAs in the GBC and GBC+GS groups. The ceRNA regulatory networks revealed that the hub lncRNAs- *DIO3OS* and *MSTRG.16633.1* identified in the GBC group and hub lncRNAs- LINC00825 and *MSTRG.53675.1* identified in the GBC+GS group were found to act as sponge or decoy and indirectly influence the activity of these hub DEmRNAs. Moreover; it was found that the hub DEmRNAs regulated by the DElncRNA and DEnlncRNA were highly downregulated except for *STAT3* which is highly upregulated in GBC+GS samples as compared to control. This indicated that lncRNA mediates the downregulation of DEmRNAs and might act as tumor suppressors in GBC pathogenesis. The hub ceRNA regulatory networks are presented in **Figure 5.7**.

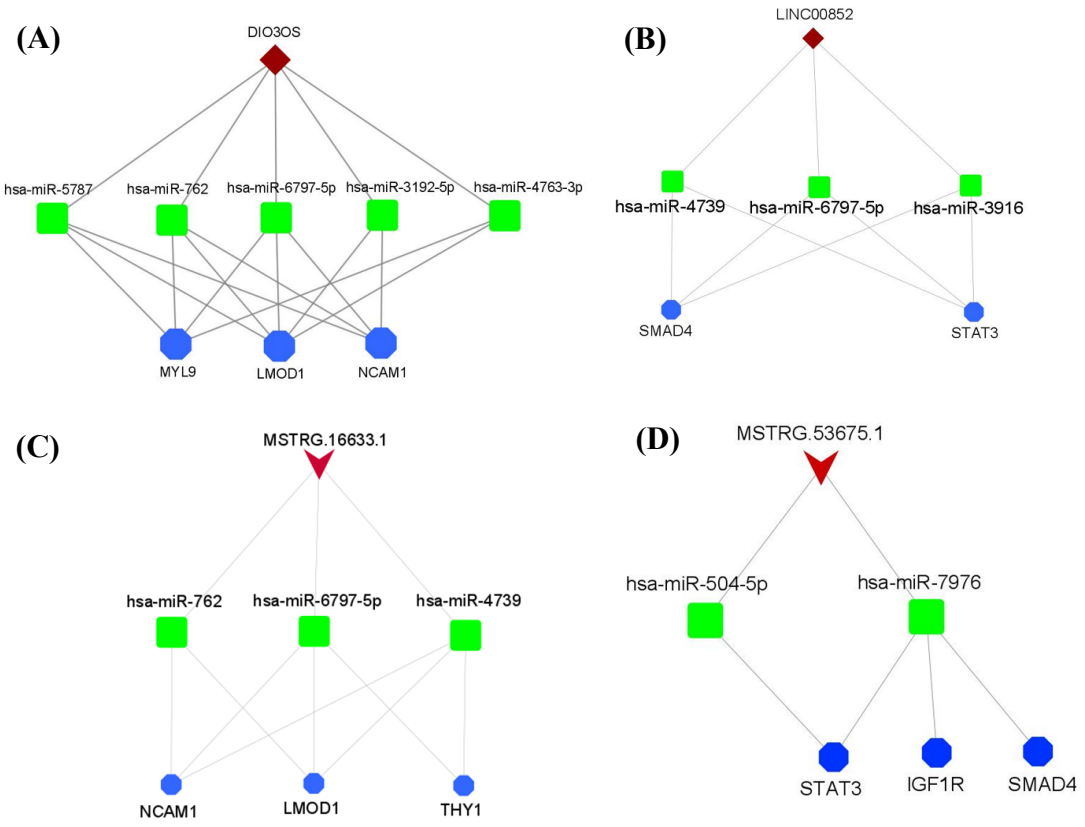


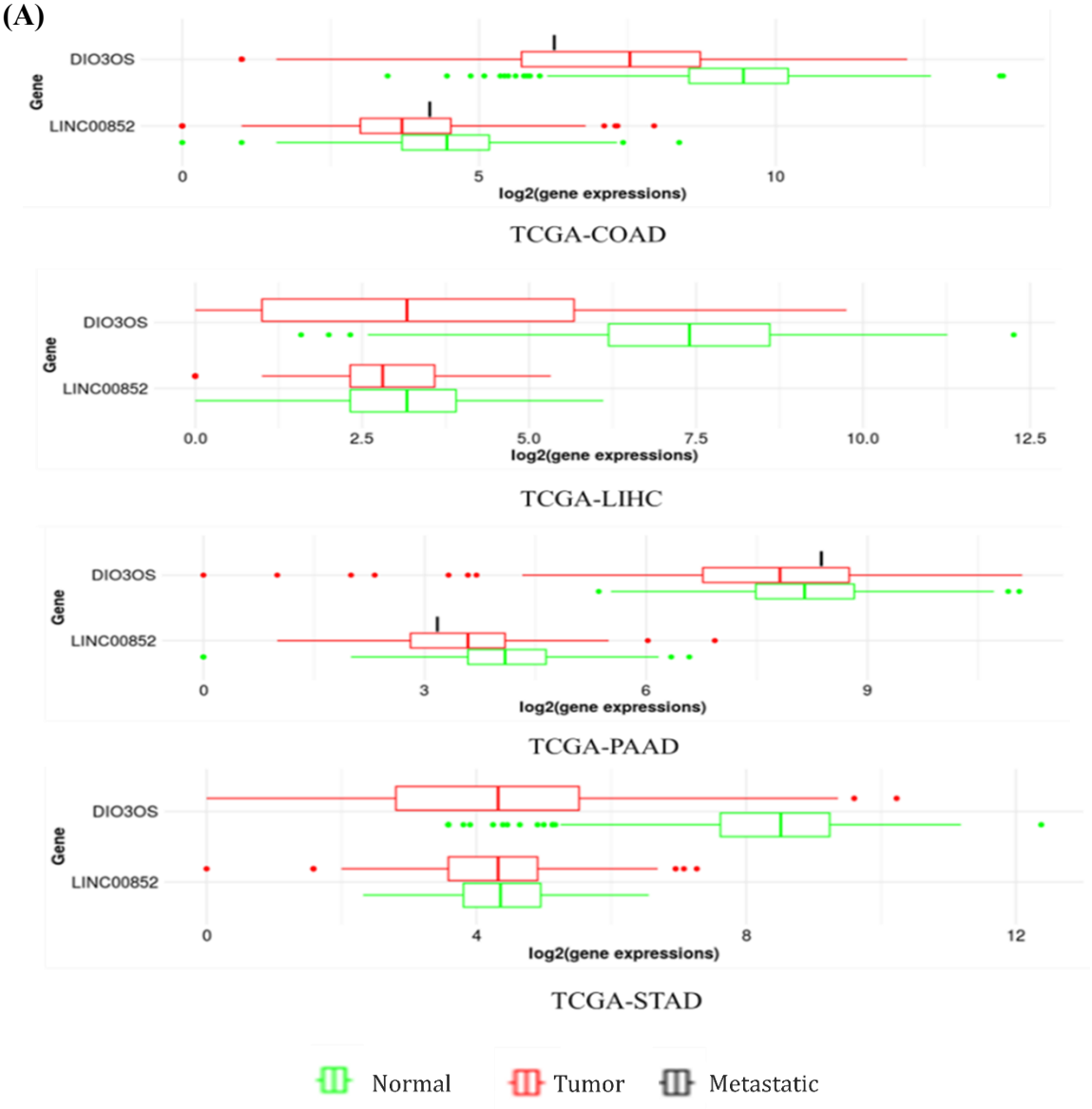
Figure 5.7: ceRNA network clusters involving the hub DElncRNAs and DEnlncRNAs targeting multiple hub miRNAs in GBC and GBC+GS groups. (A-B) ceRNA network clusters involving hub DElncRNA in GBC and GBC+GS groups respectively. **(C-D)** ceRNA network clusters involving hub DE novel lncRNA and in GBC and GBC+GS groups respectively.

5.2.6 Validation of the identified hub lncRNAs in independent datasets and through qRT-PCR

The expression of hub lncRNAs identified in each GBC group was further validated in independent TCGA datasets – Cholangiocarcinoma (CHOL), Liver hepatocellular carcinoma (LIHC), and Pancreatic ductal adenocarcinoma (PAAD) [**Figure 5.8 A**]. The expression of *DIO3OS* in CHOL and PAAD samples is upregulated whereas, in LIHC, the expression level of *DIO3OS* is significantly downregulated. Therefore, the expression level of DIO3OS identified in the GBC group significantly correlates with the *DIO3OS* expression in the LIHC dataset. The expression of lncRNA-*LINC00852* in CHOL and PAAD exhibited upregulated expression. However, no significant variation of the *LINC00852* expression is observed in the

LIHC dataset. Furthermore, the publicly available GBC transcriptome dataset (GSE139682) was analyzed, and differentially expressed lncRNAs were identified in GBC compared to normal. On comparing the DElncRNAs identified from the in-house transcriptome dataset involving the GBC and GBC+GS group with that of DElncRNAs identified from the public GBC dataset, it was revealed that only a small percentage of shared DElncRNAs were identified. The hub lncRNA- *DIO3OS* identified in the GBC group was also identified in the public GBC dataset [**Figure 5.8 B**] showcasing a consistent downtrend in its expression across both datasets. However; *LINC00852* identified in the GBC+GS group was not identified in the public dataset. Since *LINC00852* is associated with gallstone-related GBC cases, its absence in the public dataset could attributed to the transcriptome data analysis concentrating on GBC samples not associated with gallstones.

(A)



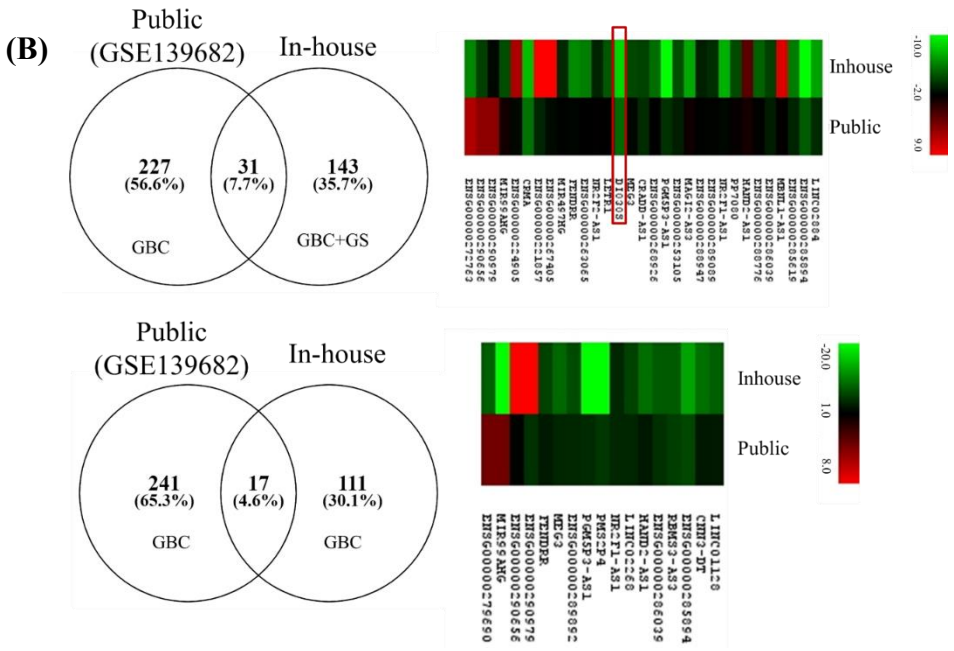


Figure 5.8: Cross-validation of the hub lncRNAs with independent public datasets. **(A)** Boxplot representing the log2foldchange expression of the hub lncRNAs in TCGA-CHOL, TCGA-LIHC, and TCGA-PAAD datasets. **(B)** Venn diagram showing the shared DElncRNAs identified between public and in-house datasets. The corresponding heatmaps showing the expression patterns of the common DElncRNAs identified.

Furthermore, the expression of the hub DElncRNAs and novel DElncRNAs were further identified in GBC and GBC+GS groups and were validated through qRT-PCR. qRT-PCR analysis showed the expression of *DIO3OS*, *LINC00852*, *MSTRG.16633.1*, and *MSTRG.53675.1* correlated with the expression level as identified from RNAseq data analysis. All the hub lncRNAs exhibited upregulation except for *DIO3OS*, which showed decreased expression in GBC in comparison to control samples. The Δ^{ct} and $2^{-\Delta\Delta ct}$ methods were used to analyze the qRT-PCR data, where Δ^{ct} and $2^{-\Delta\Delta ct}$ represent the sample's expression and relative expression of the target genes respectively [**Figure 5.9**].

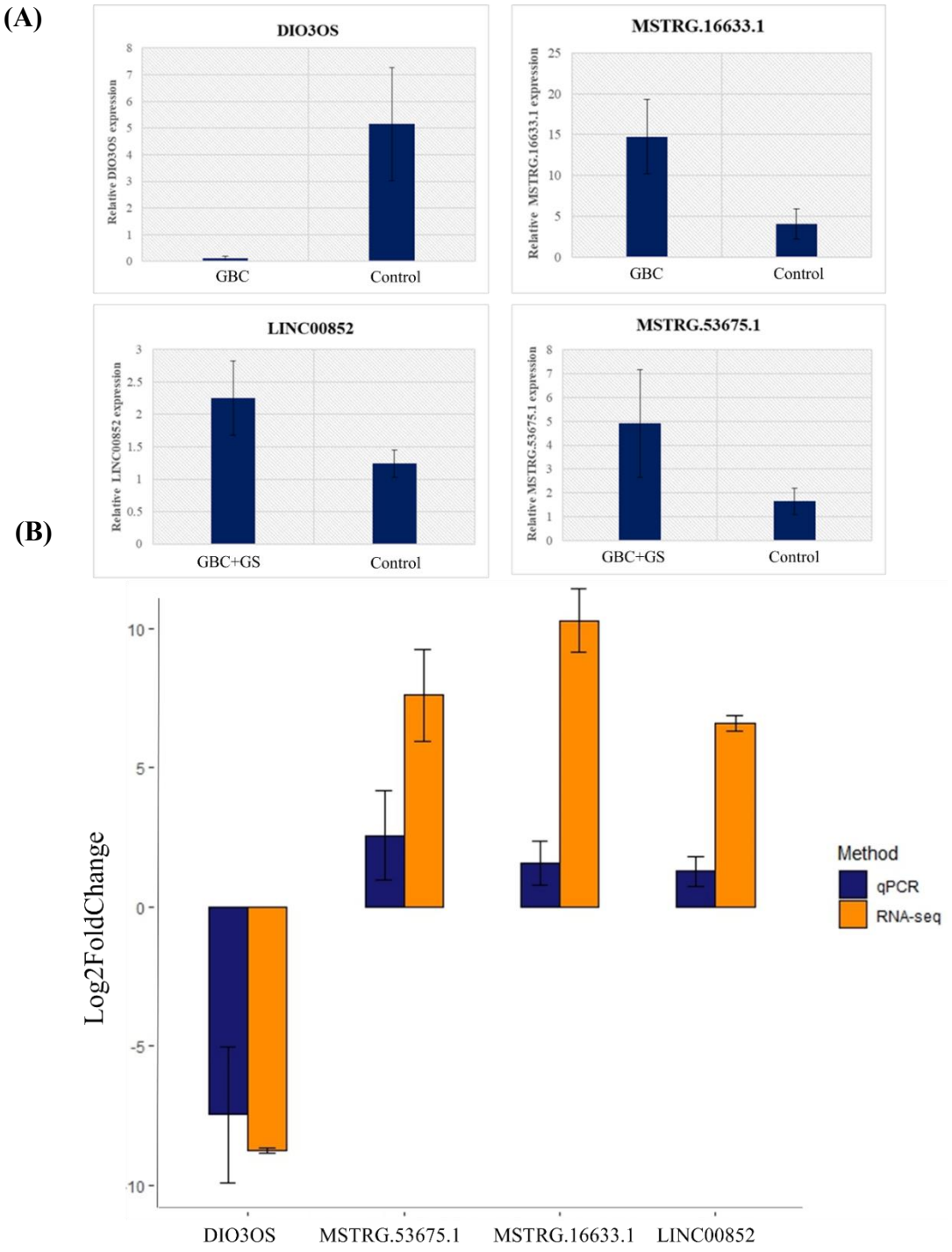


Figure 5.9: qRT-PCR validation of the hub lncRNAs identified in the GBC and GBC+GS group. (A) Bar plot representing the relative expression of lncRNAs identified using qRT-PCR data analysis in GBC compared to control. **(B)** Bar plot showing the gene expression level (log2FoldChange) of lncRNAs identified through RNAseq and qRT-PCR.

The secondary structure of the hub novel lncRNAs *MSTRG.16633.1* and *MSTRG.53675.1* identified in the GBC and GBC+GS group was predicted using the RNAfold tool by calculating the base pairing probabilities for each position in the RNA sequence and the MFE structure, which represents the most stable configuration for a given RNA sequence. The snapshot of two crucial novel lncRNA identified in GBC is presented in **Figure 5.10**.

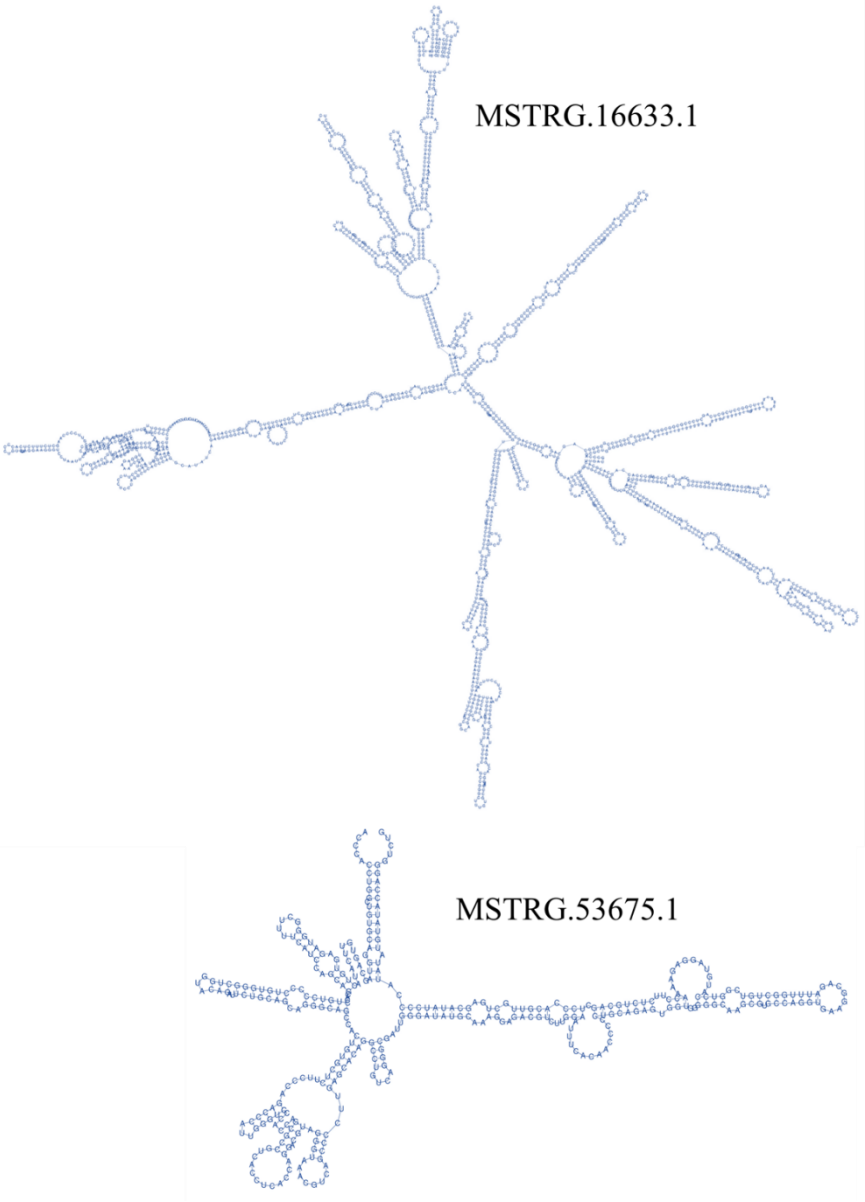


Figure 5.10: Predicted secondary structures of *MSTRG.16633.1* and *MSTRG.53675.1* identified in the GBC and GBC+GS groups respectively.

5.3 Discussions

Recent findings indicate that lncRNAs constitute a significant component of cancer biology [18]. dysregulated lncRNA expression in cancer delineates the range of disease progression and could independently predict patient outcomes [18–20]. LncRNAs play crucial roles in cancer by regulating gene expression across multiple levels: transcriptional, post-transcriptional, epigenetic, and translational [21–23]. Emerging evidence suggests that their dysregulation is extensively associated with tumorigenesis, including cell proliferation, migration, invasion, and apoptosis [24–27]. LncRNAs act as signal molecules to either promote or repress transcription, function as epigenetic regulators, and serve as scaffolds, interacting with diverse protein partners to form ribonucleoprotein complexes [23,28]. Numerous studies have suggested that lncRNAs play an important role either as oncogene (*HOTAIR*, *MALAT1*, *ANRIL*, *CCAT1*, Linc-*ITGB1*), or as tumor suppressor (*MEG3*, *LET*, *LincRNA*), or both oncogenic and tumor suppressor (*H19*) role in GBC etiopathogenesis [29]. However, the expression and mechanistic roles of most of the lncRNAs in GBC are still unclear.

Previous studies have focused on identifying potential DElncRNAs only in GBC compared to normal samples [51-53]. However, GBC patients can be further classified into two groups based on the presence (GBC+GS) and absence (GBC) of gallstones. In this study, efforts have been made to identify potential known and novel lncRNAs in the GBC and GBC+GS groups using integrative regulatory network analysis. The lncRNAs identifiedthrough the in-house lncRNA identification pipeline show distinct differential lncRNA expression in the GBC and GBC+GS groups. Expression correlation analysis of the DElncRNA – DEmRNA and DEnlncRNA - DEmRNA pairs identified the top 10 potential DElncRNA and DEnlncRNA based on the highest degree of interaction with DEmRNA in GBC and GBC+GS group. The DEmRNAs significantly coexpressing with the DE lncRNA and DEnlncRNA expression in the GBC group were identified to be involved in multiple metabolic pathways and signaling pathways including TNF-alpha Signaling via NF-kB, mTORC1, and Wnt-beta catenin signaling. The xenobiotic metabolism by cytochrome p450 and bile secretion were identified as the most enriched pathways. A recent study has revealed that variant genotypes of CYP1A1-MspI and CYP1A1-Ile462Val polymorphisms were associated with an increased risk of GBC in the patient group without stones [30]. Moreover, no relationship between *CYP1A1* polymorphisms and gallstone disease was observed in the Chinese population [30]. The gallstone-independent GBC cases could be due to *CYP1A1* genetic variants altering xenobiotic

metabolism. *CYP17*, *CYP1A1*, and *CYP1B1* are crucial enzymes involved in the biosynthesis of estrogen and testosterone hormones and xenobiotic metabolism. Furthermore, the increased hepatic excretion of cholesterol and xenobiotics can cause bile toxicity, which may lead to GBC development [31]. The DEmRNAs significantly coexpressing with the DElncRNA and DEnlncRNA expression in the GBC+GS group are found to be enriched in multiple cancer-related pathways, the most prominent of which is the base-excision repair pathway fatty acid metabolism pathways.

Further, by integrating the expression correlation and ceRNA regulatory networks, the potential hub lncRNAs and miRNAs identified in the GBC and GBC+GS groups, viz., *DIO3OS* and *LINC00852* emerged as the highly interacting lncRNAs, while *MSTRG.16633.1* and *MSTRG.53675.1* were recognized as the most prominently interacting novel lncRNAs in the GBC and GBC+GS groups, respectively. DIO3OS is an antisense lncRNA transcribed from the DIO3 gene imprinted locus, which encodes type 3 deiodinase (*D3*) [32-33]. The primary function of the human D3 gene is to maintain the thyroid hormone (TH) in fetuses and adults [33]. The expression of *DIO3OS* is detected in many tissues but shows decreased expression levels in several cancers. Recent studies have demonstrated a potential link between *DIO3OS* inactivation and tumor pathogenesis [34]. The expression of *DIO3OS* was found to be significantly decreased in HCC tissues and cells and is linked to poor prognosis in HCC patients [35]. However, the expression level of *DIO3OS* lncRNA is upregulated in several cancers including thyroid and pancreatic cancer [34,36]. This suggests that *DIO3OS* plays the role of tumor suppressors and oncogenes in cancer pathogenesis. This is the first study that reported the association of *DIO3OS* in GBC particularly in GBC cases without gallstones.

LINC00852 is a novel lncRNA that was initially discovered in lung adenocarcinoma. It acts as an oncogene and promotes lung adenocarcinoma cell proliferation, migration, and invasion [37]. A recent study showed that overexpressed *LINC00852* facilitates the infiltration of osteosarcoma cells by acting as a ceRNA [38]. *LINC00852* has been reported to be dysregulated in several cancers. For instance, *LINC00852* overexpression is associated with poor survival in gastric cancer patients [39]. *LINC00852* was found to be up-regulated in lung adenocarcinoma spinal metastases and lung adenocarcinoma cells and promotes lung adenocarcinoma cell proliferation while inhibiting apoptosis in vitro [37]. Furthermore, a recent study demonstrated that *LINC00852* overexpression contributes to the proliferation and migration of ovarian [40] cancer and prostate cancer cells [41]. These findings suggest that

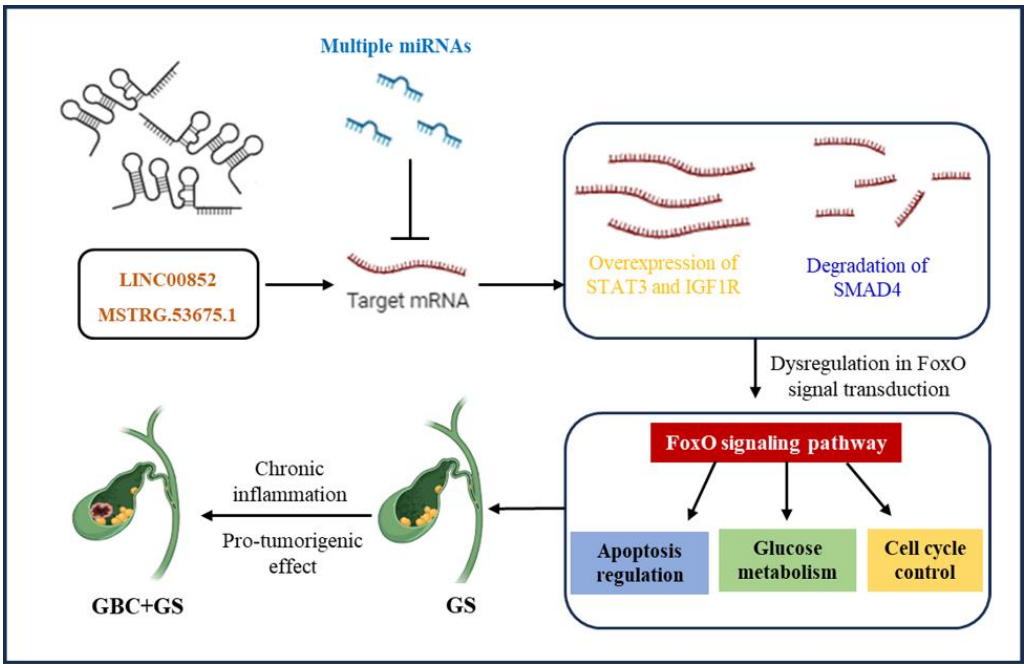
LINC00852 acts as an oncogene and regulates the proliferation and progression of several cancers. LINC00852 is highly overexpressed in GBC+GS cases as compared to control, which indicates its potential oncogenic function in GBC cases with gallstones.

From the ceRNA regulatory networks, it was observed that both the hub DElncRNAs (*DIO3OS* and *LINC00852*) and hub DEnlncRNAs (*MSTRG.16633.1* and *MSTRG.53675.1*) were found to interact with multiple miRNAs. Based on degree centrality, the top five miRNAs were identified in each ceRNA network. Among these hub miRNAs, *miR-6797-5p* and *miR-762* were present across all the ceRNA networks. This suggests a potential involvement of these miRNAs in GBC pathophysiology. miRNA *miR-6797-5p*, a microRNA, was not extensively studied compared to more well-known microRNAs in the context of cancer. The exploration of its role in cancer was relatively new and less documented. However, there were a few studies indicating its potential involvement in diseases including cancer. For instance, it was identified as a target of circRNA in the early stages of Alzheimer's disease [42]. Moreover, in hepatocellular and pancreatic ductal carcinoma, *miR-6797* was observed to be regulated by multiple circRNAs [43-44]. *miR-762* holds significance in several cancer types including nasopharyngeal squamous cell carcinoma (NSCLC), ovarian cancer, and breast cancer. It exhibits overexpression across several cancer types, detectable in both tissue and serum, suggesting the potential for *miR-762* to serve as a diagnostic and prognostic biomarker for cancer [45–47]. The role of *miR-6797-5p* and *miR-762* remain unexplored in GBC until now. As a result, these hub miRNAs can be considered novel findings and are being reported for the first time in this study.

In the GBC+GS group, LINC00852 indirectly targets the hub DEmRNAs- *STAT3* and *SMAD4* by interacting with multiple miRNAs. Both *SMAD4* and *STAT3* act as an important signaling molecule in FOXO signaling pathways. FOXO signaling is a crucial pathway that influences important regulatory pathways including cell cycle control, apoptosis, and glucose metabolism [54]. Various growth factors, including insulin, are crucial for FOXO regulation across different cancers [**Figure 5.11 A**]. The primary pathways interacting with FOXO signaling in cancer include the PI3K/AKT pathway, Ras-MEK-ERK, IKK, and AMPK pathway [54,55]. Therefore; targeting FOXO signaling pathways therapeutically holds promise for discovering and developing effective molecules against aggressive carcinomas including GBC [48].

In the GBC group, both the hub lncRNAs- *DIO3OS* and *MSTRG.16633* targets the DEmRNA (*LMOD1*, *NCAM1*, *MYL9*) integral to cell adhesion processes—showed substantial downregulation in GBC as compared to the control. Cell adhesion molecules (CAMs) are essential in cell-to-cell and cell-to-ECM interactions, impacting malignant cell adhesion properties crucial for cancer development and progression [56]. Modifications in CAM expression or function influence not only cell adhesion but also signaling pathways involved in tumor invasion and metastasis [57]. Dysregulation in cell-adhesion pathways also regulates classic signal transduction pathways, thereby, contributing to the metastatic behaviour of cancer cells [58]. The identified lncRNAs, *DIO3OS* and *MSTRG.16633*, are significant players in the epithelial-mesenchymal-transition (EMT) process as they target key genes involved in cell-adhesion processes [**Figure 5.11 B**]. This could potentially explain the rapid progression of GBC in patients without gallstones.

(A)



(B)

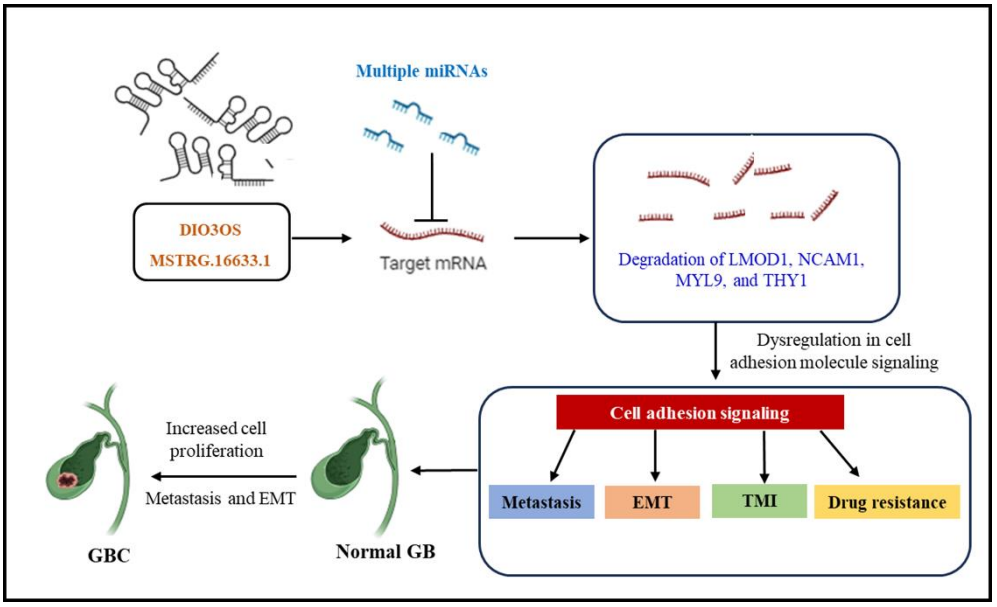


Figure 5.11: Mechanistic illustration of the action of hub DElncRNA and DEnlncRNA in GBC and GBC+GS pathogenesis. (A) Dysregulated lncRNAs (*LINC00852* and *MSTRG.53675*) identified in the gallstone-associated GBC group show that the hub lncRNAs were found to interact with potential hub DEmRNAs involved in FoxO signaling pathways which induce pro-tumorigenic effects by modulating crucial regulatory pathways and promote tumorigenesis in gallbladder with gallstones. **(B)** In the GBC group, the lncRNAs (*DIO3OS* and *MSTRG.16633*) identified indirectly interact with the cell-adhesion molecules and contribute to tumor invasion through EMT and tumor-microenvironment interaction, which ultimately leads to increased cell proliferation and metastatic gallbladder carcinogenesis.

5.4 Summary

In the previous chapter, it was observed that the DEGs identified in each case studies are highly downregulated and involved in signal transduction, cell-adhesion and immune signaling pathways. In case case study 4, the DEGs identified in two GBC groups- GBC and GBC+GS were found to be associated with distinct and diverse pathlogical pathways. The potential DEGs identified in GBC group are related to cell adhesion signaling whereas, cell cycle and immune response pathways are the predominantly associated with the hub DEGs identified in GBC+GS group. To investigate the coding-noncoding interactome, integrative regulatory analysis of the known and novel lncRNA has been performed in GBC and GBC+GS group. This chapter presents the identification of unique regulatory lncRNAs and target mRNAs in GBC and GBC+GS groups using integrated gene network-based approaches. This is the first study that

has reported the association of lncRNA- *DIO3OS* and *LINC00852* and *miR-762* and *miR-6797-5p* in GBC pathogenesis. Furthermore, the novel lncRNA sequences and their secondary structure have also been identified in GBC and GBC+GS groups independently. Therefore, this study has significantly advanced our molecular understanding of ncRNA involvement in GBC pathogenesis. The identified lncRNAs and miRNAs may serve crucial functions as diagnostic, prognostic, and therapeutic markers in GBC, offering promising avenues for further exploration and functional validation.

Bibliography

[1] Spizzo, R., Almeida, M. I., Colombatti, A., & Calin, G. A. Long non-coding RNAs and cancer: a new frontier of translational research?. *Oncogene*, 31(43): 4577-4587, 2012.

[2] Rinn, J. L., & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annual review of biochemistry*, 81: 145-166, 2012.

[3] Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et.al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome research*, 22(9): 1775-1789, 2012.

[4] Mercer, T. R., Dinger, M. E., Sunkin, S. M., Mehler, M. F., & Mattick, J. S. Specific expression of long noncoding RNAs in the mouse brain. *Proceedings of the National Academy of Sciences*, 105(2): 716-721, 2008.

[5] Khandelwal, A., Malhotra, A., Jain, M., Vasquez, K. M., & Jain, A. The emerging role of long non-coding RNA in gallbladder cancer pathogenesis. *Biochimie*, 132: 152-160, 2017.

[6] Peng, W. X., Koirala, P., & Mo, Y. Y. LncRNA-mediated regulation of cell signaling in cancer. *Oncogene*, 36(41): 5661-5667, 2017

[7] Khorkova, O., Hsiao, J., & Wahlestedt, C. Basic biology and therapeutic implications of lncRNA. *Advanced drug delivery reviews*, 87, 15-24, 2015.

[8] Schmitt, A. M., Garcia, J. T., Hung, T., Flynn, R. A., Shen, Y., Qu, K., et.al. An inducible long noncoding RNA amplifies DNA damage signaling. *Nature genetics*, 48(11): 1370-1376, 2016.

[9] Jariwala, N., & Sarkar, D. Emerging role of lncRNA in cancer: a potential avenue in molecular medicine. *Annals of translational medicine*, 4(15), 2016.

[10] Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafreniere, R. G., Xing, Y., Lawrence, J., & Willard, H. F. The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell*, 71(3): 527-542, 1992.

[11] Feil, R., Walter, J., Allen, N. D., & Reik, W. Developmental control of allelic methylation in the imprinted mouse Igf2 and H19 genes. *Development*, 120(10): 2933-2943, 1994.

[12] Liu, B., Shen, E. D., Liao, M. M., Hu, Y. B., Wu, K., Yang, P., et.al. Expression and mechanisms of long non-coding RNA genes MEG3 and ANRIL in gallbladder cancer. *Tumor Biology*, 37: 9875-9886, 2016.

[13] Wang, S. H., Wu, X. C., Zhang, M. D., Weng, M. Z., Zhou, D., & Quan, Z. W. Long noncoding RNA H19 contributes to gallbladder cancer cell proliferation by modulated miR-194-5p targeting AKT2. *Tumor Biology*, 37: 9721-9730, 2016.

[14] Wu, X. S., Wang, X. A., Wu, W. G., Hu, Y. P., Li, M. L., Ding, Q., et.al. MALAT1 promotes the proliferation and metastasis of gallbladder cancer cells by activating the ERK/MAPK pathway. *Cancer biology & therapy*, 15(6): 806-814, 2014.

[15] Ma, M. Z., Li, C. X., Zhang, Y., Weng, M. Z., Zhang, M. D., Qin, Y. Y., et.al. Long non-coding RNA HOTAIR, a c-Myc activated driver of malignancy, negatively regulates miRNA-130a in gallbladder cancer. *Molecular cancer*, 13: 1-14, 2014.

[16] Ma, M. Z., Kong, X., Weng, M. Z., Zhang, M. D., Qin, Y. Y., Gong, W., et.al. Long non-coding RNA-LET is a positive prognostic factor and exhibits tumor-suppressive activity in gallbladder cancer. *Molecular carcinogenesis*, 54(11): 1397-1406, 2015.

[17] Ma, M. Z., Chu, B. F., Zhang, Y., Weng, M. Z., Qin, Y. Y., Gong, W., & Quan, Z. W. Long non-coding RNA CCAT1 promotes gallbladder cancer development via negative modulation of miRNA-218-5p. *Cell death & disease*, 6(1): e1583-e1583, 2015.

[18] Prensner, J. R., & Chinnaiyan, A. M. The emergence of lncRNAs in cancer biology. *Cancer discovery*, 1(5): 391-407, 2011.

[19] Prensner, J. R., Iyer, M. K., Balbin, O. A., Dhanasekaran, S. M., Cao, Q., Brenner, J. C., et.al. Transcriptome sequencing identifies PCAT-1, a novel lincRNA implicated in prostate cancer progression. *Nature biotechnology*, 29(8): 742, 2011.

[20] Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., et.al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. nature, 464(7291): 1071-1076, 2010.

[21] Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., & Shen, F. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. I*nternational journal of molecular sciences*, 20(22): 5573, 2019.

[22] Ponting, C. P., Oliver, P. L., & Reik, W. Evolution and functions of long noncoding RNAs. *Cell*, 136(4): 629-641, 2009.

[23] Wang, K. C., & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. *Molecular cell*, 43(6): 904-914, 2011.

[24] Huarte, M. The emerging role of lncRNAs in cancer. *Nature medicine*, 21(11): 1253-1261, 2015.

[25] Peng, F., Wang, R., Zhang, Y., Zhao, Z., Zhou, W., Chang, Z., et.al. Differential expression analysis at the individual level reveals a lncRNA prognostic signature for lung adenocarcinoma. *Molecular cancer*, 16(1): 1-12, 2017.

[26] Xu, Y., Zhang, X., Hu, X., Zhou, W., Zhang, P., Zhang, J., et.al. The effects of lncRNA MALAT1 on proliferation, invasion, and migration in colorectal cancer through regulating SOX9. *Molecular medicine*, 24: 1-15, 2018.

[27] Ahadi, A. Functional roles of lncRNAs in the pathogenesis and progression of cancer. *Genes & Diseases*, 8(4): 424-437, 2021.

[28] Mercer, T. R., Dinger, M. E., & Mattick, J. S. Long non-coding RNAs: insights into functions. *Nature reviews genetics*, 10(3): 155-159, 2009.

[29] Tulsyan, S., Hussain, S., Mittal, B., Saluja, S. S., Tanwar, P., Rath, G. K., et.al. A systematic review with in silico analysis on transcriptomic profile of gallbladder carcinoma. *In Seminars in Oncology*, 47(6), 398-408. WB Saunders, 2020.

[30] Park, S. K., Andreotti, G., Sakoda, L. C., Gao, Y. T., Rashid, A., Chen, J., et.al. Variants in hormone-related genes and the risk of biliary tract cancers and stones: a population-based study in China. *Carcinogenesis*, 30(4): 606-614, 2009.

[31] Pilgrim, C. H., Groeschl, R. T., Christians, K. K., & Gamblin, T. C. Modern perspectives on factors predisposing to the development of gallbladder cancer. *Hpb*, 15(11): 839-844, 2013.

[32] Hernandez, A., Martinez, M. E., Croteau, W., & Germain, D. L. S. Complex organization and structure of sense and antisense transcripts expressed from the DIO3 gene imprinted locus. *Genomics*, 83(3): 413-424, 2004.

[33] Hernandez, A. Structure and function of the type 3 deiodinase gene. *Thyroid*, 15(8): 865-874, 2005.

[34] Wang, Y., Wang, J., Wang, C., Chen, Y., & Chen, J. DIO3OS as a potential biomarker of papillary thyroid cancer. *Pathology-Research and Practice*, 229, 153695, 2022.

[35] Wang, Y., Sun, P., Hao, X., Cao, D., Liu, J., & Zhang, D. Decreased DIO3OS expression predicts poor prognosis in hepatocellular carcinoma and is associated with immune infiltration. *Biochemical Genetics*, 1-16, 2023.

[36] Cui, K., Jin, S., Du, Y., Yu, J., Feng, H., Fan, Q., & Ma, W. Long noncoding RNA DIO3OS interacts with miR-122 to promote proliferation and invasion of pancreatic cancer cells through upregulating ALDOA. *Cancer Cell International*, 19(1): 1-10, 2019.

[37] Liu, P., Wang, H., Liang, Y., Hu, A., Xing, R., Jiang, L., et.al. LINC00852 promotes lung adenocarcinoma spinal metastasis by targeting S100A9. *Journal of Cancer*, 9(22): 4139, 2018.

[38] Li, Q., Wang, X., Jiang, N., Xie, X., Liu, N., Liu, J., et.al. Exosome-transmitted linc00852 associated with receptor tyrosine kinase AXL dysregulates the proliferation and invasion of osteosarcoma. *Cancer Medicine*, 9(17): 6354-6366, 2020.

[39] Lu, Y., Meng, Q., Bai, L., Wang, R., Sun, Y., Li, J., et.al. LINC00858 stabilizes RAN expression and promotes metastasis of gastric cancer. *Biology Direct*, 17(1): 1-15, 2022.

[40] Qiao, Z. W., Jiang, Y., Wang, L., Wang, L., Jiang, J., Zhang, J. R., & Mu, P. LINC00852 promotes the proliferation and invasion of ovarian cancer cells by competitively binding with miR-140-3p to regulate AGTR1 expression. *BMC cancer*, 21(1): 1-14, 2021.

[41] Zhang, H., Du, Y., Xin, P., & Man, X. The LINC00852/miR-29a-3p/JARID2 axis regulates the proliferation and invasion of prostate cancer cell. *BMC cancer*, 22(1): 1269, 2022.

[42] Zheng, D., Tahir, R. A., Yan, Y., Zhao, J., Quan, Z., Kang, G., et.al. Screening of Human Circular RNAs as Biomarkers for Early Onset Detection of Alzheimer's Disease. *Frontiers in Neuroscience*, 16, 878287, 2022.

[43] Zhou, L., Wang, Q., Hou, J., Wu, X., Wang, L., & Chen, X. Upregulation of hsa_circ_0002003 promotes hepatocellular carcinoma progression. *BMC cancer*, 23(1): 611, 2023.

[44] Li, Q., Geng, S., Yuan, H., Li, Y., Zhang, S., Pu, L., et.al. Circular RNA expression profiles in extracellular vesicles from the plasma of patients with pancreatic ductal adenocarcinoma. *FEBS Open Bio*, 9(12): 2052-2062, 2019.

[45] Li, Y., Huang, R., Wang, L., Hao, J., Zhang, Q., Ling, R., & Yun, J. microRNA-762 promotes breast cancer cell proliferation and invasion by targeting IRF 7 expression. *Cell proliferation*, 48(6): 643-649, 2015.

[46] Lai, P. S., Chang, W. M., Chen, Y. Y., Lin, Y. F., Liao, H. F., & Chen, C. Y. Circulating microRNA-762 upregulation in colorectal cancer may be accompanied by Wnt-1/β-catenin signaling. *Cancer Biomarkers*, 32(2): 111-122, 2021.

[47] Chen, L., Li, Y., & Lu, J. Identification of circulating miR-762 as a novel diagnostic and prognostic biomarker for non-small cell lung cancer. *Technology in cancer research & treatment*, 19, 1533033820964222, 2020.

[48] Wen, L., & Han, Z. Identification and validation of xenobiotic metabolism-associated prognostic signature based on five genes to evaluate immune microenvironment in colon cancer. *Journal of Gastrointestinal Oncology*, 12(6): 2788, 2021.

[49] Williams, J. A., & Phillips, D. H. Mammary expression of xenobiotic metabolizing enzymes and their potential role in breast cancer. *Cancer research*, 60(17): 4667-4677, 2000.

[50] Tamási, V., Monostory, K., Prough, R. A., & Falus, A. Role of xenobiotic metabolism in cancer: involvement of transcriptional and miRNA regulation of P450s. *Cellular and Molecular Life Sciences*, 68: 1131-1146, 2011.

[51] Dixit, R., Pandey, M., Rajput, M., & Shukla, V. K. Unravelling of the comparative transcriptomic profile of gallbladder cancer using mRNA sequencing. *Molecular Biology Reports*, 49(7): 6395-6403, 2022

[52] Zuo, M., Rashid, A., Wang, Y., Jain, A., Li, D., Behari, A. et.al. RNA sequencing-based analysis of gallbladder cancer reveals the importance of the liver X receptor and lipid metabolism in gallbladder cancer. *Oncotarget*, 7(23): 35302, 2016.

[53] Ebata, N., Fujita, M., Sasagawa, S., Maejima, K., Okawa, Y., Hatanaka, Y., et.al. Molecular classification and tumor microenvironment characterization of gallbladder cancer by comprehensive genomic and transcriptomic analysis. *Cancers*, 13(4): 733, 2021.

[54] Farhan, M., Wang, H., Gaur, U., Little, P. J., Xu, J., et.al. FOXO signaling pathways as therapeutic targets in cancer. *International journal of biological sciences*, 13(7): 815, 2017.

[55] Ho, K. K., Myatt, S. S., & Lam, E. W. Many forks in the path: cycling with FoxO. *Oncogene*, 27(16): 2300-2311, 2008.

[56] Myatt, S. S., & Lam, E. W. F. The emerging roles of forkhead box (Fox) proteins in cancer. *Nature Reviews Cancer*, 7(11): 847-859, 2007.

[57] Makrilia, N., Kollias, A., Manolopoulos, L., & Syrigos, K. Cell adhesion molecules: role and clinical significance in cancer. *Cancer investigation*, 27(10): 1023-1037, 2009.

[58] Cavallaro, U. G. O., & Christofori, G. Multitasking in tumor progression: signaling functions of cell adhesion molecules. *Annals of the New York Academy of Sciences*, 1014(1): 58-66, 2004.